Anlage 1

Fragebogen zur Datenerhebung

Nr. ______

Liebe Teilnehmerin, lieber Teilnehmer,

herzlichen Dank, dass Sie sich bereit erklärt haben, an dieser Studie teilzunehmen.

Im Rahmen unserer Forschung möchten wir Sie bitten, sich ein paar Minuten Zeit zu nehmen, um einen **Produkttest** zu machen: einen Wallwasher testen. Von besonderem Interesse ist es für uns, wie Sie die Anwendbarkeit des Wallwashers für zu Hause beurteilen. Um die **Wartezeit zu überbrücken**, würden wir Sie zudem bitten, einige Aufgaben zu lösen und einige Fragen zu beantworten.

Anschließend folgen vier Fragen, die sich auf den Wallwasher beziehen.

Achten Sie bitte auf die Anweisungen des Versuchsleiters.

Wir möchten Sie darauf hinweisen, dass alle Antworten **anonym** erhoben werden und nur für wissenschaftliche Zwecke verwendet werden. Ihre Teilnahme ist freiwillig und kann jederzeit abgebrochen werden.
Um etwaige Fragen während der Befragung zu klären, stellen Sie diese bitte sofort, bevor Sie mit der Bearbeitung starten.

Bitte konzentrieren Sie sich ausschließlich auf die Aufgabe, da die Ergebnisse dieser Studie für die weitere Forschung ausschlaggebend sein werden.

Blättern Sie nun bitte um

Als erstes beantworten Sie bitte folgende Frage:

Wie geht es Ihnen **im Moment**?

sehr schlecht ... sehr gut

1	2	3	4	5	6	7	8	9

Ich fühle mich im Augenblick …

		gar nicht	etwas	einiger-maßen	ziemlich	über-wiegend	völlig
1.	fröhlich	☐	☐	☐	☐	☐	☐
2.	vergnügt	☐	☐	☐	☐	☐	☐
3.	gutgelaunt	☐	☐	☐	☐	☐	☐
4.	heiter	☐	☐	☐	☐	☐	☐

		gar nicht	etwas	einiger-maßen	ziemlich	über-wiegend	völlig
5.	ruhig	☐	☐	☐	☐	☐	☐
6.	nervös	☐	☐	☐	☐	☐	☐
7.	gelassen	☐	☐	☐	☐	☐	☐
8.	ausgeglichen	☐	☐	☐	☐	☐	☐
9.	locker	☐	☐	☐	☐	☐	☐

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

1 Aufgabe zur visuellen Verarbeitung

Nun folgen 4 Aufgaben zur visuellen Wahrnehmung mit Zeitbegrenzung. Dies wird folgendermaßen umgesetzt:
Wenn der Versuchsleiter Sie dazu auffordert, blättern Sie bitte um und verbinden die Punkte im nachfolgenden Bild so schnell und so genau wie möglich. Vermeiden Sie es, Punkte auszulassen und versuchen Sie, innerhalb der vorgegebenen Zeit so viel wie möglich zu schaffen.

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

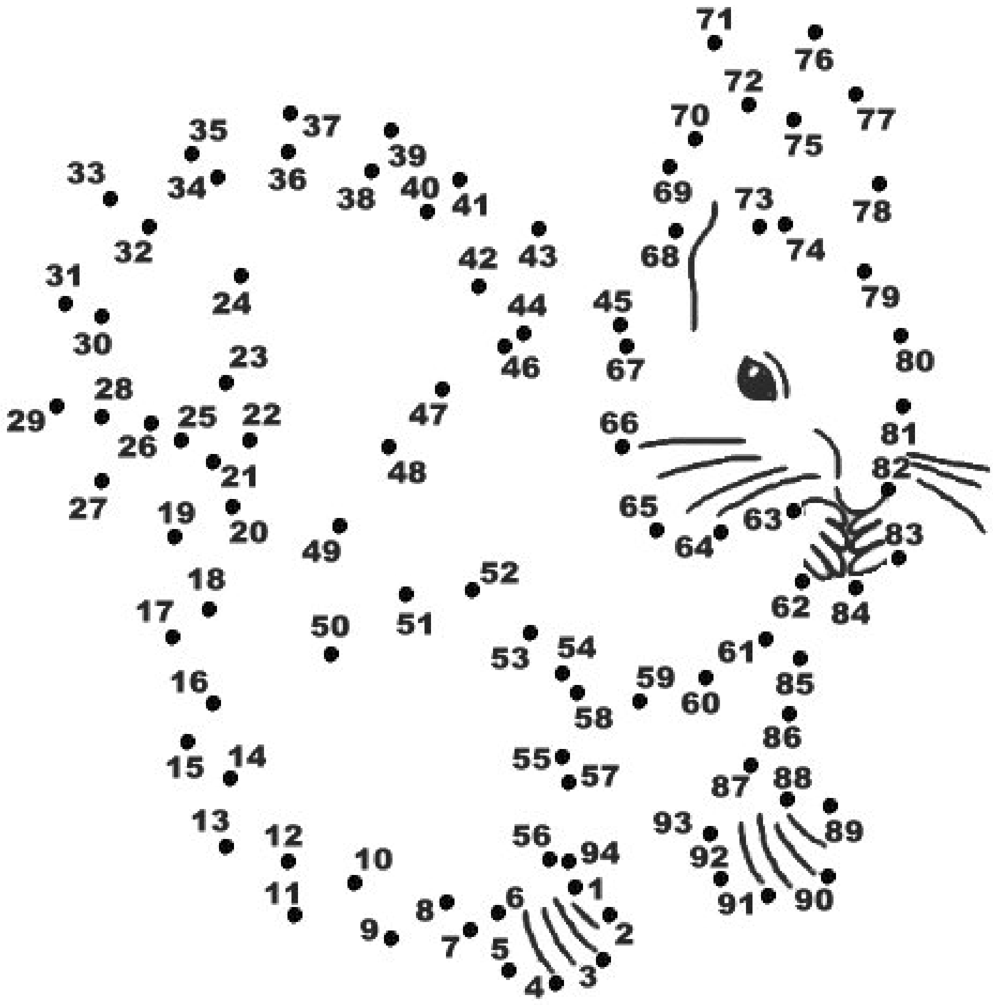
1.1. „Zahlenbild Esel“

Bitte denken Sie daran, so schnell und so genau wie möglich zu arbeiten.

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

1.2. Zahlenbild „Eichhörnchen“

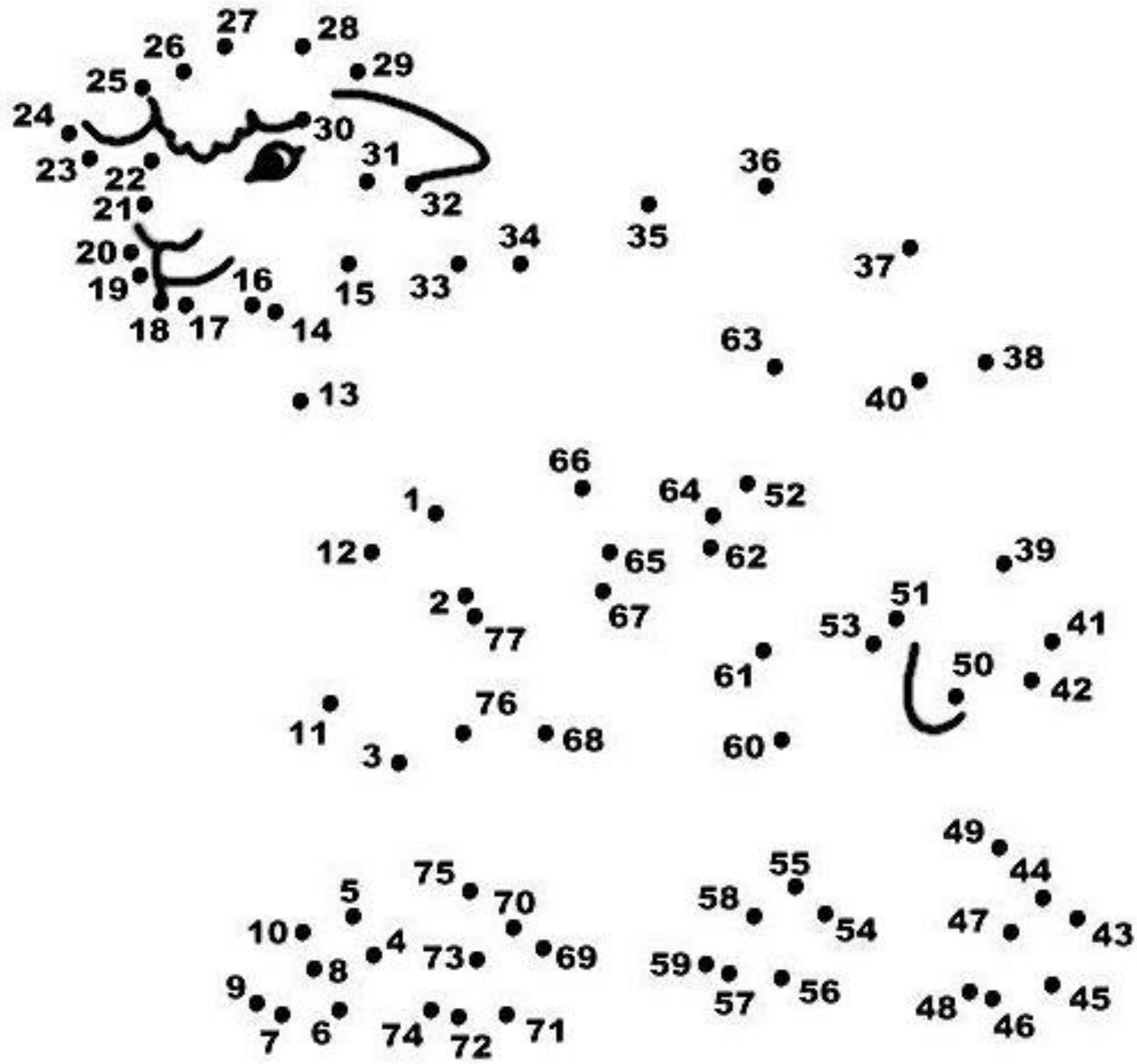
Bitte denken Sie daran, so schnell und so genau wie möglich zu arbeiten.



Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

1.3. Zahlenbild „Lamm“

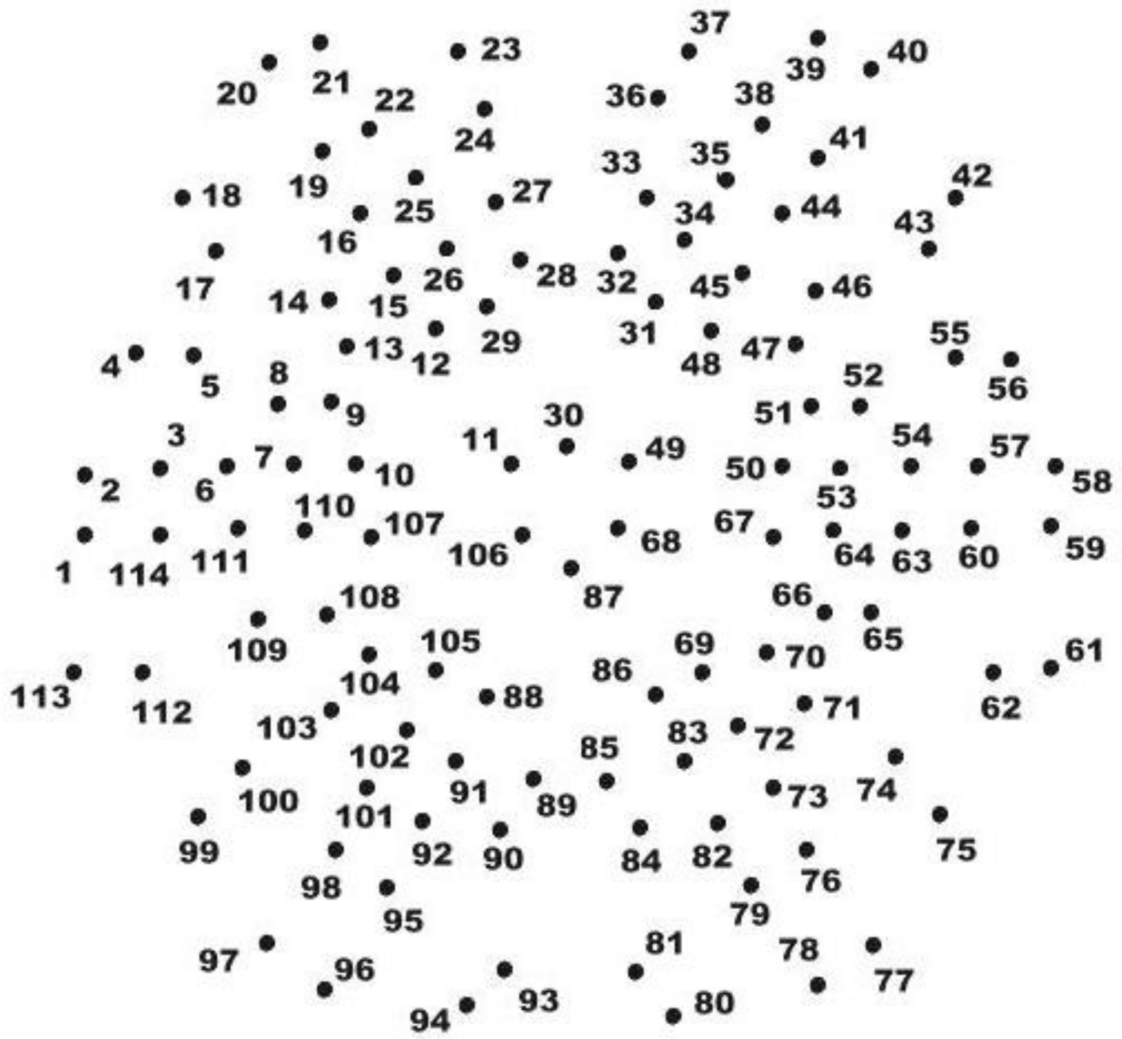
Bitte denken Sie daran, so schnell und so genau wie möglich zu arbeiten.



Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

1.4. Zahlenbild „Schneekristall“

Bitte denken Sie daran, so schnell und so genau wie möglich zu arbeiten.



Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

Diese Fragen beziehen sich auf die Zahlenverbindungsaufgaben.

1. Wie schwierig waren die Aufgaben für Sie?

sehr leicht sehr schwer

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

2. Wie viel Spaß haben Ihnen die Aufgaben gemacht?

gar nicht sehr viel

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

3. Wie motiviert waren Sie, die Aufgaben zu lösen?

überhaupt nicht motiviert sehr motiviert

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

2. Nutzung einer Getränkedose

Überlegen Sie sich einmal, was man alles mit einer **leeren Getränkedose** machen kann. Produzieren Sie so viele Ideen, wie möglich. Denken Sie an praktische Dinge, an Freizeitbetätigungen, an künstlerische Verwendungsbereiche – an einfach alles, was Ihnen einfällt. Es gibt hier **keine "richtige"** Lösung.

Für alle Teilnehmenden ist die Zeit auf **zwei Minuten** (inkl. lesen dieser Einleitung) festgelegt.

Bitte hören Sie auf zu schreiben, sobald der Versuchsleiter **`STOPP`** gesagt hat.

Schreiben Sie in diesem Feld Ihre Ideen auf!

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

Diese Fragen beziehen sich auf die Aufgabe über die Nutzung einer Getränkedose.

1. Wie schwierig war die Aufgabe für Sie?

sehr leicht — sehr schwer

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

2. Wie viel Spaß hat Ihnen die Aufgabe gemacht?

gar nicht — sehr viel

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

3. Wie motiviert waren Sie, die Aufgabe zu lösen?

überhaupt nicht motiviert — sehr motiviert

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

2 Außerirdisches Leben

Nun, die letzte Aufgabe: Wir interessieren uns auch dafür, wie sich Menschen **außerirdisches Leben** vorstellen. Stellen Sie sich dazu vor, dass Sie in eine andere Galaxie in unserem Universum reisen und dort einen Planeten vorfinden, der sich erheblich von der Erde unterscheidet. Auf Ihrer Reise stoßen Sie dabei auf eine Kreatur, die Sie in Ihrer Vorstellung antreffen. Es geht hierbei **nicht um Ihr Zeichentalent**, sondern um die Frage, wie Sie sich außerirdisches Leben **vorstellen**, um die Unterschiede in den Vorstellungen über Außerirdische zu untersuchen. Damit alle Teilnehmer gleich viel Zeit aufwenden, ist die Zeichenzeit auf **drei Minuten** (inkl. dieser Einleitung) festgelegt.

Zeichnen Sie Ihre Kreatur hier:

Blättern Sie um, wenn der Versuchsleiter Sie dazu auffordert.

Diese Fragen beziehen sich auf die Aufgabe über das außerirdische Leben.

1. Wie schwierig war die Aufgabe für Sie?

sehr leicht — sehr schwer

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

2. Wie viel Spaß hat Ihnen die Aufgabe gemacht?

gar nicht — sehr viel

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

3. Wie motiviert waren Sie, die Aufgabe zu lösen?

überhaupt nicht motiviert — sehr motiviert

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

Blättern Sie nun bitte um.

Aufgabe zur Produktbewertung

Wir würden Sie nun bitten, den Wallwasher hinsichtlich nachfolgender Dimensionen zu bewerten.

---

1. Wie interessiert wären Sie, wenn das Modell im Vorbeigehen in einem Geschäft entdecken würden?

	1	2	3	4	5	6	7	
gar nicht interessiert								sehr interessiert

---

2. Wie gut könnten Sie sich vorstellen, den Wallwasher zu Hause zu nutzen?

	1	2	3	4	5	6	7	
gar nicht gut								sehr gut

---

3. Wie praktisch hinsichtlich der Anwendbarkeit zu Hause schätzen Sie den Wallwasher insgesamt ein?

	1	2	3	4	5	6	7	
gar nicht praktisch								sehr praktisch

---

4. Wie schätzen Sie die Qualität des Wallwashers ein?

	1	2	3	4	5	6	7	
sehr schlecht								sehr gut

Blättern Sie nun bitte um.

Bewertung des Raumes

Wir würden Sie nun bitten, die Atmosphäre im Raum zu bewerten, in dem Sie sich gerade befinden.

Dieser Raum ist…

	ganz wenig oder gar nicht	etwas	einiger-maßen	ziemlich	über-wiegend	völlig
1. fröhlich	☐	☐	☐	☐	☐	☐
2. inspirierend	☐	☐	☐	☐	☐	☐
3. lebendig	☐	☐	☐	☐	☐	☐
4. stimulierend	☐	☐	☐	☐	☐	☐
5. formell	☐	☐	☐	☐	☐	☐
6. kühl	☐	☐	☐	☐	☐	☐
7. kalt	☐	☐	☐	☐	☐	☐
8. sachlich	☐	☐	☐	☐	☐	☐
9. intim	☐	☐	☐	☐	☐	☐
10. behaglich	☐	☐	☐	☐	☐	☐
11. geborgen	☐	☐	☐	☐	☐	☐
12. kuschelig	☐	☐	☐	☐	☐	☐
13. gemütlich	☐	☐	☐	☐	☐	☐
14. ungemütlich	☐	☐	☐	☐	☐	☐
15. angespannt	☐	☐	☐	☐	☐	☐
16. beängstigend	☐	☐	☐	☐	☐	☐
17. bedrohlich	☐	☐	☐	☐	☐	☐
18. beklemmend	☐	☐	☐	☐	☐	☐

Und ein Paar Abschlussfragen:

Wie geht es Ihnen **im Moment**?

sehr schlecht | sehr gut

1	2	3	4	5	6	7	8	9

Ich fühle mich im Augenblick …

	gar nicht	etwas	einiger-maßen	ziemlich	über-wiegend	völlig
1. fröhlich	☐	☐	☐	☐	☐	☐
2. vergnügt	☐	☐	☐	☐	☐	☐
3. gutgelaunt	☐	☐	☐	☐	☐	☐
4. heiter	☐	☐	☐	☐	☐	☐

	gar nicht	etwas	einiger-maßen	ziemlich	über-wiegend	völlig
5. ruhig	☐	☐	☐	☐	☐	☐
6. nervös	☐	☐	☐	☐	☐	☐
7. gelassen	☐	☐	☐	☐	☐	☐
8. ausgeglichen	☐	☐	☐	☐	☐	☐
9. locker	☐	☐	☐	☐	☐	☐

Was ist Ihrer Meinung nach der **genaue Zweck** der Untersuchung (kurz und stichwortartig)?

____

____

Haben Sie Bemerkungen oder Kritikpunkte?

____

____

Wurde bei Ihnen eine Farbfehlsichtigkeit festgestellt?
Ja ☐ Nein ☐

Nun benötigen wir noch einige **demographische Daten** von Ihnen.

Alter: _____ Jahre Geschlecht: w ☐ m ☐

Beschäftigungsstatus:
Student ☐ Berufstätig ☐ anderes ☐

Wenn Student:
Studiengang: __________________ Fachsemester: _________________

Deutschkenntnisse: sehr gut oder Muttersprache ☐ nicht gut ☐

Herzlichen Dank für Ihre Teilnahme!
Wenden Sie sich an den Versuchsleiter, um Ihre Belohnung zu bekommen!

Anlage 2

Tabellen zur Ergebnisausgabe

Gruppenstatistiken					
	Lichtbedingung	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Wie schwierig war die Aufgabe für Sie? (UU)	weiß	49	5,49	2,607	0,372
	blau	51	4,71	2,46	0,344
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (UU)	weiß	49	4,47	2,042	0,292
	blau	51	6,02	2,093	0,293
Wie motiviert waren Sie, die Aufgabe zu lösen? (UU)	weiß	49	5,33	2,174	0,311
	blau	51	5,88	1,774	0,248
Wie schwierig war diese Aufgabe für Sie? (Alien)	weiß	49	4,98	1,995	0,285
	blau	51	4,94	1,782	0,25
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (Alien)	weiß	49	4,61	1,579	0,226
	blau	51	5,12	1,829	0,256
Wie motiviert waren Sie, die Aufgabe zu lösen? (Alien)	weiß	49	5,94	1,886	0,269
	blau	51	6,02	1,772	0,248
Stimmung nachher (t2)	weiß	49	6,9	1,246	0,178
	blau	51	6,53	1,567	0,219
Stimmung vorher (t1)	weiß	49	6,96	1,29	0,184
	blau	51	5,18	11,803	1,653
Anzahl aller Nennungen (UU)	weiß	49	6,27	2,099	0,3
	blau	51	6,9	2,663	0,373
Anzahl kreativer Nennungen (Quantität UU)	weiß	49	2,02	1,942	0,277
	blau	51	2,69	2,015	0,282
Prozentualer Anteil kreativer Nennungen (UU) in %	weiß	49	30,73469388	23,5953169	3,370759557
	blau	51	39,85490196	24,17589968	3,385303084
Qualitätsmaß (UU)	weiß	49	4,063061224	0,997973372	0,142567625
	blau	51	4,465882353	1,038441479	0,145410892
Originalität des Aliens	weiß	49	3,067959184	1,055373116	0,150767588
	blau	51	3,470588235	1,240408661	0,173691954
Ähnlichkeit des Aliens	weiß	49	2,837142857	1,375283304	0,196469043
	blau	51	2,549019608	1,374457355	0,192462526
Kreative Leistung	weiß	49	-0,1989	0,69285	0,09898
	blau	51	0,1388	0,65118	0,09118

Anmerkungen:
(UU): "Unusual Uses"-Aufgabe
(Alien): Imaginationsaufgabe

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit			T-Test für die Mittelwertgleichheit			
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
Wie schwierig war die Aufgabe für Sie? (UU)	Varianzen sind gleich	0,615	0,435	1,547	98	0,125	0,784	0,507	-0,222	1,789
	Varianzen sind nicht gleich			1,545	97,062	0,126	0,784	0,507	-0,223	1,791
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (UU)	Varianzen sind gleich	0	0,996	-3,747	98	0	-1,55	0,414	-2,371	-0,729
	Varianzen sind nicht gleich			-3,749	97,975	0	-1,55	0,414	-2,371	-0,73
Wie motiviert waren Sie, die Aufgabe zu lösen? (UU)	Varianzen sind gleich	3,692	0,058	-1,404	98	0,164	-0,556	0,396	-1,342	0,23
	Varianzen sind nicht gleich			-1,398	92,656	0,165	-0,556	0,398	-1,345	0,234
Wie schwierig war diese Aufgabe für Sie? (Alien)	Varianzen sind gleich	2,574	0,112	0,102	98	0,919	0,038	0,378	-0,712	0,788
	Varianzen sind nicht gleich			0,101	95,779	0,919	0,038	0,379	-0,714	0,79
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (Alien)	Varianzen sind gleich	0,904	0,344	-1,477	98	0,143	-0,505	0,342	-1,185	0,174
	Varianzen sind nicht gleich			-1,481	96,907	0,142	-0,505	0,341	-1,183	0,172
Wie motiviert waren Sie, die Aufgabe zu lösen? (Alien)	Varianzen sind gleich	0,017	0,898	-0,221	98	0,826	-0,081	0,366	-0,807	0,645
	Varianzen sind nicht gleich			-0,221	96,973	0,826	-0,081	0,366	-0,808	0,646
Stimmung nachher (t2)	Varianzen sind gleich	1,623	0,206	1,299	98	0,197	0,369	0,284	-0,195	0,932
	Varianzen sind nicht gleich			1,305	94,73	0,195	0,369	0,282	-0,192	0,929
Stimmung vorher (t1)	Varianzen sind gleich	2,763	0,1	1,051	98	0,296	1,783	1,696	-1,583	5,149
	Varianzen sind nicht gleich			1,072	51,243	0,289	1,783	1,663	-1,555	5,121

Anzahl aller Nennungen (UU)	Varianzen sind gleich	2,607	0,11	-1,324	98	0,189	-0,637	0,481	-1,591	0,317
	Varianzen sind nicht gleich			-1,331	94,446	0,187	-0,637	0,479	-1,587	0,313
Anzahl kreativer Nennungen (Quantität UU)	Varianzen sind gleich	0,004	0,95	-1,682	98	0,096	-0,666	0,396	-1,452	0,12
	Varianzen sind nicht gleich			-1,683	97,999	0,096	-0,666	0,396	-1,451	0,119
Prozentualer Anteil kreativer Nennungen (UU) in %	Varianzen sind gleich	0,161	0,689	-1,908	98	0,059	-9,120208083	4,779615095	-18,60519868	0,364782518
	Varianzen sind nicht gleich			-1,909	97,975	0,059	-9,120208083	4,777268776	-18,60057323	0,360157063
Qualitätsmaß (UU)	Varianzen sind gleich	0,007	0,932	-1,977	98	0,051	-0,402821128	0,203805011	-0,807265539	0,001623282
	Varianzen sind nicht gleich			-1,978	98	0,051	-0,402821128	0,203641487	-0,806941034	0,001298777
Originalität Alien	Varianzen sind gleich	0,109	0,742	-1,745	98	0,084	-0,402629052	0,230744895	-0,860534787	0,055276684
	Varianzen sind nicht gleich			-1,751	96,604	0,083	-0,402629052	0,22999948	-0,859138001	0,053879897
Ähnlichkeit Alien	Varianzen sind gleich	1,549	0,216	1,048	98	0,297	0,288123249	0,275027404	-0,257659682	0,833906181
	Varianzen sind nicht gleich			1,048	97,835	0,297	0,288123249	0,275030742	-0,257677784	0,833924282
Kreative Leistung	Varianzen sind gleich	0,018	0,894	-2,512	98	0,014	-0,33767	0,13441	-0,60441	-0,07094
	Varianzen sind nicht gleich			-2,509	96,985	0,014	-0,33767	0,13458	-0,60477	-0,07058

Anmerkungen:
(UU): “Unusual Uses“-Aufgabe
(Alien): Imaginationsaufgabe

Korrelationen																		
		Alter	Geschlecht	Wie schwierig war die Aufgabe für Sie? (UU)	Wieviel Spaß hat Ihnen die Aufgabe gemacht? (UU)	Wie motiviert waren Sie, die Aufgabe zu lösen? (UU)	Wie schwierig war diese Aufgabe für Sie? (Alien)	Wieviel Spaß hat Ihnen die Aufgabe gemacht? (Alien)	Wie motiviert waren Sie, die Aufgabe zu lösen? (Alien)	Stimmung nachher (t2)	Stimmung vorher (t1)	Anzahl Nennungen (UU)	Anzahl kreativer Nennungen (UU)	Anteil kreativer Nennungen (UU) in %	Qualitätsmaß (UU)	Originalität Alien	Ähnlichkeit Alien	Kreative Leistung
Alter	r	1	0,001	0,063	-0,129	0,086	-0,007	0,132	0,153	0,081	-,198*	-0,077	0,066	0,112	0,085	-0,102	0,107	0,002
	p		0,988	0,533	0,202	0,397	0,944	0,19	0,129	0,423	0,048	0,447	0,512	0,267	0,399	0,314	0,291	0,988
	N	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Geschlecht	r		1	0,02	0,195	0,19	-0,037	0,131	0,108	0,143	-0,043	0,103	0,139	0,037	0,056	0,023	0,072	0,055
	p			0,84	0,052	0,059	0,714	0,192	0,283	0,157	0,67	0,306	0,166	0,712	0,577	0,818	0,479	0,587
	N		100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Wie schwierig war die Aufgabe für Sie? (UU)	r			1	-,400**	-,387**	,239*	-0,133	-0,067	0,052	-0,095	-0,066	-0,05	-0,072	-0,087	-,258**	0,004	-,217*
	p				0	0	0,017	0,187	0,507	0,609	0,347	0,513	0,621	0,479	0,389	0,01	0,965	0,03
	N			100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (UU)	r				1	,665**	-0,127	,324**	0,168	-0,069	0,003	0,056	0,123	0,164	,205*	0,146	-0,171	,236*
	p					0	0,208	0,001	0,095	0,494	0,98	0,577	0,222	0,104	0,04	0,148	0,088	0,018
	N				100	100	100	100	100	100	100	100	100	100	100	100	100	100
Wie motiviert waren Sie, die Aufgabe zu lösen? (UU)	r					1	-,237*	,259**	,400**	0,131	-0,022	0,076	0,107	0,108	0,145	0,118	-0,046	0,176
	p						0,018	0,009	0	0,194	0,827	0,454	0,29	0,285	0,151	0,243	0,649	0,08
	N					100	100	100	100	100	100	100	100	100	100	100	100	100
Wie schwierig war diese	r						1	-,370**	-,393**	-0,099	-0,071	-,391**	-0,174	-0,034	0,007	-0,058	0,051	-0,029
	p							0	0	0,328	0,48	0	0,084	0,739	0,942	0,567	0,613	0,773

Aufgabe für Sie? (Alien)	N						100	100	100	100	100	100	100	100	100	100	100	100
Wieviel Spaß hat Ihnen die Aufgabe gemacht? (Alien)	r							1	,554**	0,125	-0,049	0,196	0,158	0,142	0,191	0,115	0,037	,208*
	p								0	0,217	0,631	0,05	0,117	0,158	0,057	0,254	0,717	0,038
	N							100	100	100	100	100	100	100	100	100	100	100
Wie motiviert waren Sie, die Aufgabe zu lösen? (Alien)	r								1	,298**	-0,092	,226*	0,194	0,193	,221*	0,106	0,003	,224*
	p									0,003	0,362	0,024	0,053	0,054	0,027	0,296	0,98	0,025
	N								100	100	100	100	100	100	100	100	100	100
Stimmung nachher (t2)	r									1	0,089	-0,035	-0,101	-0,101	-0,056	-0,082	0,17	-0,089
	p										0,378	0,73	0,315	0,32	0,582	0,418	0,09	0,378
	N									100	100	100	100	100	100	100	100	100
Stimmung orher (t1)	r										1	0,055	0,059	0,055	0,063	0,029	0,079	0,063
	p											0,588	0,563	0,588	0,536	0,778	0,436	0,534
	N										100	100	100	100	100	100	100	100
Anzahl Nennungen (UU)	r											1	,467**	0,049	-0,034	,204*	0,08	0,098
	p												0	0,632	0,736	0,041	0,43	0,334
	N											100	100	100	100	100	100	100
Anzahl kreative Nennungen (UU)	r												1	,851**	,727**	0,147	-0,03	,617**
	p													0	0	0,144	0,765	0
	N												100	100	100	100	100	100
Anteil kreativer Nennungen (UU) in %	r													1	,889**	0,11	-0,119	,713**
	p														0	0,274	0,238	0
	N													100	100	100	100	100
	r														1	0,129	-0,08	,806**

Qualitätsmaß (UU)	p															0,2	0,428	0
	N														100	100	100	100
Originalität Alien	r															1	-,302**	,692**
	p																0,002	0
	N															100	100	100
Ähnlichkeit Alien	r																1	-,239*
	p																	0,017
	N																100	100
Kreative Leistung	r																	1
	p																	
	N																	100

Anmerkungen:
r: Pearson-Korrelationskoeffizient
p: Zweiseitiges Signifikanzniveau der Korrelation
N: Anzahl Probanden
(UU): "Unusual Uses"-Aufgabe
(Alien): Imaginationsaufgabe
*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.
**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Statistiken

		Alter	Geschlecht der Versuchsperson	Beschäftigungsstatus der Versuchsperson	Studiengang der Versuchsperson, falls sie Student ist	Fachsemester der Versuchsperson, falls sie Student ist	Deutschkenntnisse der Versuchsperson
N	Gültig	100	100	100	100	96	100
	Fehlend	0	0	0	0	4	0
Mittelwert		22,45	1,75	1,01		2,61	1
Standardabweichung		3,95	0,435	0,1		8,667	0

Alter

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	18	4	4	4	4
	19	9	9	9	13
	20	17	17	17	30
	21	18	18	18	48
	22	15	15	15	63
	23	14	14	14	77
	24	5	5	5	82
	25	7	7	7	89
	26	4	4	4	93
	28	2	2	2	95
	29	1	1	1	96
	31	2	2	2	98
	41	1	1	1	99
	45	1	1	1	100
	Gesamt	100	100	100	

Geschlecht der Versuchsperson

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	männlich	25	25	25	25
	weiblich	75	75	75	100
	Gesamt	100	100	100	

Beschäftigungsstatus der Versuchsperson

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	99	99	99	99
	2	1	1	1	100
	Gesamt	100	100	100	

Studiengang der Versuchsperson, falls sie Student ist

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig		1	1	1	1
	Allg. Finanzverwaltung	3	3	3	4
	Fakultät 2	1	1	1	5
	Finanzverwaltung	6	6	6	11
	GPO 2011	1	1	1	12
	Grundschule PO	1	1	1	13
	Grundschullehramt	3	3	3	16
	GS (lehramt)	1	1	1	17
	Hauptstudium	1	1	1	18
	Lehramt	2	2	2	20
	Lehramt Primärstufe	1	1	1	21
	Lehramt Sek 1	2	2	2	23
	Lehramt WHR	2	2	2	25
	Public Management	58	58	58	83
	Sondernpädagokik	3	3	3	86
	Sonderpädagogik	4	4	4	90
	Steuer/Wirtschaftsrecht	1	1	1	91
	Steuerrecht	5	5	5	96
	Steuerverwaltung	2	2	2	98
	Steur- und Wirtschaftsrecht	1	1	1	99
	WTTR PO 2011	1	1	1	100
	Gesamt	100	100	100	

Fachsemester der Versuchsperson, falls sie Student ist

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	-77	1	1	1	1
	1	28	28	29,2	30,2
	2	8	8	8,3	38,5
	3	37	37	38,5	77,1
	4	3	3	3,1	80,2
	6	4	4	4,2	84,4
	7	1	1	1	85,4
	8	6	6	6,3	91,7
	9	1	1	1	92,7
	10	5	5	5,2	97,9
	11	1	1	1	99
	12	1	1	1	100
	Gesamt	96	96	100	
Fehlend	System	4	4		
Gesamt		100	100		

Deutschkenntnisse der Versuchsperson

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	100	100	100	100

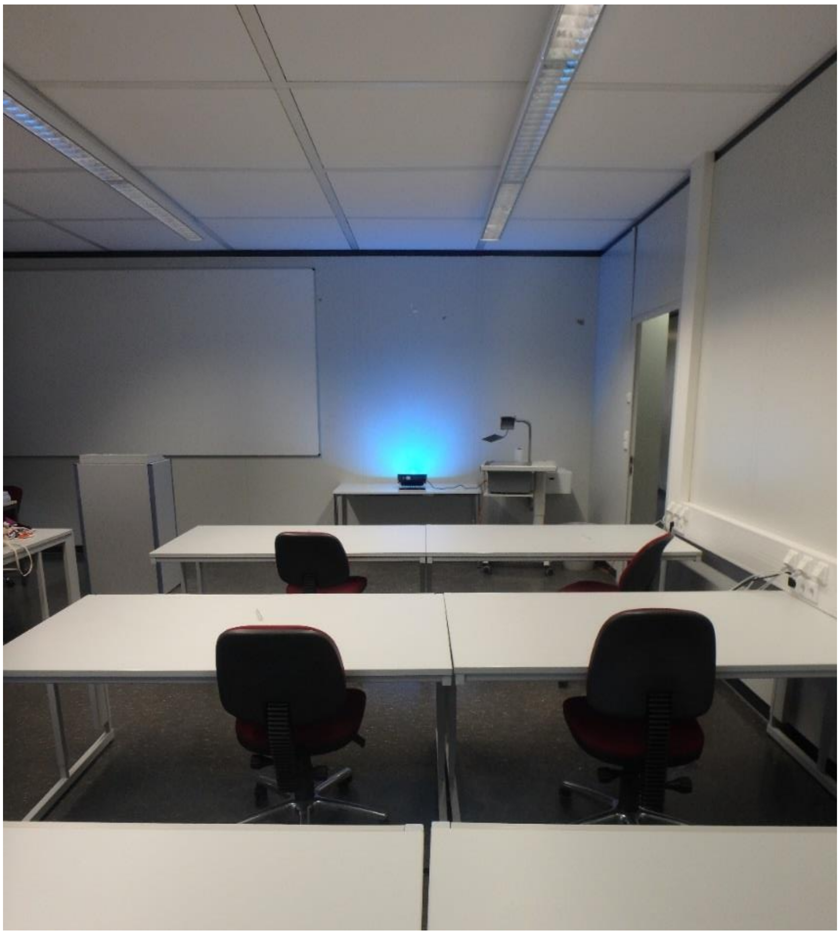
festgestellte Farbfehlsichtigkeiten

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	5	5	5	5

Anlage 3

Bilder von den Versuchsräumen

Raum 5.109



Raum 6.007

Raum 6.108

Anlage 4

Bundesamt für Strahlenschutz: Infrarot-Strahlung

Bundesamt für Strahlenschutz

Startseite Themen Optische Strahlung Infrarot-Strahlung Was ist Infrarot-Strahlung?

Was ist Infrarot-Strahlung?

Infrarotstrahlung (IR-Strahlung) - auch als Wärmestrahlung bezeichnet - ist Teil der optischen Strahlung und damit Teil des elektromagnetischen Spektrums (siehe Abbildung). Sie schließt sich in Richtung größerer Wellenlängen an das sichtbare Licht an. Ihr Wellenlängenbereich reicht von 780 Nanometer bis 1 Millimeter.

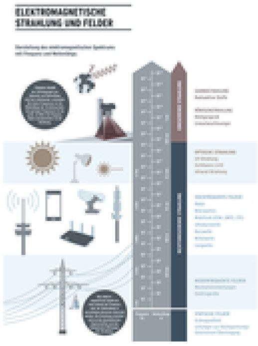
Unterteilung in IR-A-, IR-B- und IR-C-Strahlung

Infrarotstrahlung wird unterteilt in

- die kurzwellige IR-A-Strahlung mit einem Wellenlängenbereich von 780 bis 1400 Nanometer,
- die IR-B-Strahlung (1400 bis 3000 Nanometer) und
- den langwelligen Teilbereich, die IR-C-Strahlung (3000 Nanometer bis 1 Millimeter).

Sonne als wichtigste Quelle für Infrarot-Strahlung

Die wichtigste natürliche Quelle für Infrarot-Strahlung ist die Sonne. Infrarot-Strahlung hat einen Anteil von 50 Prozent an der Sonnenstrahlung, die den Erdboden erreicht. Außerdem gibt die durch die Sonneneinstrahlung erwärmte Erde Infrarot-Strahlung ab.



Elektromagnetisches Spektrum

Wärmehaushalt der Erde

Durch die in der Atmosphäre enthaltenen natürlichen und künstlichen Gase wie Wasser, Kohlendioxid, Ozon, Methan und Fluorchlorkohlenwasserstoffe (FCKWs) wird die von der Erde abgegebene Infrarot-Strahlung absorbiert. Dies führt zu einer zusätzlichen Erwärmung der Erde.

Dieser Prozess ist für den Wärmehaushalt der Erde und damit auch für die globale Erwärmung (Klimawandel) von entscheidender Bedeutung.

Anlage 5

Bundesamt für Strahlenschutz: UV-Strahlung

Bundesamt für Strahlenschutz

Wie wirkt UV-Strahlung?

Natürliche wie künstliche UV-Strahlung wirkt auf unseren Körper ein. In erster Linie sind davon unsere Augen und unsere Haut betroffen.

Wie tief UV-Strahlung in Auge und Haut eindringt, ist von der Wellenlänge abhängig: UV-A mit längeren Wellenlängen von 315 Nanometer bis 400 Nanometer dringt tiefer in Auge und Haut ein als die kurzwelligere UV-B-Strahlung mit Wellenlängen von 280 Nanometer bis 315 Nanometer.

In den Zellen des Auges und der Haut wird UV-Strahlung aufgenommen (absorbiert) und bewirkt dort unterschiedliche Veränderungen. Die wichtigste Veränderung ist die Schädigung des Erbguts (DNA) durch UV-Strahlung. UV-Strahlung schädigt auch in geringer Dosis die Erbsubstanz in den Hautzellen – weit bevor ein Sonnenbrand entsteht. Reparatursysteme in den Zellen beseitigen die Schäden am Erbgut in aller Regel wieder. Aber

- häufige,
- lang anhaltende und
- intensive UV-Bestrahlungen sowie
- Sonnenbrände

überbelasten diese Systeme. Die gesetzten Schäden werden dann nicht mehr vollständig beziehungsweise nicht fehlerfrei repariert und können zu bleibenden Erbgutveränderungen (Mutationen) werden. Damit steigt das Risiko für Hautkrebs. Die **Internationale Agentur für Krebsforschung [http://www.iarc.fr/]** (International Agency for Research on Cancer, IARC) hat daher im Jahr 2009 die UV-Strahlung der Sonne und künstliche UV-Strahlung in Solarien in die höchste **Risikogruppe 1 "krebserregend für den Menschen" [http://monographs.iarc.fr/ENG/Classification/index.php]** eingestuft.

UV-Strahlung hat kurzfristige und langfristige Wirkungen

Es lassen sich kurzfristige (akute) und langfristige (chronische) Wirkungen unterscheiden. Erstere treten unmittelbar oder Minuten, Stunden und Tage nach UV-Belastung auf,

Anlage 6

Fraunhofer IAO: LightFusion

FRAUNHOFER-INSTITUT FÜR ARBEITSWIRTSCHAFT UND ORGANISATION IAO

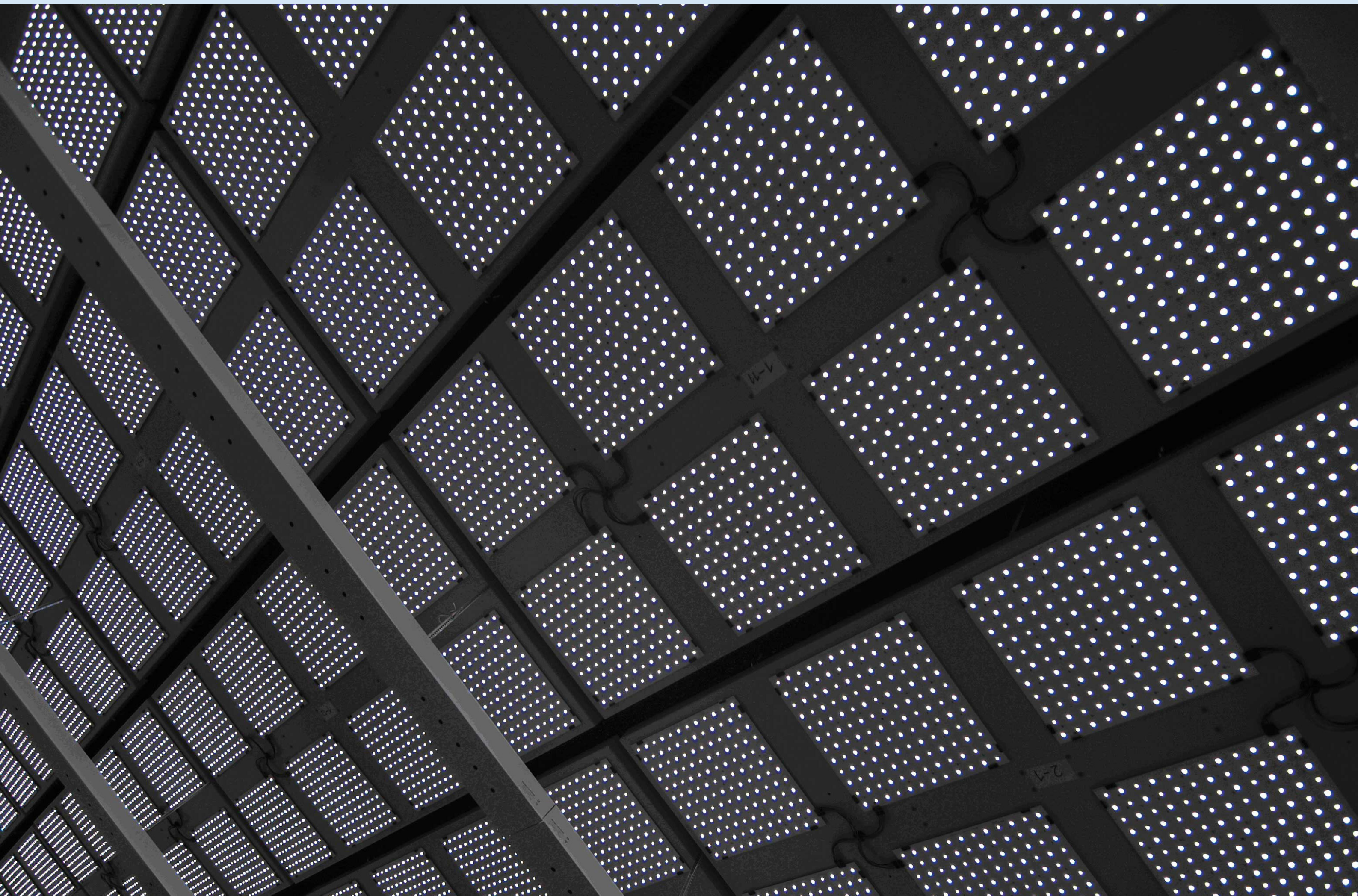
HINTERGRUNDINFORMATIONEN

LightFusion

NEUE ANSÄTZE FÜR LICHT UND DISPLAY AM ARBEITSPLATZ

Univ.-Prof. Dr.-Ing. Dr.-Ing. E.h. Dieter Spath, Dipl.-Ing. Matthias Bues, Dr.-Ing. Martin Braun, Dipl.-Ing. Designer Oliver Stefani
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Nobelstraße 12, 70569 Stuttgart

dieter.spath@iao.fraunhofer.de | matthias.bues@iao.fraunhofer.de | martin.braun@iao.fraunhofer.de | oliver.stefani@iao.fraunhofer.de



1. EINLEITUNG

Die Licht- und Displaytechniken befinden sich gleichermaßen in einem Umbruchprozess. Dies ist im Wesentlichen durch die technologische Entwicklung bei den Leuchtdioden (LED) und organischen Leuchtdioden (OLED) bedingt. LED besitzen zahlreiche vorteilhafte Eigenschaften, wie geringer Raumbedarf, hohe Leuchtdichten, lange Lebensdauer und hohe Energieeffizienz. In vielen Anwendungsbereichen der Lichttechnik haben LED deshalb andere Leuchtmittel bereits verdrängt; Beispiele hierfür sind Signalleuchten, Fahrzeug- und Fassadenbeleuchtung. Dieser Prozess dauert an; Studien zufolge wird im Jahr 2020 der LED-Anteil an der allgemeinen Beleuchtung erwartungsgemäß bei über 90 Prozent liegen (Jacob 2006).

Auch die Displaytechnik erfährt erhebliche Wandlungen durch diese Entwicklung. Aktuell bedeutet das vor allem den zunehmenden Einsatz von LED als Lichtquelle heutiger Displaytechnologien (z.B. für LC-Flachdisplays oder Datenprojektoren). Damit werden kompaktere, energieeffizientere Displaysysteme mit besserer Farbwiedergabe möglich. Der nächste große Schritt wird der Ersatz der Flüssigkristalle durch OLEDs sein. Damit werden noch kompaktere Flachdisplays möglich, die zudem auch auf biegsamen Substraten aufgebracht werden können. Kleinere Displays für Mobilgeräte sind bereits marktgängig; in einigen Jahren werden auch großflächige, hochauflösende OLED-Displays verfügbar sein.

Die skizzierten Entwicklungen führen zu völlig neuen Gestaltungsmöglichkeiten in der Beleuchtungs- und Displaytechnik. Leuchtende Tapeten, die je nach Bedarf auch Informationen darstellen, werden genauso vorstellbar wie multimodale »Smart Windows« (Oltean 2006), die alle Zustände vom transparenten Fenster bis zur diffusen Lichtquelle einnehmen können. Alle Flächen eines Raums, von Wänden bis hin zu Möbeln, können so zu Licht- und Informationsquellen werden. In Verbindung mit adaptiven optischen Systemen können einzelne Raumbereiche gezielt präzise ausgeleuchtet werden.

2. LightFusion – EIN INTEGRIERTER ANSATZ

Der gegenwärtige Wissensstand über Licht und dessen Anwendung bezieht sich vor allem auf direkte Wirkungen der Beleuchtung, die sich anhand von quantitativen Parametern (z. B. Intensität, Spektralverteilung, Lichtfarbe) beschreiben lassen. Die aktuelle Lichtforschung am Fraunhofer IAO berücksichtigt verstärkt auch indirekte Wirkungen des Lichts auf den Menschen. Einflüsse lassen sich in die Dimensionen der visuellen Wahrnehmung, der emotionalen Lichtstimmung und der physiologisch-biologischen Wirkungen gliedern. In der Arbeitsgestaltung blieben indirekte Wirkungen des Lichts etwa auf Stimmung und Gesundheit bislang weitgehend unberücksichtigt. Zudem wurden Wirkungen des Lichts zumeist eindimensional bzw. isoliert betrachtet: So existieren Empfehlungen und Richtlinien für Beleuchtungsstärken am Arbeitsplatz, Leuchtdichten und Auflösung von Bildschirmen etc.; erst in jüngerer Vergangenheit wurden Beleuchtungssysteme entwickelt, die die spektrale Veränderung des Sonnenlichts im Tagesverlauf nachzubilden versuchen.

Der LightFusion-Ansatz zielt auf eine integrierende Betrachtung aller drei Aspekte von Licht unter Nutzung der Gestaltungsmöglichkeiten, die sich durch LED- und OLED-Technologien ergeben, insbesondere großflächige Displays, präzise steuerbare Lichtquellen und Flächenleuchten auf OLED-Basis. Darüber hinaus wird auch das Tageslicht einbezogen, was unter Gesichtspunkten der Energieeffizienz und Gesunderhaltung gleichermaßen wichtig ist. Ziel ist die Schaffung einer dynamischen, individualisierbaren Lichtsituation, die günstige Voraussetzungen für die menschliche Leistungsfähigkeit schafft und dabei gleichzeitig gesundheitsfördernde Ressourcen aktiviert. Besondere Bedeutung kommt hierbei der Beeinflussung des circadianen Systems zu.

Systemtechnisch manifestiert sich dieser integrierte Ansatz in der integrierten Steuerung und Regelung aller Lichtquellen und -senken eines Raums. In einer verallgemeinerten Betrachtung kann jede Fläche des Raums Licht aussenden oder reflektieren.

Wird die Charakteristik jeder Fläche erfasst, kann mit Methoden aus der Computergraphik die räumliche Verteilung des Lichteinfalls theoretisch für jeden Punkt im Raum berechnet werden. Durch Modulation der Lichtquellen und Displays sowie durch die Änderung der Reflexionseigenschaften passiver Flächen mittels E-Paper-Technologien wird diese Lichtverteilung physikalisch umgesetzt.

3. LICHT AM WISSENSARBEITSPLATZ DER ZUKUNFT

3.1 Wirkungen des Lichts

3.1.1 Grundlagen

Licht ist für den Menschen ein lebensnotwendiges Medium. Neben der Informationsaufnahme beeinflusst Licht nahezu den gesamten vitalen Bereich des Menschen (d.h. die Rhythmik der Organe bzw. Lebensfunktionen).

Der gegenwärtige Wissensstand über Licht und dessen Anwendung bezieht sich vornehmlich auf direkte Wirkungen des Lichts zur Informationsaufnahme (d.h. Intensität, Spektralverteilung, Lichtfarbe etc.). Im Kontext der Lichtforschung sollen verstärkt die indirekten Wirkungen von Licht auf den Menschen berücksichtigt werden. In eine umfassende Betrachtung des Lichts werden

- visuelle Wahrnehmungsgrundlagen (d.h. Informationsvermittlung),
- psychisch-emotionale Wirkungen (d.h. Lichtstimmungen) und damit verbundene Raummilieus (d.h. Beleuchtung, Tages- und Kunstlicht),
- physiologisch-biologische Wirkungen auf Gesundheitszustand und Leistungsfähigkeit (z. B. circadianer Rhythmus, SAD, Melatonin, Vitamin D)

einbezogen.

Das Erfordernis einer innovativen Lichtforschung und daraus resultierende Verbesserungen der Beleuchtungs- und Anzeigesysteme werden durch die wachsende Unzufriedenheit der arbeitenden Personen belegt: Mehr als 50 Prozent der Befragten einer zufälligen Stichprobe in einer Studie von Çakir/Çakir (1998) beschreiben die Beleuchtungssituation an ihren Arbeitsplätzen mit negativen Attributen wie unangenehm, unfreundlich und grell.

Mit zunehmender Entfernung des Arbeitsplatzes vom Fenster bzw. mit abnehmender Wirkung des Tageslichts nehmen Befindlichkeitsstörungen am Büroarbeitsplatz zu. Künstliche Beleuchtung wird von 59 Prozent der befragten Büromitarbeiter nicht nur als unangenehm, sondern auch als ständige Beeinträchtigung der Gesundheit empfunden (Çakir/Çakir 1998). Trotzdem werden immer noch viele Büro-Arbeitsplätze unter weitgehendem Ausschluss von Tageslicht betrieben und die negativen Folgen künstlicher Beleuchtung häufig unzureichend wahrgenommen. Dies liegt u. a. am hohen Anpassungsvermögen des menschlichen Auges an veränderte Lichtsituationen, welches über mehrere Zehnerpotenzen hinweg reicht (Fisch 2000). Ein weiterer Grund ist darin zu sehen, dass die Ermüdung vieler Elemente des Sehapparats kaum messbar ist und möglicherweise auch weniger stark wahrgenommen wird, so dass keine wirksamen Schutzmechanismen ergriffen werden.

3.1.2 Visuelle Wahrnehmung

Die Relevanz visueller Wahrnehmungen wird deutlich, wenn man deren Anteil an der menschlichen Sinneswahrnehmung betrachtet, der schätzungsweise 85 Prozent beträgt (Fisch 2000). Visuelle Wahrnehmungsaspekte (d.h. Sehleistung und Informationsaufnahme) gelten beim derzeitigen Stand von Wissenschaft und Technik als weitgehend gelöst. Von besonderer Bedeutung bei der Gestaltung mit Displays ist die Kontrastempfindlichkeit des Auges. Sie ist als die Fähigkeit definiert, Gitter mit niedrigem Kontrast zu diskriminieren. Obgleich sich die Wahrnehmung an sehr große Kontraste anpassen kann (8-9 Größenordnungen) ist der lokale Kontrast einer spezifischen Region nur auf 1:100 bis 1:150 begrenzt (Wandell 1995; Vos/van Meeteren 1972).

Die räumliche Auflösung des Auges ist individuell; sie wird von optischen Eigenschaften sowie von der Zellenverteilung auf der Retina bestimmt. Diese hängen nichtlinear vom Helligkeitsniveau der Szene ab (Wandell 1995; Waren 2000). Bei der Sehschärfe (Visus) unterscheidet man zwischen folgenden Werten:

- Punktschärfe bei hellem Umgebungslicht: 1-1.2 Bogenminuten (0.35-0.3 Millimeter pro unterscheidbarem s/w Pixelpaar auf 1 m)
- Punktschärfe bei niedrigem Umgebungslicht: 2-3 Bogenminuten (0.87-0.58 Millimeter pro unterscheidbarem s/w Pi-xelpaar auf 1 m)
- Gittersehschärfe: 1-2 Bogenminuten (Die Fähigkeit, ein Muster von hellen und dunklen Stäben von einer konstanten grauen Fläche zu unterscheiden)
- Buchstabenschärfe: 5 Bogenminuten (Die Fähigkeit Buchstaben zu erkennen. Der sogenannte »Snellen-Index« ist der Standard zum Messen dieser Fähigkeit.)
- Stereoschärfe: 10 Bogenminuten (Die Fähigkeit, Gegenstände in der Tiefe noch aufzulösen).
- Noniussehschärfe: 10 Bogensekunden (Die Fähigkeit zu unterscheiden, ob zwei gleichgerichtete gerade Linien etwas gegeneinander verschoben sind).

Relevant ist auch der Grenzwert für das Wahrnehmen von Flimmern, die sogenannte Critical Flicker Frequenzy (CFF). Wenngleich diese ebenfalls vom Umbebungslicht und der Position auf der Retina abhängt, man geht davon aus, dass Frequenzen oberhalb von 60 Herz nicht wahrgenommen werden.

Beleuchtungssysteme werden nach den bekannten Parametern des visuellen Systems gestaltet. Die hierbei einbezogenen vier lichttechnischen Grundparameter sind der Lichtstrom, die Lichtstärke, die Leuchtdichte und die Beleuchtungsstärke. Neben diesen lichttechnischen Grundgrößen werden in den Beleuchtungsnormen folgende Charakteristika betrachtet: Blendung, Lichtrichtung und Schattigkeit sowie Farbwiedergabe und Lichtfarbe (DIN EN 12464-1). Aufgrund des Wissensstands wird im Bereich der visuellen Wahrnehmung derzeit kein vordringlicher arbeitswissenschaftlicher Forschungsbedarf gesehen. Dieser ergibt sich erst bei Wechselwirkungen mit den weiteren Wirkungsdimensionen des Lichts, wie sie im Folgenden dargestellt werden.

3.1.3 Psychisch-emotionale Wirkung

Die Sonne hat einen wärmenden symbolischen Charakter (Baumeier 2000). Lorincz (1960) schreibt dem Sonnenlicht eine entspannende Wirkung sowie eine Steigerung des Wohlbefindens, des Appetits und der Qualität des Schlafs zu. Dass Licht zur Behandlung von Stress und Stimmungsschwankungen angewendet werden kann (Liberman 2005), belegt seine Bedeutung auf psychisch-emotionaler Ebene. Klinische Studien haben die Wirksamkeit der Lichttherapie zur Behandlung verschiedener Gemütskrankheiten bewiesen (Wetterberg 1993; Lam 1998).

Licht wirkt vornehmlich unbewusst. Ausgewogene Beleuchtung und angenehme Lichtfarben dienen kurz- und langfristig dem Wohlbefinden (Fisch 2000). Welche Stimmung ein bestimmtes Licht bei einem Menschen hervorruft, hängt jedoch von dessen Erfahrungen und Vorlieben ab. Für alle Menschen gleichermaßen angenehme Licht- und Beleuchtungsbedingungen existieren daher nicht. Für die Lichtgestaltung ist daher bedeutsam, dass die Beleuchtung individuell und kontextspezifisch anpassbar ist.

Nach Ehrenstein (2002) hat der Mensch während der ergotropen Tagphase normalerweise ein erhöhtes Bedürfnis nach Licht. Während der trophotropen Nachtphase besteht physiologische Lichtscheu beziehungsweise eine Bevorzugung warmen Lichts, wie es Kerzen und Glühlampen emittieren. Dies sollte im Rahmen einer dynamischen Lichtgestaltung berücksichtigt werden.

Kühlweiße fluoreszierende Beleuchtung erzeugte bei Schulkindern verstärkt Hyperaktivität, Erschöpfung, Reizbarkeit und Aufmerksamkeitsstörungen, wohingegen Vollspektrumlampen die Überwindung von Lese- und Lernproblemen herbeiführten (Liberman 2005). Um die kognitive und emotionale Ebene positiv zu beeinflussen, sollen daher Beleuchtungssysteme eingesetzt werden, welche dem Tageslichtspektrum nahekommen.

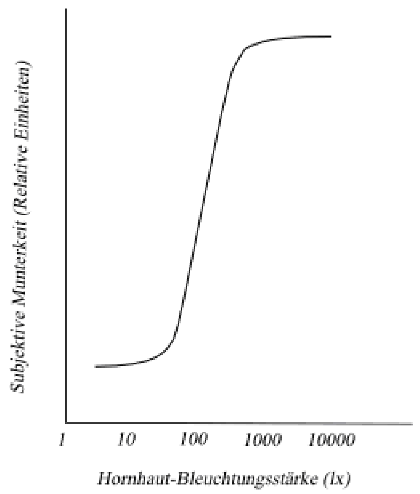


Abbildung 1: Subjektive Munterkeit in Abhängigkeit der Hornhaut-Beleuchtungsstärke (Krüger 2007).

Ein weiterer Aspekt, der auch auf emotionaler Ebene relevant wird, ist die Blendung. Blendung wird häufig mit den Attributen farblos und kalt verbunden (stärker noch als bestimmte Lichtfarben). Farblosigkeit entsteht darüber hinaus häufig durch graue Einrichtungen und kann dann durch die Beleuchtung kaum vermieden werden. Çakir/Çakir (1998) betonen, dass es durch blendende künstliche Beleuchtung zur Beeinträchtigung der Stimmung kommen kann. Eine erhöhte Belastung des Sehapparats, wie sie durch ungünstige Sehaufgaben oder höhere Blendung verursacht werden kann, führt zu einer höheren inneren Anspannung und stärkerer psychischer Ermüdung (Çakir 1979).

Um vorzeitiger Ermüdung entgegenzuwirken, wird eine Dynamisierung der Beleuchtung empfohlen. Was zunächst als Nachteil der Tageslichtbeleuchtung angesehen wird, offenbart sich unter Aufrechterhaltung bestimmter Mindestbeleuchtungsstärken als besonderer Vorteil für die menschliche Tätigkeitsausführung.

Weitere Anhaltspunkte zur subjektiven Präferenz bei der Beleuchtungsstärke gibt Krüger (2007): Als Stimmungsparameter kann das Bewertungspaar »angenehm – unangenehm« herangezogen werden. Bei einer Variabilität der Beleuchtung von 200 Lux bis 600 Lux und einem Indirektanteil der Beleuchtung zwischen 0 Prozent und 100 Prozent werden höhere Beleuchtungsstärken und ein höherer Indirektanteil als angenehmer bewertet. Höhere Beleuchtungsstärken führen darüber hinaus zu höherer subjektiver Munterkeit, wie in Abbildung 1 dargestellt.

Neben steigender Beleuchtungsstärke bewirkt auch ein steigender Direktanteil der Beleuchtung höhere Munterkeit, wohingegen indirekte Beleuchtung eher beruhigend wirkt (Krüger 2007).

Psychisch-emotionale Wirkungen beziehen sich grundsätzlich auf die Subjektivität des Individuums und sind daher anhand objektiver Kriterien zuweilen schwer zu erfassen. Lichtgestaltung zur Stimmungsbeeinflussung erfolgt zumeist intuitiv. Zur Licht- und Farbgestaltung sind die Grundsätze der Harmonie zu berücksichtigen, die sich in Vielfalt ausdrückt. Richtige Farbgebung durch Licht verbessert die Wahrnehmung, steigert die Motivation und verringert Fehlleistungen, hebt das allgemeine Befinden, schafft Ordnung und verbessert dadurch Orientierung und Erholung. Neben den Grundsätzen zu Erzeugung von Farbharmonien (z. B. Harmonie durch Komplementärkontraste, Mengenkontraste, Helligkeitskontraste, Warm-Kalt-Kontrast) werden insbesondere auch Richtlinien zur Vermeidung von Disharmonien berücksichtigt. Insbesondere der Warm-Kalt-Kontrast sollte bei der Beleuchtung berücksichtigt werden, da sogenannte Symphaticotoniker bzw. Prasymphatikoniker unterschiedlich auf warmes und kaltes Licht reagieren. Durch Einwirkung von farbigem Licht können Adrenalin und Insulin den Blutzuckerspiegel erhöhen respektive senken. Darüber unterstützen Kontraste auch die visuelle Wahrnehmung (Çakir/Çakir 1998). Frieling (1992) empfiehlt, eine einseitige Farbreizung (z. B. durch Licht) zu vermeiden und besser eine dynamische Farbreizung zu verwenden.

3.1.4 Physiologisch-biologische Wirkung des Lichts auf die Gesundheit

Der biologische Wirkungsbereich des Lichts unterscheidet sich erheblich von der visuellen Wirkung. Besondere Aufmerksamkeit wird der circadianen Wirkung des Lichts geschenkt, weil sie das gesundheitliche Befinden des Menschen wesentlich beeinflusst. Über das Auge (und die Haut) reguliert das Licht sämtliche Hormondrüsen und den gesamten Zell-Stoffwechsel. Licht triggert die innere Uhr nach tages-, wochen- und jahreszeitlichen Rhythmen. Die Kommunikation mit den verschiedenen Systemen des Körpers beruht auf der Ausschüttung des Hormons Melatonin in den Blutkreislauf (Rea 2002).

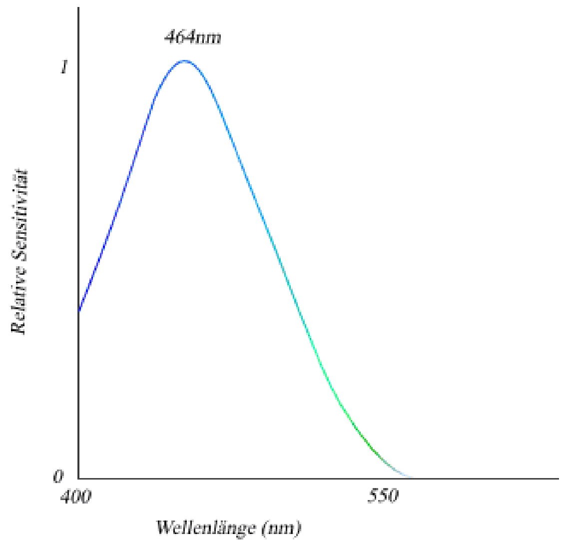


Abbildung 2: Circadiane Empfindlichkeitskurve (Brainard et al. 2001).

Im Auge befinden sich drei Arten von Zapfen (blau, rot, grün) als Photorezeptoren, welche beim photopischen Sehen den visuellen Prozess einleiten. Hinzu kommen die Stäbchen, die ein skotopisches Sehen ermöglichen. Anders als Zapfen und Stäbchen dient der Photorezeptor Melanopsin in der Netzhaut des Auges nicht dem Sehen, sondern steuert das von der Zirbeldrüse produzierte Melatonin (Baumeier 2000). Die Empfindlichkeit des circadianen Photorezeptors Melanopsin in Abhängigkeit der Wellenlänge zeigt Abbildung 2 (Brainard et al. 2001). Das Maximum der Melatoninsuppression und damit der circadianen Aktivierung liegt bei einer Wellenlänge von ca. 460 Nanometern, was blauem Licht entspricht (Thapan 2001).

Die Produktion von Melatonin wird durch Lichteinfluss unterdrückt (Wirz-Justice/Roenneberg 2004). Melatonin löst Müdigkeit aus und steuert somit den Schlaf-Wach-Rhythmus sowie andere circadiane Organfunktionen. Durch die Tagesrhythmik von Lichtstärke und Lichtfarbe werden somit Leistungs- bzw. Regenerationsphasen des Menschen unterstützt (Spath et al. 2004).

Der Einfluss des Zeitgebers Licht auf den endogenen Oszillator hängt auch vom Tageszeitpunkt ab: Befindet sich der endogene Oszillator in der Tagphase, reagiert er kaum auf Lichtreize, zu Beginn der Nacht verschiebt sich die Rhythmik durch Lichteinfluss nach hinten und zu Ende der Nacht nach vorne (Roenneberg et al. 2003). Laut Figueiro (2002) ist die Unterstützung der Circadianrhythmik durch Bestrahlung in bereits genannter Wellenlänge und Beleuchtungsstärke morgens sinnvoll.

Die räumliche Verteilung, welche bei der visuellen Wahrnehmung von großer Relevanz ist, scheint für das circadiane System keine allzu große bedeutende Rolle zu spielen (Rea 2002). Die Beschreibung der Auslöser für die Suppression von Melatonin ermöglicht sowohl ein besseres Verständnis der circadianen Rhythmik als auch eine gezielte Beeinflussung derselben hin zu einer gesundheitsförderlichen Gestaltung von Beleuchtungssystemen. Brainard et al. (2001) schlagen daher für Arbeitsplätze neue Herangehensweisen bei der Planung von Beleuchtungssystemen vor, welche sowohl das visuelle als auch das circadiane System berücksichtigen. Dies erfordert die Entwicklung einer speziell auf die Stärke der circadianen Beeinflussung ausgelegten Lichtmessung. Das photometrische System ist mit der visuellen Empfindlichkeit des Auges gewichtet. Diese unterscheidet sich jedoch deutlich von der Empfindlichkeit des Melanopsin. Differenzen bestehen bezüglich des Spektrums, aber auch Quantität, Zeitpunkt und Dauer der Lichteinwirkung. In Anlehnung an photometrische Größen definierten Gall et al. (2004) einen circadianen Wirkungsfaktor, durch dessen Multiplikation mit der Beleuchtungsstärke sich die »circadiane Beleuchtungsstärke« ergibt.

Die organisatorische Arbeitsgestaltung schafft häufig Bedingungen, die gegen die Funktionsmuster der menschliche Physiologie verstoßen: War man vor der Erfindung der Glühbirne gezwungen, sich an den tageszeitliche Lichtrhythmus anzupassen, so wurde mit Einführung des künstlichen Lichts die Möglichkeit geschaffen, den Arbeitsrhythmus durch andere Vorgaben festzulegen (Spath et al. 2004). Das Arbeiten in den frühen Morgen- oder späten Abendstunden kann daher insbesondere in den Wintermonaten eine gewisse Belastung darstellen.

Eine biologisch wirksame Lichtgestaltung bedingt eine dynamische Beleuchtung (hinsichtlich spektraler, räumlicher und zeitlicher Dimension). Dynamisierung erfolgt vor dem Hintergrund einer Individualisierung von Beleuchtung, d.h. der tendenziellen Abkehr von normierten Lichtkonzepten. Bei der Dynamisierung kann die sinnvolle Variation von Lichtintensität, Lichtfarbe und Lichtverteilung zu einem Erlebniswert führen. Wesentlich ist hier die Unterscheidung zwischen dem unbewussten Erleben und dem bewussten Verändern.

3.1.5 Prinzipien der Arbeitsplatzgestaltung

Zukunftsweisende Beleuchtungs- und Anzeigesysteme sollen geeignete Leistungsvoraussetzungen bei geistiger bzw. körperlicher Arbeit schaffen und zur Gesunderhaltung des arbeitenden Menschen beitragen. Neben Aspekten der Sehleistung soll die Beleuchtung eine Aktivierung des Organismus bewirken, damit die Tätigkeit nicht vorschnell zur Ermüdung und zu ermüdungsbedingten Aufmerksamkeitsfehlern führt. Um die Raum- und Arbeitsplatzbeleuchtung auf die Bedürfnisse des Menschen abzustimmen, sind folgende Kriterien zu berücksichtigen:

- Einhaltung der lichttechnischen Gütemerkmale
- bei sinnvoller dynamischer Veränderung des Lichts
- unter Einbeziehung des natürlichen Tageslichts.

Wesentlich für innovative Lichtkonzepte ist, dass Beleuchtung nicht als statische, möglichst gleichmäßige Installation in einen Raum und seinen begrenzenden Flächen aufgefasst wird. Beleuchtung soll vielmehr als dynamisches Design eines visuellen Raumklimas verstanden werden.

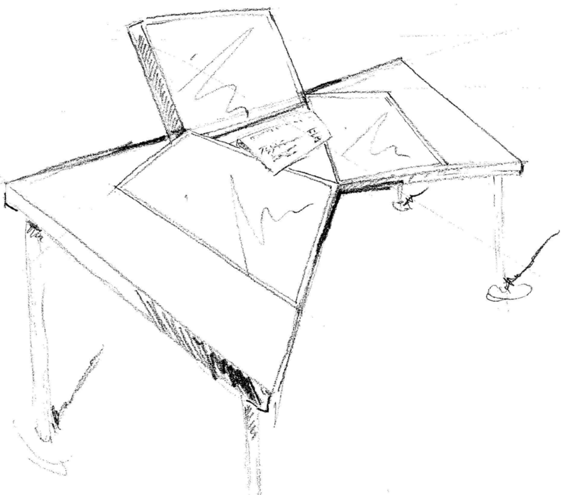


Abbildung 3: Arbeitsplatzkonzept mit drei integrierten Flachbildschirmen.

Studien zeigen, dass viele Monitore am Arbeitsplatz zu hoch positioniert werden. Dem menschlichen Auge fällt es jedoch wesentlich leichter, auf tief platzierte Objekte zu akkomodieren (Bücher werden beim Lesen üblicherweise auch nicht auf Augenhöhe gehalten). Bereits in den 1990er Jahren wurden ergonomische Arbeitsplatzkonzepte mit tief in den Schreibtisch integrierten Monitoren entwickelt. Allerdings konnten sich diese Konzepte aufgrund der sperrigen CRT-Monitore seinerzeit nicht durchsetzen. Die heute weit verbreitete Flachbildschirmtechnologie ermöglicht es hingegen, großflächige, hochauflösende Bildschirme direkt in die Tischfläche zu integrieren. Wenngleich das papierlose Büro vielfach propagiert wird, stapeln sich nach wie vor Papierberge auf vielen Schreibtischen. Warum also nicht mit intuitiven Interaktionskonzepten die gesamte Schreibtischfläche als Display nutzen und Zettel nur noch virtuell auf dem Schreibtisch ablegen, verschieben und weitergeben? Abbildung 3 zeigt ein deratiges Konzept in Verbindung mit einem zusätzlichen, herkömmlich platzierten Monitor, der dem Benutzer die Möglichkeit des Haltungswechsels und des aufrechten Blicks für entfernte Objekte (z. B. Übersichtsgrafiken, Videokonferenzen etc.) bietet. Solch ein Arbeitsplatzkonzept stellt auch eine Lösung des Problems der optimalen Arbeitsplatzausleuchtung dar: Für Arbeiten am Computer werden 300-500 Lux empfohlen. Arbeiten mit Papierdokumenten erfordern aber mehr Licht. Durch die selbst-leuchtende Bildschirmfläche im Arbeitstisch wäre eine ausreichende Helligkeit gewährleistet.

4. VOM BILDSCHIRMARBEITSPLATZ ZUM INFORMATIONSRAUM

Ansatz der arbeitswissenschaftlichen Forschung ist die Anpassung der (Licht-) Technik an den arbeitenden Menschen, so dass dieser beeinträchtigungsfreie sowie gesundheits- und leistungsförderliche Arbeitsbedingungen vorfindet. Ziele einer methodisch fundierten Planung, Gestaltung und Evaluierung von Beleuchtungs- und Anzeigesystemen sind gleichermaßen produktive, effiziente und gesundheitsgerechte Arbeitsbedingungen.

Der heutige Wissensarbeiter verbringt einen großen Teil seiner Arbeitszeit vor einem Bildschirm von etwa 50 cm Diagonale. Dieser bildschirmzentrierte Arbeitsplatz ist mehr durch technische Gegebenheiten als durch tatsächliche Anwendungserfordernisse geprägt worden. Neben den ergonomischen Schwächen prägen die begrenzte Bildschirmfläche und die dadurch limitierte gleichzeitig darstellbare Informationsmenge die Arbeitsabläufe.

In Meeting- und Präsentationssituationen haben sich große Displayflächen als vorteilhaft erwiesen. Das erscheint zunächst naheliegend wegen der größeren Betrachteranzahl und des damit erforderlichen Betrachtungsabstandes; viel wichtiger ist jedoch die größere Pixelanzahl und damit die größere Menge an Informationen, die gleichzeitig dargestellt und verknüpft werden können. Im Projekt Z-VISUM des Fraunhofer IAO wurde ein universell nutzbarer Meeting- und Präsentationsraum konzipiert und realisiert (Blach 2007). Dieser Raum verfügt als zentrales Element über eine Displaywand mit einer Auflösung von 3072 x 1536 Pixel. Auf dieser Wand können Bildinformationen aus verschiedenen Quellen flexibel kombiniert werden. Darüber hinaus verfügt das Z-VISUM-Labor über eine stereoskopische Displaywand mit 280 x 210 cm und einer Auflösung von 1400 x 1050 Pixel. Auf Basis dieser Infrastruktur wurden verschiedene Anwendungsszenarien aus den Bereichen Produktentwicklung, Digitale Produktion, Business Performance Management und Softwaremanagement entwickelt. Das Thema Licht wurde in diesem Projekt nur hinsichtlich der Anpassung der Beleuchtung an die Randbedingungen der Displaysysteme berücksichtigt. Maßgeblich ist hier der geforderte Kontrast bei gegebener maximaler Leuchtdichte der Displays, der durch die Anpassung der Umgebungshelligkeit beeinflusst werden kann.

5. DER nLightened WORKPLACE

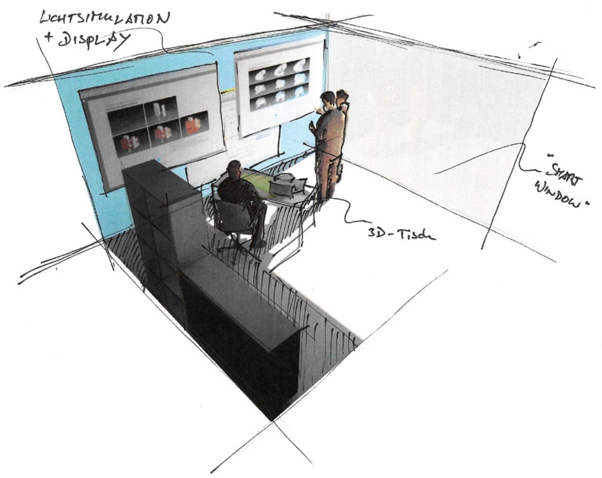


Abbildung 4: nLightenend Workplace (Konzept).

Das Konzept des nLightened Workplace überträgt Erkenntnisse aus Meeting- und Präsentationssituationen auf den Individualarbeitsplatz und misst dabei dem Thema Licht entscheidende Bedeutung zu. Verbesserte ergonomische Bedingungen sind dabei ein wichtiges Ziel; wesentlich ist jedoch die Entwicklung neuer Arbeitsprozesse und Interaktionsformen in der Wissensarbeit. Der nLightened Workplace wird die Grenzen des klassischen Bildschirmarbeitsplatzes überwinden und diesen durch das verallgemeinerte Konzept der digitalen Arbeitsfläche ersetzen. Diese ist idealerweise immer in derjenigen Größe, Lage und Auflösung verfügbar, die der Anwender zur aktuellen Aufgabenbewältigung benötigt. Damit werden neue individuelle, aber auch kooperative Arbeitsweisen möglich.

Ein Beispielszenario ist der Arbeitsprozess des Designers, der in allen Phasen der Produktgestaltung, von den ersten Entwurfsskizzen über die Diskussion mit Teamkollegen bis zur Ausarbeitung und hochwertigen Visualisierung des Entwurfs, vom nLightened Workplace optimal unterstützt wird.

Der LightFusion-Ansatz betrachtet zudem die Tatsache, dass Displays auch Lichtquellen sind, aus denen am Arbeitsplatz ein großer Teil der vom Auge aufgenommenen Lichtmenge stammt. Neuartige OLED-Displays und Displays mit LED-Beleuchtung sind einerseits sehr energieeffizient und bieten eine hervorragende Farbwiedergabe, andererseits emittieren diese Displays im Gegensatz zu herkömmlichen Displays blaues Licht mit einem sehr ähnlichen Spektrum, wie es von den Melanopsinempfängern im Auge absorbiert wird. Dies legt die Vermutung nahe, dass durch solche Displays die Müdigkeit des Betrachters verringert werden kann. Auswirkungen von neuartigen LC-Displays mit LED-Beleuchtung bzw. OLED Displays auf die circadiane Rhythmik des arbeitenden Menschen sind noch weitgehend unerforscht.

Bei einer Untersuchung der Wirkung heller bzw. dunkler CRT-Monitoren auf die Müdigkeit hat sich eine signifikant höhere Melatoninkonzentration beim dunklen Monitor herausgestellt (Higuchi et al. 2003). Cajochen et al. (2005) wiesen die Wirkung bereits geringer Mengen an monochromatischem Licht der Wellenlänge um 460 Nanometer am Abend auf die Melatoninsuppression und somit auf die Müdigkeit nach.

Weitere Studien von Desan et al. (2007), Glickman et al. (2006) und Wright et al. (2001) zeigen die Möglichkeit der Melatoninsuppression durch LED-Licht im Blauspektrum. Jasser et al. (2006) stellten zwar fest, dass weißes Licht von nur 18 Lux (bei einer Aussetzung von Probanden über zwei Stunden während der Nacht) bereits eine signifikante Reduzierung der Melatoninsuppression bei anschließender Beleuchtung durch monochromatisches 460 Nanometer-Licht verursacht, eine Verstärkung der Melatoninsuppression durch LED-beleuchtete Displays wurde bislang aber nicht untersucht.

Im Rahmen der Entwicklung des nLightened Wokplace werden deshalb Untersuchungen zur spektralen Emission und der Reaktion des Menschen auf neuartige Displays (d.h. Konzentration, Müdigkeit, Vigilanz etc.) durchgeführt. Es werden Ergebnisse erwartet, die die Entwicklung von Informationsdisplays wesentlich beeinflussen könnten, und zwar hinsichtlich einer steuerbaren Stimulation der Melanopsinempfänger durch das Displaysystem.

Neben der physiologischen Wirkung kann eine spektral variable Beleuchtung auch psychologische Wirkungen (z. B. Stimmungen) erzeugen und auf diese Weise adaptiv und sogar über physiologische Messmethoden selbst-adaptiv je nach Bedarf entspannende oder anregende Lichtsituationen vorschlagen.

Durch die enge Zusammenarbeit mit den Universitären Psychiatrischen Kliniken Basel werden Display- und Beleuchtungs-Prototypen auf psychologische und physiologische Wirkungen getestet und validiert.

Das prototypische Konzept des nLightened Workplace sieht einen Raum mit einer Nutzfläche von 15 m^2 vor. Die Wandflächen sind modular austauschbar und werden damit wahlweise als Displays, Leuchtflächen oder passive Flächen verwendet. Hinzu kommen Punktlichtquellen, die automatisch positionierbar sind und damit das physische »Highlighting« einzelner Gegenstände im Raum ermöglichen. Die durch den modularen Ansatz erreichte Variabilität ist auch deshalb wichtig, weil damit neue Displaytechnologien – sobald prototypisch verfügbar – schnell in das System integriert werden können.

Die Sonne ist eine Quelle von kostenlosem Licht in unerschöpflicher Menge. Verschiedenste Tageslichtsysteme, die das natürliche Lichtpotenzial nutzbar machen, indem sie z. B. Tageslicht vom Fenster in die Raumtiefe transportieren, existieren bereits. Aber auch großflächige, passive Displays (LCD) können Tageslicht als Lichtquelle nutzen, was unter Aspekten der Energieeinsparung vorteilhaft ist. Eine Kombination aus variablem Diffusor und einem LCD-Panel kann Tageslichtflächen in Räumen flexibel als Displayflächen, aber auch als Fenster nutzbar machen. Eine zusätzliche transparente OLED-Schicht macht die Displayfunktion eines solchen »Smart Window« darüber hinaus vom Tageslicht weitgehend unabhängig.

Alle Komponenten des nLightened Workplace, also Lichtquellen, Displays, Tageslichtelemente und passive Flächen, werden über ein zentrales System gesteuert. Dieses kennt die geometrische Verteilung aller Komponenten und deren physikalische Eigenschaften wie die maximale Leuchtdichte. Bei den Displayflächen werden zudem die jeweils dargestellten Bildinhalte berücksichtigt; diese sind für das von der Displayfläche abgestrahlte Licht maßgeblich.

Die Übermittlung der Bildinformation zu den jeweiligen Displays erfolgt netzwerkbasiert; jede Displayfläche ist ein im Netzwerk sichtbarer Knoten mit definierten Parametern zu Größe, Position, Auflösung und Leuchtcharakteristik. Die Leuchtdichtenanpassung der Displayflächen erfolgt über eine direkte Ansteuerung ihrerLichtquellen. Das zentrale Steuerungssystem ermöglicht die gezielte, dynamische Darstellung einer optimalen, in gewissen Grenzen auch vom Benutzer wählbaren Lichtsituation im Raum.

Der nLightened Workplace wird die Forschungs- und Integrationsplattform des LightFusion-Konzepts sein, einerseits also eine Testumgebung für neue Licht- und Displaytechnologien und andererseits eine prototypische Anwendungsumgebung, in der Arbeits- und Interaktionskonzepte für konkrete Anwendungsfälle entwickelt und evaluiert werden. In dieser Doppelfunktion fügt sich der nLightened Workplace nahtlos in das neu entstehende Zentrum Virtuelles Engineering (ZVE) des Fraunhofer IAO ein, das ganz im Zeichen der Erforschung und Umsetzung innovativer Formen der Wissensarbeit steht.

6. LITERATUR

Baumeier, D.: Der Einfluss von Licht auf die Psyche. Leipzig, Uni-versität, Diss., 2000.

Blach, R.: Boardroom VISUM des Fraunhofer Instituts für Ar-beitswirtschaft und Organisation IAO: Visuelles Unternehmensmanagement und die Erarbeitung von Strategien als Virtual Reality im 3D-Modell. Professional System (2007) Nr. 2, Seite 40-44.

Brainard, G.; Hanifin, J.; Greeson, J.; Byrne, B.; Glickman, G.; Gerner, E.; Rollag, M.: Action Spectrum for Melatonin Regulation in Humans: Evidence for a Novel Circadian Photore-ceptor. Journal of Neuroscience 21 (2001) Nr. 16, S. 6405-6412.

Cajochen, C.; Münch, M.; Kobialka, S.; Kräuchi, K.; Steiner, R.; Oelhafen, P.;, Orgül, S.; Wirz-Justice, A.: High Sensitivity of Human Melatonin, Alertness, Thermoregulation, and Heart Rate to Short Wavelength Light. The Journal of Clinical Endocrinology & Metabolism 90 (2005) Nr. 3, S. 1311–1316.

Çakir, A: Strukturwandel im Büro und die Änderung des Bean-spruchungsprofils. In: Armbruster, A.; Çakir, A. (Hrsg.): Das Datensichtgerät als Arbeitsmittel. Berlin: ERGONOMIC Institut, 1979.

Çakir, A.; Çakir, G.: Forschungsbericht des Projekts »Licht und Gesundheit«. Berlin: Ergonomic Institut, 1998.

Desan, P.H.; Weinstein, A. J.; Michalak, E. E., Tam, E. M.; Meesters, Y.; Ruiter, M. J.; Horn, E.; Telner, J.; I skandar, H.; Boivin, D. B.; Lam, R. W.: A controlled trial of the Litebook light-emitting diode (LED) light therapy device for treatment of Seasonal Affective Disorder (SAD). BMC Psychiatry, 2007.

DIN Deutsches Institut für Normung: DIN EN 12464-1; Beleuchtung von Arbeitsstätten; Teil 1: Arbeitsstätten in Innenräumen. Berlin: Beuth, 2003.

DIN Deutsches Institut für Normung: DIN 5031-10; Strahlungs-physik im optischen Bereich und Lichttechnik - Teil 10: Photobiologisch wirksame Strahlung, Größen, Kurzzeichen und Wirkungsspektren. Berlin: Beuth, 2000.

Ehrenstein, W.: Das Auge stellt die biologische Uhr des Menschen – Konsequenzen für Lichtnutzung und Lichttechnik. Universität Hohenheim, Angewandte Physiologie. Tagung Licht und Gesundheit 2002, Berlin.

Figueiro, M.: Research Recap: Light, Aging & the Circadian System – Riviving ‘All that Jazz?’. In: Lighting Design + Application 33 (2002) Nr. 6, S. 8-11

Frieling, H.: Farbe am Arbeitsplatz. München: Bayerisches Staatsministerium für Arbeit und Sozialordnung, 1992.

Fisch, J.: Licht und Gesundheit – Das Leben mit optischer Strahlung. Technische Universität Ilmenau. Ilmenau: Eigenverlag, 2000.

Gall, D.: Die Messung circadianer Strahlungsgrößen. Tagung Licht und Gesundheit, 26.-27.2.2004, Berlin.

Gall, D.; Vandahl, C; Bieske, K.; Schmidt, A.; Hermann, R.; Blankenhagen, C.: Die Ermittlung von Licht- und Farbfeldgrößen zur Bestimmung der spektralen Wirkung des Lichtes. Tagung LICHT 2004, Dortmund.

Glickman, G.; Byrne, B.; Pineda, C.; Hauck, W. W.; Brainard, G.: Light therapy for seasonal affective disorder with blue narrow-band light-emitting diodes (LEDs). GC Biol Psychiatry 59 (2006) Nr. 6, S. 502-507.

Higuchi, S.; Motohashi, Y.; Liu, Y.; Ahara, M.; Kaneko, Y.; Effects of VDT tasks with a bright display at night on melatonin, core temperature, heart rate, and sleepiness. J Appl Physiol 94 (2003), S. 1773-1776.

Jacob, S.: Solid-state Lighting: Today and Tomorrow. Presentation to Energy Technologies for Sustainable Building Workshop 2006, Vancouver.

Jasser, S.; Hanifin, J.; Rollag, M.; Brainard, G.: Dim Light Adaptation Attenuates Acute Melatonin Suppression in Humans. J Biol Rhythms 21 (2006), S. 394.

Krüger, H.: Visionen für eine bessere Beleuchtung. Tagung Licht- und Lebensqualität, 19./20.4.2007, Lüneburg.

Lam, R.: Seasonal affective disorder and beyond: light treatment for SAD and non-SAD disorders. Washington, DC: American Psychiatric Press, 1998.

Liberman, J.: Die heilende Kraft des Lichts – Der Einfluss des Lichts auf Psyche und Körper. 6. Aufl. München, Piper, 2005

Lorincz, A.: The physiological and pahtological changes in skin from sunburn and suntan. In: Journal of the American Medical Association 173 (1960), S. 1227-1240.

Oltean, M.: Switchable Glass: A possible medium for Evolvable Hardware. NASA conference on Adaptive Hardware Systems, IEEE CS Press, 2006, S. 81-87

Rea, M.: Light – Much More Than Vision. In: Light and Human Health: EPRI/LRO 5th International Lighting Research Symposium. Palo Alto: The Lighting Research Office of the Electric Power Research Institute, 2002, S. 1-15.

Roenneberg, T.; Daan, S.; Merrow, M.: The art of entrainment. Journal of Biological Rhythms 18 (2003), S. 183-194.

Spath, D.; Braun, M.; Grunewald, P.: Gesundheits- und leistungsförderliche Gestaltung geistiger Arbeit – Arbeitsgestaltung unter Einbeziehung menschlicher Eigenzeiten und Rhythmen. Bielefeld: Schmidt, 2004.

Thapan, K.: An action Spectrum for melatonin suppression: evi-dence for a novel non-rod, non-cone photoreceptor system in humans. Journal of Physiology 535 (2001) Nr. 1, S. 261-267.

Vos J. J.; van Meeteren, A.: Resolution and contrast sensitivity at low luminances, Vision Research (1972), 12:825

Wandell, B. A.: Foundations of Vision. Sunderland: Sinauer, 1995.

Ware, C.: Information Visualization - Perception for Design. San Francisco: Morgan Kauffmann, 2000.

Wetterberg, L.: Light and biological rhythms in man. Stockholm: Pergamon, 1993.

Wirz-Justice, A.; Roenneberg, T.: Circadiane und Saisonale Rhythmen. Zentrum für Chronobiologie, Psychatrische Universitätsklinik, Basel; Zentrum für Chronobiologie, Institut für Medizinische Psychologie, Ludwig-Maximilians-Universität, München, 2004.

Wright, H. R.; Lack, L. C.; Partridge, K. J.: Light emitting diodes can be used to phase delay the melatonin rhythm. Journal of Pineal Research, 31 (2001) Nr. 4, S. 350-355.

Anlage 7

Fraunhofer IAO: Office 21-Studie: Information Work 2009

Fraunhofer
IAO

FRAUNHOFER-INSTITUT FÜR ARBEITSWIRTSCHAFT UND ORGANISATION IAO

Dieter Spath (Hrsg.), Jörg Kelter, Stefan Rief, Wilhelm Bauer, Udo-Ernst Haner

OFFICE 21®-STUDIE

INFORMATION WORK 2009

Über die Potenziale von Informations- und Kommunikationstechnologien bei Büro- und Wissensarbeit

Dieter Spath (Hrsg.), Jörg Kelter, Stefan Rief, Wilhelm Bauer, Udo-Ernst Haner

OFFICE 21®-STUDIE

INFORMATION WORK 2009

Über die Potenziale von Informations- und Kommunikationstechnologien bei Büro- und Wissensarbeit

IMPRESSUM

Herausgeber
Univ.-Prof. Dr.-Ing. Dr.-Ing. E.h. Dieter Spath

Autoren
Dr.-Ing. Jörg Kelter
Dipl.-Ing. Stefan Rief
Dr.-Ing. Wilhelm Bauer
Dipl.-Wirtsch.-Ing. Udo-Ernst Haner

Fraunhofer-Institut für Arbeitswirtschaft
und Organisation IAO, Stuttgart

Druck
Fraunhofer IRB
Nobelstraße 12, 70569 Stuttgart

Auslieferung und Vertrieb
Fraunhofer-Institut für Arbeitswirtschaft
und Organisation IAO
Nobelstraße 12
70569 Stuttgart
Tel.: +49 711 / 970 - 54 90
Fax: +49 711 / 970 - 54 61
joerg.kelter@iao.fraunhofer.de

Erscheinungsjahr
2009

ISBN-Nr.: 978-3-8396-0007-8

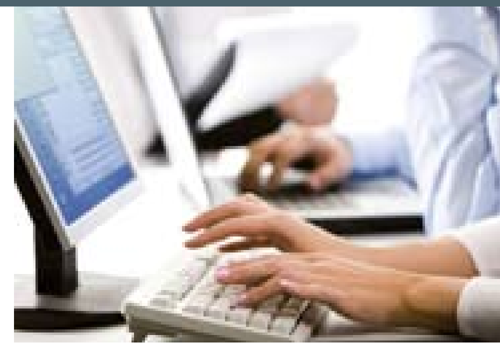
IAO-Shop
http://shop.iao.fraunhofer.de/

INHALT

VORWORT

Bild © pat fauve - Fotolia.com

ETWA 17 MIO. MENSCHEN IN DEUTSCHLAND ARBEITEN ÜBERWIEGEND IN BÜROS.



Dies entspricht einem Anteil von etwa 45% aller Erwerbstätigen. Dabei ist der Anteil der Beschäftigten, die auch im Umfeld von Fertigungs- und Produktionsanlagen in büroähnlichen Situationen arbeiten noch nicht vollständig berücksichtigt. Der tatsächliche Anteil von Menschen, die »Büroarbeit« ausüben, dürfte also noch um einiges höher liegen.

Während der Produktionssektor in den vergangenen Jahrzehnten erhebliche Produktivitätssteigerungen vollzogen hat, war in den Büros davon eher wenig zu spüren. Zwar haben sich mit der allgegenwärtigen Informations- und Kommunikationstechnik die Prozesse auch hier verändert, aber inwieweit Büroarbeit dadurch wirklich produktiver geworden ist, bleibt umstritten. Wenn überhaupt, dann hat sich die Produktivität bislang eher evolutionär entwickelt und erwartet jetzt den ganz großen Schub. Wissensarbeit produktiv zu machen, ist die große Managementaufgabe dieses Jahrhunderts. Insbesondere in turbulenten Zeiten wird deutlich, dass wir am Standort Deutschland nur dann erfolgreich sein können, wenn wir das geistige Potenzial der Büro- und Wissensarbeiter höchst effizient und effektiv einsetzen. Und dabei gilt es zu berücksichtigen, dass sich Büro- und Wissensarbeit zukünftig für immer mehr Menschen vielfältig vernetzt, mobil und nicht nur am eigenen Arbeitsplatz im Büro abspielen wird, sondern ebenso beim Kunden, im Hotel, am Flughafen oder auch zuhause. Die Anlässe und Orte der Arbeitserbringung werden vielfältiger.

Wenn wir im globalen Wettbewerb bestehen wollen, dann müssen wir uns an der Spitze orientieren, Bestleistungen erbringen und Innovationen in schneller Folge umsetzen. Aber diese entstehen nicht unbedingt inmitten von Papierbergen, an antiquierten Arbeitsplätzen und mit längst überholter Technik. Spitzenleistungen entstehen nur, wenn hochqualifizierte Menschen mit performanter und an die Prozessbedarfe angepasster Infrastruktur und Technik mit Engagement und Freude arbeiten können.

Im Mittelpunkt der aktuellen OFFICE 21®-Studie »Information Work 2009« steht daher insbesondere die Frage, unter welchen Voraussetzungen und in welchem Maße die informations- und kommunikationstechnische Ausstattung dazu beitragen kann, um Informations- und Wissensarbeiter sowohl als Individuum als auch im Team optimal zu unterstützen, deren ganze Produktivität und Kreativität zu entfalten und Herausforderungen, z. B. in Bezug auf eine zunehmende Mobilität und räumlich verteilte Arbeitsweisen zu meistern.

Bild © pressmaster - Fotolia.com

1 OFFICE 21®-STUDIE »INFORMATION WORK 2009«

1.1 HINTERGRUND

Im Verbundforschungsprojekt OFFICE 21® (www.office21.de) entwickelt das Fraunhofer IAO gemeinsam mit Partnerunternehmen auf Basis von Trendanalysen, wissenschaftlichen Studien sowie Best Practice Untersuchungen weitgreifende Szenarien und konkrete Lösungen zur Entwicklung innovativer und zukunftsorientierter Arbeits- und Bürowelten.

Die vorliegende empirische Studie »Information Work 2009« ist Bestandteil eines aktuellen Forschungsschwerpunktes im Verbundforschungsprojekt OFFICE 21® und baut auf den wissenschaftlichen Erkenntnissen und Forschungsfragen vorangegangener Projektphasen auf.

Die dargestellten Ergebnisse basieren auf der Auswertung der Angaben von 1.020 Studienteilnehmern, die im Zeitraum von Januar 2008 bis Dezember 2008 an der webbasierten Erhebung »Information Worker Check« (IWC) teilgenommen haben (www.iw.web-erhebung.de). Dieses Portal ist auch weiterhin frei zugänglich und soll für die Fortsetzung dieser Studienreihe auch zukünftig genutzt werden.

Wir bedanken uns an dieser Stelle bei allen Teilnehmern der Studie. Ebenso gilt Herrn Christian Köhler für sein großes Engagement bei der Auswertung der empirischen Daten und seine stetige Diskussionsbereitschaft in allen Phasen der Studie unser besonderer Dank.

Dank sagen wir vor allem aber allen Projektpartnern des Verbundforschungsprojektes OFFICE 21® für die Initiierung und Förderung dieser empirischen Studie.

Bild © Nmedia - Fotolia.com

1.2 ZIELSETZUNG UND VORGEHENSWEISE

„Wissensarbeit produktiv zu machen ist die große Managementaufgabe dieses Jahrhunderts, so wie es die große Aufgabe des vergangenen Jahrhunderts war, manuelle Arbeit produktiv zu machen."
Peter Drucker (1909-2005)

Ausgelöst durch technische Innovationen und neue Formen der Unternehmens- und Arbeitsorganisation wird in vielen Bereichen Arbeit insgesamt komplexer und stellt höhere Anforderungen an die Qualifikationen bzw. Kompetenzen der Beschäftigten. Betroffen sind hiervon nicht nur bestimmte Branchen oder Berufe, sondern nahezu alle Wirtschaftsbereiche. In der Folge ergeben sich insbesondere für Büro- und Wissensarbeit vielfältige Konzepte, Lösungen und Produkte, die es erlauben, ein sehr heterogenes und differenziertes Arbeits-, Technik-, Raum- und Büroumfeld zu schaffen.

Doch was brauchen Menschen, deren Aufgabe im Wesentlichen darin besteht, Wissen zu erwerben, zu erzeugen, zu bündeln oder anzuwenden wirklich? Insbesondere wenn man diese Fragestellung darauf fokussiert, welche informations- und kommunikationstechnische Ausstattung geeignet ist, diese Aufgaben erfolgreich und performant zu bewältigen, stellt man fest, dass hierzu wenige aktuelle Erkenntnisse vorliegen.

Wie also kann insbesondere Informationsarbeit – im englischen Sprachraum als »Information Work« bezeichnet – in Verbindung mit einer geeigneten informations- und kommunikationstechnischen Ausstattung produktiv gestaltet werden? Wie sieht die aktuelle und gängige Praxis aus, und wo lassen sich Wirkungszusammenhänge aufzeigen?

Im Fokus der vorliegenden empirischen Studie standen daher folgende Leitfragen:

- Welche informations- und kommunikationstechnische Ausstattung ist für welche unterschiedlichen Typen von Wissensarbeitern charakteristisch?
- Wie gut können unterschiedliche Technologien und Werkzeuge Informationsarbeit unterstützen?
- Welche IT-Anwendungen und Funktionen werden wie intensiv genutzt bzw. was davon ist für wen wie sinnvoll und unterstützt die Performance wirklich?
- Lassen sich generelle Wirkungszusammenhänge aufzeigen?

Ziel der OFFICE 21®-Studie »Information Work 2009« war es, diese Fragestellungen zu untersuchen, um insbesondere im Hinblick auf die Nutzung informations- und kommunikationstechnischer Lösungen, ein differenziertes Bild über Arbeitsweisen und Anwendungsverhalten unterschiedlicher Typen von Wissensarbeitern zu erhalten.

Als Grundlage der in Form einer empirischen Langzeit- und Trendstudie konzipierten Onlinebefragung haben Experten des Instituts dazu das Portal zum »Information-Worker-Check« als ein webbasiertes Selbstbewertungs-System für Büro- und Wissensarbeiter entwickelt (www.iw.web-erhebung.de).

Der »Information-Worker-Check« ist dabei als eine webbasierte Umfrage konzipiert, die jedem Befragungsteilnehmer eine direkte und einfache Möglichkeit bietet, die eigene Arbeitssituation unter den genannten Aspekten kritisch zu analysieren und zu bewerten. Darüber hinaus kann jeder Teilnehmer auf Basis seiner individuellen Ergebnisauswertungen sich stets mit aktuellen Benchmarks im Vergleich zu Anderen einordnen.

Da der »Information-Worker-Check« seit Ende 2008 auch in einer englischsprachigen Version verfügbar ist, sind über die vorliegende Studie hinaus zukünftig auch Auswertungen in einem internationalen Kontext möglich.

1.3 INHALTE UND AUSWERTUNGSBEREICHE

Der Fragenkatalog des »Information-Worker-Check« umfasst im Hinblick auf die o. g. Leitfragen folgende Themenblöcke:

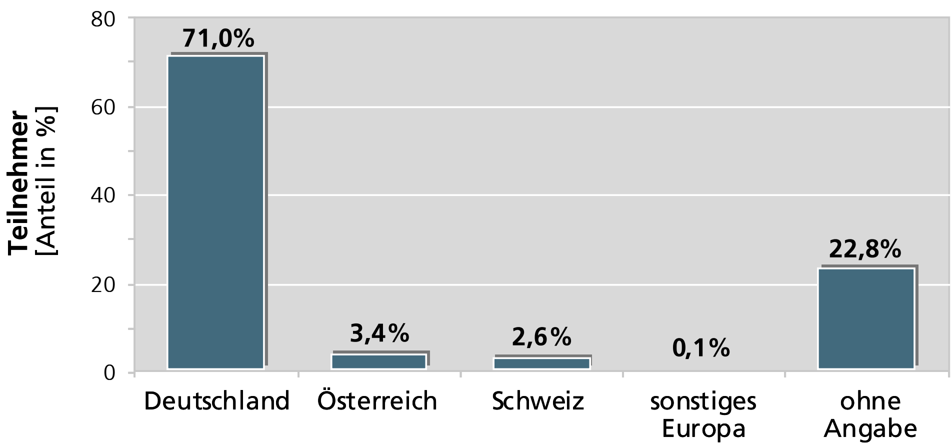
- Grundsätzliche Aspekte und Merkmale der Arbeit
- Mobilität und Flexibilität der Arbeit
- IuK-Ausstattung und deren Nutzung
- Verfügbare System-/Programmfunktionen
- Sicherheit und Datenschutz
- Bewertung der Arbeitsbedingungen
- Allgemeine Angaben

2 TEILNEHMERSTRUKTUR

Grundlage der nachfolgenden Auswertungen der empirischen OFFICE 21®-Studie »Information Work 2009« sind die seit Freischaltung des Portals von Januar 2008 bis Ende Dezember 2008 vorliegenden Daten von insgesamt 1.020 Teilnehmern.

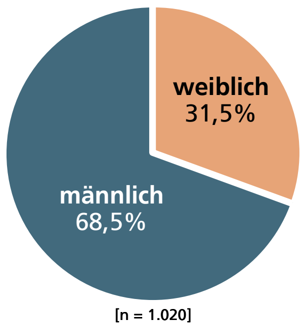
Hinsichtlich der regionalen Herkunft bzw. dem Sitz des Unternehmens, in dem die Teilnehmer beschäftigt sind, überwiegt erwartungsgemäß Deutschland mit 71% (Abb. 1). Auf Österreich, die Schweiz bzw. das restliche Europa entfallen insgesamt ca. 6% der Teilnehmer. Da diese Frage jedoch nicht von allen vollständig beantwortet wurde - 22,8% der Teilnehmer haben hierzu keine Angabe gemacht - sind diese Angaben als Relativwerte zu verstehen.

ABB. 1 *Teilnehmerstruktur nach regionaler Herkunft (Sitz des Unternehmens)*



Von den Teilnehmern sind insgesamt 31,5% weiblich und 68,5% männlich (Abb. 1). Unter Berücksichtigung der thematischen und technologieorientierten Grundausrichtung der Studie und bezogen auf den allgemeinen Anteil an weiblichen »Informations- und Wissensarbeitern«, die im Bereich Telekommunikation, Informationstechnologie und Medien tätig sind (ca. 26%), kann dies als eine angemessene Verteilung angesehen werden. Im Vergleich mit dem Anteil männlicher Beschäftigter in allen Büroberufen insgesamt (knapp 50%), sind Männer in der untersuchten Stichprobe stärker vertreten.

ABB. 2 *Teilnehmerstruktur nach Geschlecht*



Die Verteilung hinsichtlich der Altersstruktur der Teilnehmer ist erwartungskonform. Erfreulicherweise sind auch über 50-jährige mit einem Anteil von gut 20% vertreten. Die Gruppe der 30 – 49-jährigen bildet mit insgesamt ca. 68% der Teilnehmer den Schwerpunkt. Auf die unter 30-jährigen entfallen knapp 12% (Abb. 3).

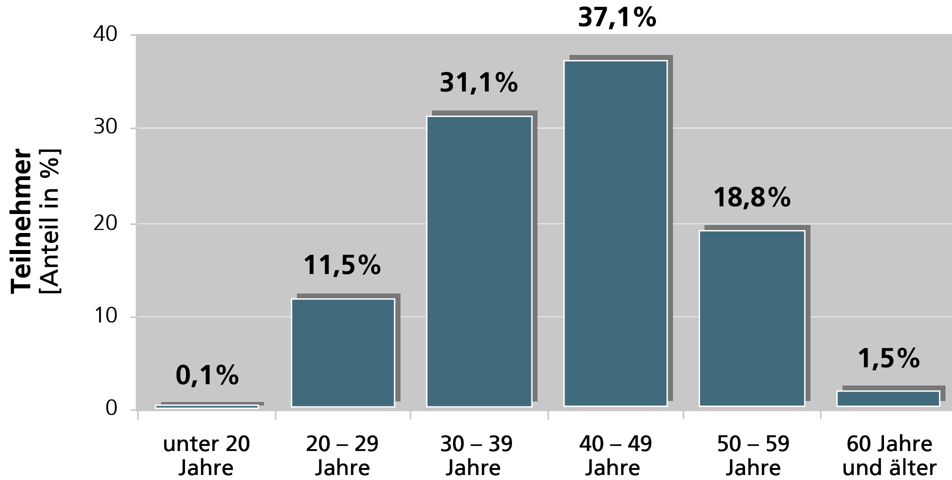


ABB. 3 *Teilnehmerstruktur nach Alter*

Bei der Frage nach der Personalverantwortung zeigt sich, dass die Zielgruppe der Entscheider im mittleren und oberen Management mit einem Anteil von insgesamt ca. 22% sehr gut erreicht werden konnte (Abb. 4). Die Mehrheit der Teilnehmer verfügt über »keine Personalverantwortung« (54,1%) bzw. ist z. B. als Teamleitung für einen kleineren Bereich verantwortlich (24,1%).

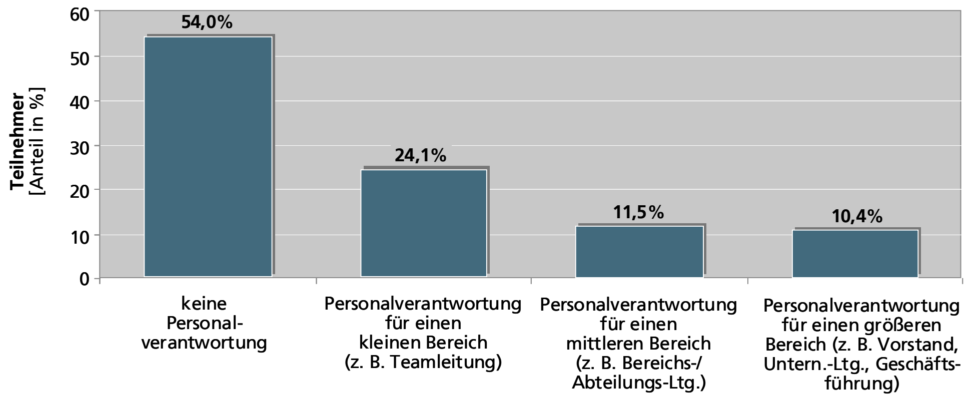
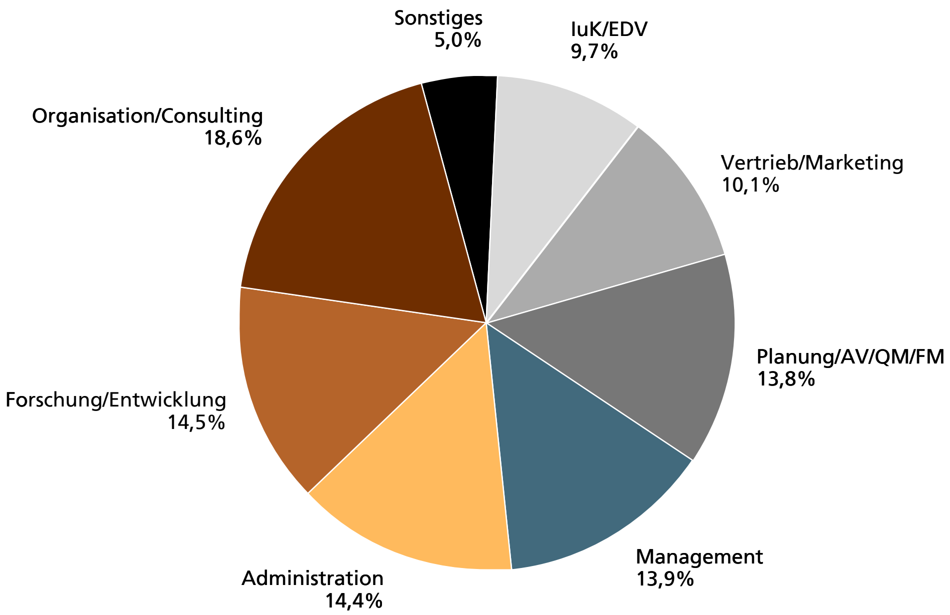


ABB. 4 *Teilnehmerstruktur nach Personalverantwortung*

Fasst man einzelne Aufgaben- und Tätigkeitsbereiche wie in Abb. 5 dargestellt zusammen, so ergibt sich eine breite Streuung unterschiedlicher Ausprägungen und Formen von Informations- und Wissensarbeit.

ABB. 5 *Teilnehmerstruktur – verdichtet nach Aufgaben- und Tätigkeitsclustern*



Neben dem Tätigkeitscluster Organisation/Consulting mit einem Anteil von 18,6% der Teilnehmer, entfallen auf F&E, Administration, Management und planerische Aufgaben jeweils ca. 14% der Nennungen. Aktivitäten im Vertrieb/Marketing (10,1%) sowie Informations- und Kommunikationstechnik/EDV/Rechenzentrum (9,7%) runden das Spektrum ab.

Bei der Aufschlüsselung der Teilnehmer nach Branchen überwiegen Industrie- und Dienstleistungsunternehmen mit einem Gesamtanteil von knapp 70%.

Auf den Bereich öffentlicher Einrichtungen und Verbände entfällt ein Anteil von 14,9%. Mit jeweils ca. 12% sind die Bereiche Maschinenbau/Investitions- und Konsumgüter, Beratung/Kommunikation sowie IT/Medien/Telekommunikation als gewichtige Branchen vertreten (Abb. 6).

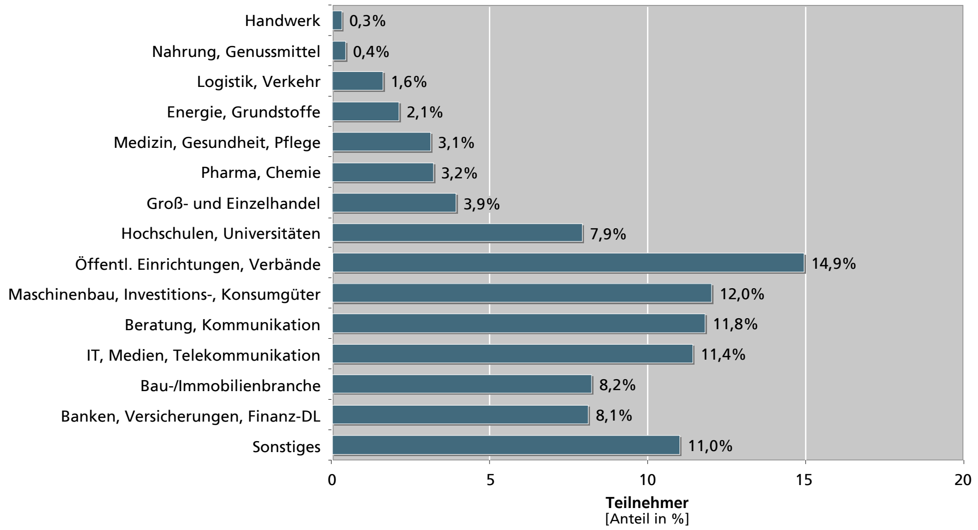


ABB. 6 *Teilnehmerstruktur nach Branchen*

Hinsichtlich der Unternehmensgröße sind kleinere und mittlere Unternehmen (kmU´s) repräsentativ vertreten - ca. 60% der Teilnehmer arbeiten in Unternehmen mit einer Beschäftigtenzahl unter 500 Personen (Abb. 7). Großunternehmen mit bis zu 2.000 Beschäftigen finden sich mit einem Anteil von 14,6% wieder. In noch größeren Konzernen sind insgesamt 27,7% der Teilnehmer tätig.

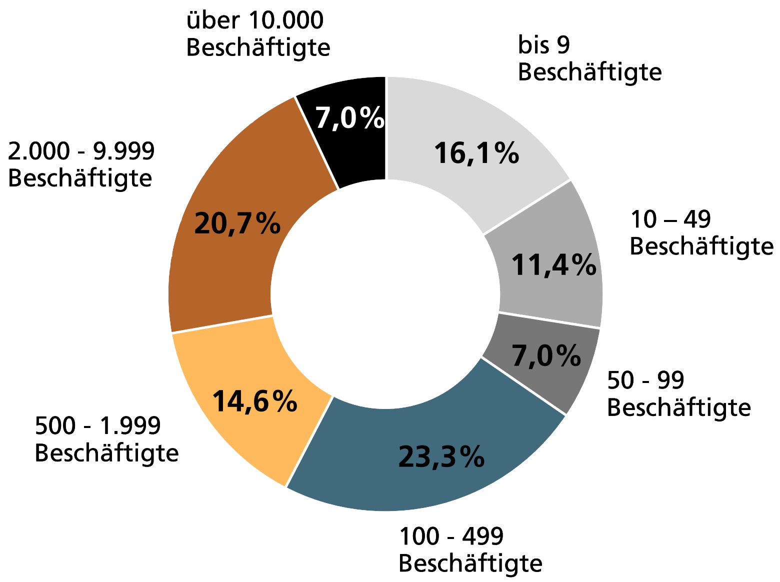


ABB. 7 *Teilnehmerstruktur nach Unternehmensgröße (Anzahl Beschäftigte)*

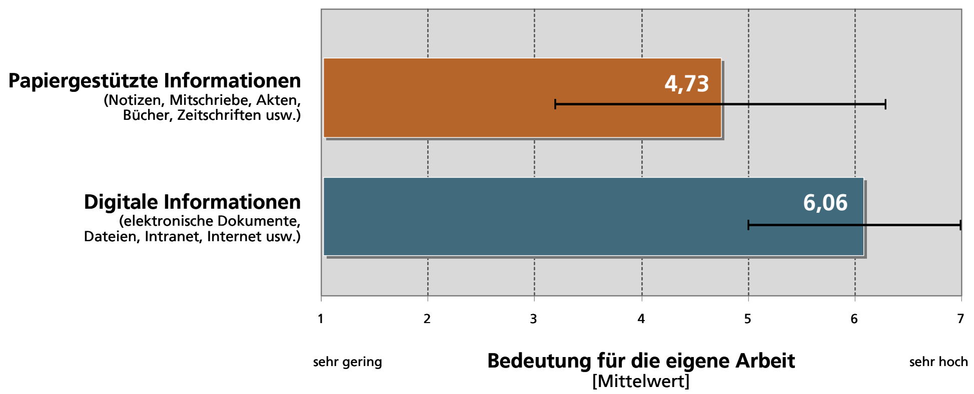
3 ERGEBNISSE

3.1 EXPLORATIVE ANALYSEN UND GESAMTAUSWERTUNGEN

3.1.1 Digitalisierung

Informationen werden nach wie vor über ganz unterschiedliche Kanäle und Wege verteilt und festgehalten. Wie aus Abb. 8 hervorgeht, haben jedoch hinsichtlich der Bedeutung für die eigene Arbeit digitale Informationen einen sehr viel höheren Stellenwert (MW=6,06) im Vergleich zu papiergestützten Informationen (MW=4,73).

ABB. 8 *Bedeutung papiergestützter Informationen vs. digitaler Informationen für die tägliche Arbeit*



Um eine Orientierung darüber zu erhalten, in welcher Form heutzutage Wissensarbeiter auf benötigte Informationen zugreifen bzw. in welcher Form sie ihre Daten, Informationen und Unterlagen aufbereitet haben, zeigt die in Abb. 9 dargestellte Verteilung des »Digitalisierungsgrades« eine eindeutige Tendenz auf.

Das digitale Zeitalter ist mittlerweile Realität in den meisten Büros. Nur bei 3,6% der Befragten dominiert Papier noch als typischer Informationsträger. Bei 41,2% der Teilnehmer hingegen ist »sehr viel digitalisiert«, d. h. mehr als ¾ aller benötigten Daten liegen in digitaler Form vor. Bei weiteren 41,4% ist zumindest mehr als die Hälfte aller Daten in digitaler Form verfügbar.

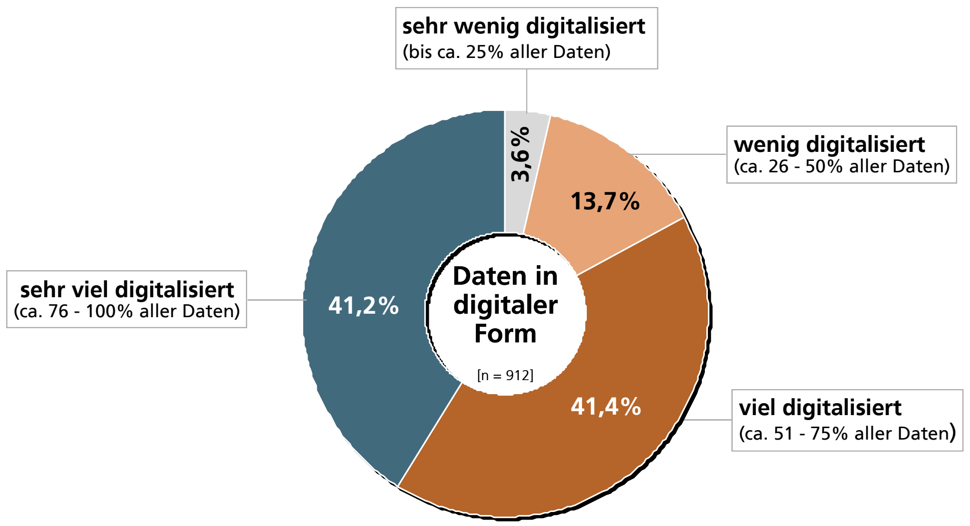


ABB. 9 *Anteil an Daten, Informationen und Unterlagen in digitalisierter Form (Häufigkeitsverteilung)*

Zwingende Rückschlüsse, inwieweit damit jedoch immer auch auf eine Papierablage verzichtet wird bzw. in welchem Umfang Dubletten sowohl in digitaler Form als auch in Papier vorgehalten werden, lassen sich damit jedoch noch nicht ziehen. Ein hohes Potenzial und gute Voraussetzungen zur Reduzierung des tatsächlichen Ablage- und Stauraumbedarfs lassen sich daraus jedoch sicher ableiten.

3.1.2 IuK-Ausstattung und deren Nutzen

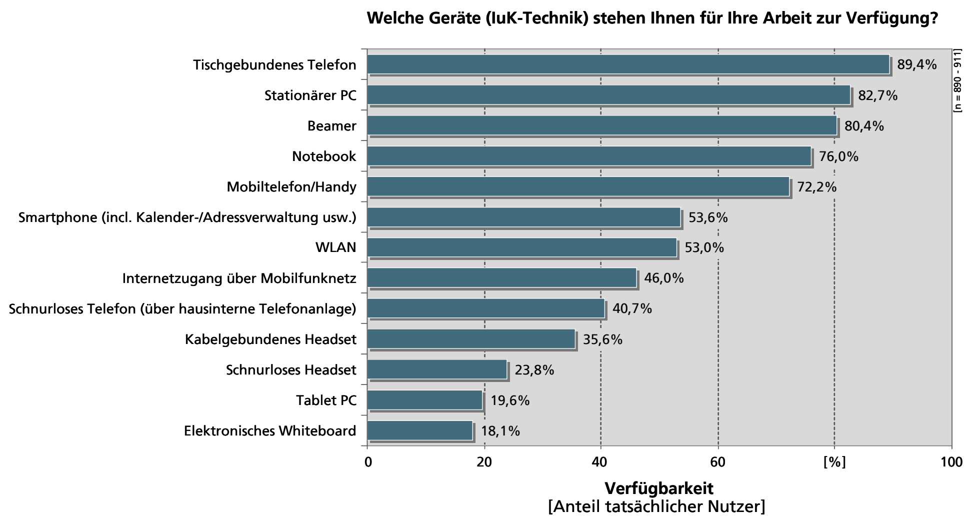
Wie die Gesamtauswertung zeigt, ergibt sich in Bezug auf die IuK-Ausstattung ein heterogenes Bild. Insgesamt betrachtet, gehören erwartungsgemäß insbesondere Telefon und Computer zu den typischen Arbeitsmitteln von Informations- und Wissensarbeitern.

Für die allermeisten der Befragten wird das Arbeitsmittel »Telefon« in Form eines tischgebundenen Apparates (89,4%) zur Verfügung gestellt. Daneben werden jedoch von vielen auch Mobiltelefon/Handy (72,2%), Smartphone (53,6%) sowie schnurlose Telefone, die als Komponente einer hausinternen Telefonanlage zur Verfügung stehen (40,7%), genutzt (Abb. 10).

Stationäre PC (82,7%) und Notebooks (76,0%) stehen bei vielen Anwendern offensichtlich parallel zur Verfügung. Tablet PC kommen derzeit bei insgesamt knapp 20% der Befragten zum Einsatz.

Darüber hinaus gehören Beamer (80,4%), WLAN (53,0%) oder der Internetzugang über ein Mobilfunknetz (46%) für einen Großteil der Befragten zum Alltag.

ABB. 10 *Für die eigene Arbeit zur Verfügung stehende IuK-Ausstattung und Anteil tatsächlicher Nutzer (Gesamtauswertung)*



Bei den zur Verfügung stehenden IT-Funktionen/Technologien sind insbesondere Teamserver (85,3%), Teamkalender (71,1%), der Zugang via Webportal auf eigene bzw. benötigte Daten im Unternehmensnetzwerk (59,9%), Groupware-Lösungen (55,3%) und Dokumenten-Management-Systeme (48,9%) weitverbreitete Anwendungen, die zur Verfügung stehen (Abb. 11).

Bei der Frage nach dem tatsächlichen Nutzen für die eigene Arbeit ergibt sich aus der Sicht derer, die diese Funktionen auch tatsächlich anwenden, das in Abb. 12 dargestellte Meinungsbild. Der höchste Nutzen wird Teamservern (MW=5,44), Groupware-Lösungen (MW=4,86) und Teamkalendern (MW=4,72) attestiert.

Auffällig – und überraschenderweise am Ende dieser Liste – stehen aus Sicht der Anwender Videokonferenzen mit einem deutlich unterdurchschnittlichen, geringen Nutzen (MW=2,47). Inwieweit hier möglicherweise technische Gründe (Bedienerfreundlichkeit, Qualitäts-/Ausstattungsmerkmale usw.), organisatorische oder monetäre Aspekte ausschlaggebend für dieses Urteil gewesen sein mögen, wäre lohnend an anderer Stelle noch tiefergehend zu untersuchen.

ABB. 11 *Für die eigene Arbeit zur Verfügung stehende IT-Funktionen/Technologien und Anteil tatsächlicher Nutzer (Gesamtauswertung)*

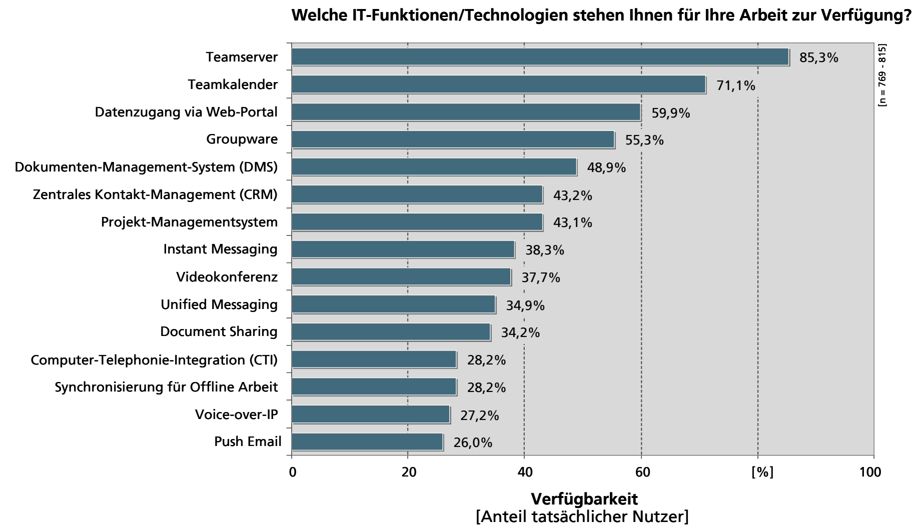
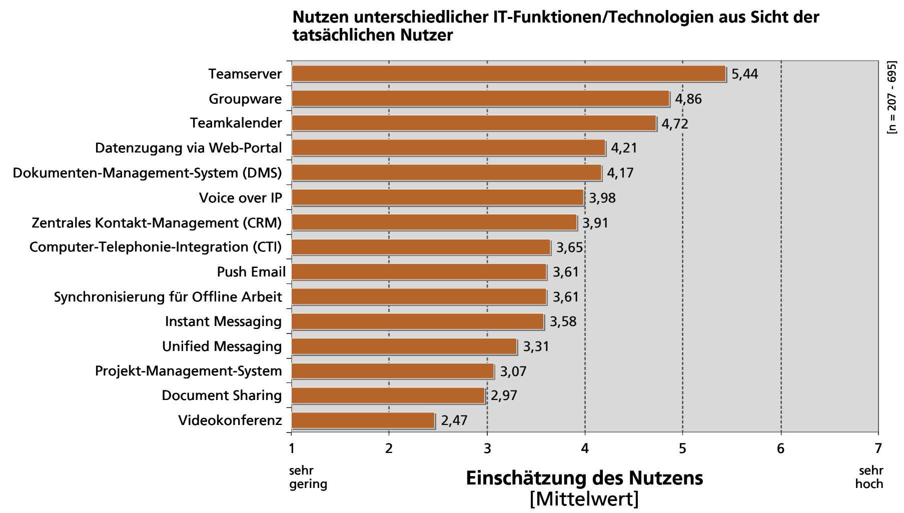


ABB. 12 *Nutzen unterschiedlicher IT-Funktionen/Technologien aus Sicht der tatsächlichen Nutzer (Gesamtauswertung)*



Da diese ersten orientierenden Analysen auf einer Gesamtauswertung aller Teilnehmer basieren und ohne Berücksichtigung individueller Tätigkeitsschwerpunkte und typischer Arbeitsmerkmalen vorgenommen worden sind, soll nachfolgend eine weitere Differenzierung anhand geeigneter Typologien für Informations- und Wissensarbeiter vorgenommen werden.

3.2 ZUR ABGRENZUNG UND DEFINITION VON WISSENSARBEIT

Das Themenfeld »Wissensarbeit« – im englischen Sprachraum als »Knowledge Work« bezeichnet – und in Abgrenzung zur »Task Work« bzw. sachbearbeitungsorientierten Tätigkeiten, ist insbesondere in der letzten Dekade zum Gegenstand organisations-soziologischer Forschung geworden. Übereinstimmend kann festgehalten werden, dass wissensbasierte Arbeitsprozesse weiter zugenommen haben und als Schlüssel zum Erfolg unserer entwickelten Wirtschaft angesehen werden. Folgerichtig gelten Wissensarbeiter, ihre Erfahrungen und ihr Know-how als entscheidende Aktivposten, über die viele Unternehmen heute verfügen.

Charakteristisch für Wissensarbeit ist, dass diese häufig komplex, wenig determiniert und folglich schwer in vorgegebenen Abläufen standardisierbar ist. Wissensarbeit ist hochgradig sowohl personen- aber auch kommunikationsorientiert und wird immer mehr in übergreifenden Teams erbracht. Wissensarbeit schafft ständig neues Wissen und baut auf Erfahrungen Anderer auf. Dabei agieren Wissensarbeiter stark autonom und sind somit wenig direkt »anleitbar«. Darüber hinaus stellt Wissensarbeit neue Anforderungen an die Arbeitsprozessorganisation, betriebliche Steuerungssysteme, die Gestaltung der Arbeitsplätze bzw. der Büroumgebung insgesamt und nicht zuletzt an die Führung und Motivation von Mitarbeitern.

Wie bereits eingangs erwähnt, wird in diesem Zusammenhang Informationsarbeit als ein Teilaspekt von Wissensarbeit verstanden, der insbesondere mit der Anwendung von informations- und kommunikationstechnischer Ausstattung in Verbindung gebracht werden kann.

Zur weiteren Abgrenzung und Definition sind folgende Merkmale für Wissensarbeit kennzeichnend:

- Wissensarbeit ist ergebnis- und prozessoffen, komplex und wenig standardisiert,
- Wissensarbeit ist vorwiegend »Kopfarbeit«, immateriell und schwierig »von Außen« zu überprüfen oder zu kontrollieren,
- Wissensarbeit ist ganz wesentlich »people business«, entsteht also in direkter Kollaboration und Kommunikation mit anderen Menschen,
- Wissensarbeit umfasst Informationssuchen, Problemanalysen, Verhandlungen, die Beobachtung und den direkten Kontakt mit Menschen in immer neuen Ausgangssituationen (»tacit interactions«),
- durch das Einbringen, Hinterfragen, Revidieren, Bewerten und Verknüpfen von Wissen entsteht wiederum neues Wissen,
- Wissensarbeit steht damit in einem Spannungsfeld zwischen konzentrierter Einzelarbeit und Gruppen-/ Teamarbeit,
- Ihre erfolgreiche Ausführung ist stark von der Motivation des Wissensarbeiters und zugestandenen bzw. nutzbaren Freiräumen abhängig.

Damit aber wird bei Wissensarbeitern insbesondere die Ausrichtung auf deren Motivation, Unterstützung und Integration in die strategische Gesamtausrichtung des Unternehmens zu einer wesentlichen Führungsaufgabe.

Aufbauend auf diesem Grundverständnis lässt sich Wissensarbeit demzufolge entlang der drei Grunddimensionen Komplexität, Neuartigkeit und Autonomie beschreiben (Abb. 13).

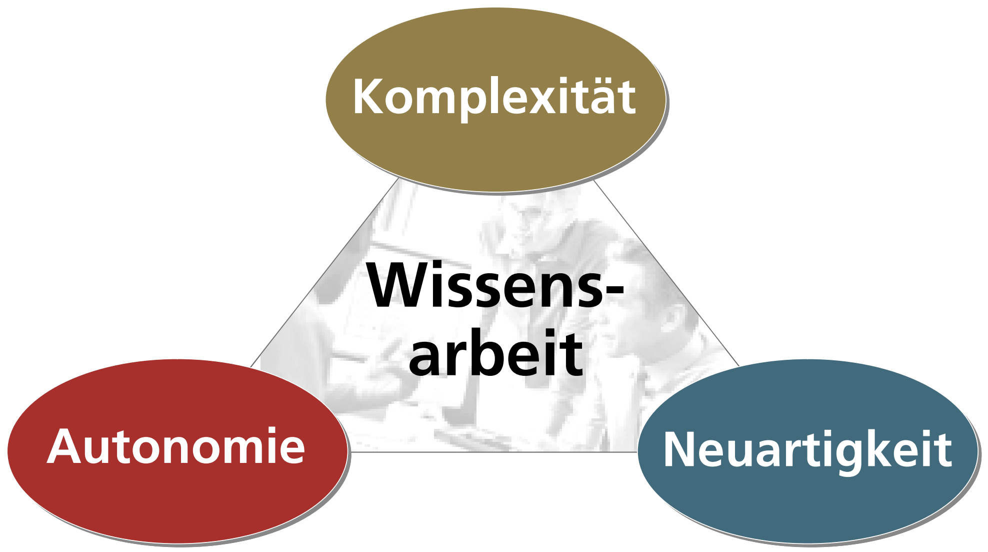


ABB. 13 *Grunddimensionen von Wissensarbeit*

3.4 MODELL UND VORGEHENSWEISE ZUR IDENTIFIKATION UNTERSCHIEDLICHER WISSENSARBEITSTYPEN

Wie eingangs bereits ausgeführt, zielte eine der zentralen Fragestellungen der OFFICE 21®-Studie »Information Work 2009« darauf ab zu untersuchen, ob und inwieweit sich Unterschiede in der informations- und kommunikationstechnischen Ausstattung und deren Nutzung in Abhängigkeit von bestimmten Tätigkeitsschwerpunkten und typischen Arbeitsmerkmalen feststellen lassen.

Als Ausgangspunkt für die Beantwortung dieser Fragestellungen galt es daher zunächst zu untersuchen, wo und wie sich aufgrund unterschiedlicher inhaltlicher Ausprägungen geeignete Typologien von Wissensarbeit finden lassen.

Um diese Unterschiede identifizieren zu können, wurde ein Schwerpunkt der empirischen Erhebung auf die Abfrage zahlreicher Einzelmerkmale typischer Informations- und Wissensarbeit gelegt. Dies umfasst z. B. Fragen nach dem Grad der Neuartigkeit der Arbeitsaufgaben oder deren Ausmaß an Vielfältigkeit. Die abgefragten Einzelmerkmale der Arbeit wurden in Anlehnung an die in Abb. 13 aufgezeigten Grunddimensionen von Wissensarbeit und auf Basis einer faktoranalytischen Überprüfung zu merkmaltypischen Indizes zusammengefasst.

Diese Indizes bilden die Grunddimensionen von Wissensarbeit ab. In einem weiteren Schritt wurden die ermittelten Dimensionen einer Clusteranalyse unterzogen, um auf diese Weise typische Kombinationen der Ausprägungen dieser Kennwerte für unterschiedliche Typen von Wissensarbeit identifizieren und beschreiben zu können (Abb. 14).

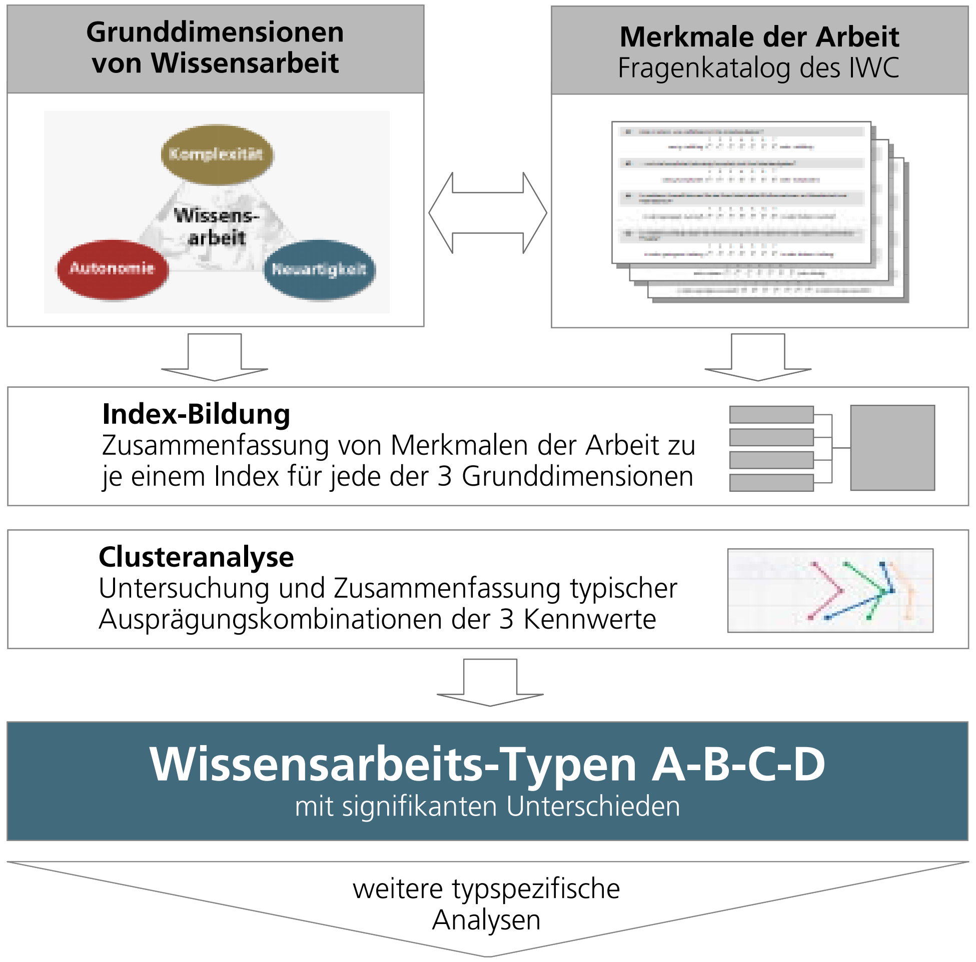


ABB. 14 *Modell und Vorgehensweise zur Identifikation unterschiedlicher Wissensarbeits-Typen*

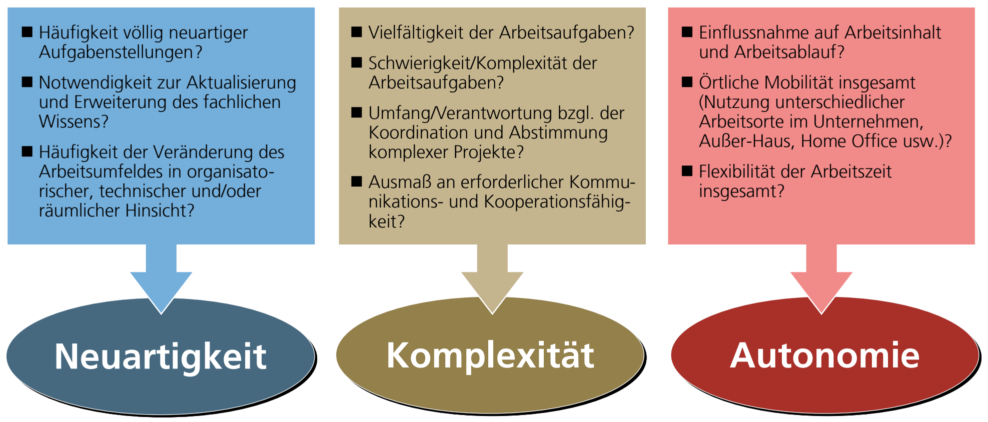
Die jeweiligen Einzelmerkmale, aus denen sich die drei Indizes zur Beschreibung der Grunddimensionen von Wissensarbeit zusammensetzen, sind in Abb. 15 aufgeführt. Durch die additive Zusammenfassung der einzelnen Merkmalsausprägungen, denen stets eine 7-stufige Antwortskala zugrunde liegt (1 = sehr gering; 7 = sehr hoch), ergibt sich der entsprechende Gesamtwert des jeweiligen Index.

Der Index über die Dimension **Neuartigkeit** der Arbeit beinhaltet neben dem Aspekt der Häufigkeit von sich verändernden und völlig neuartigen Aufgabenstellungen auch die Frage nach der Notwendigkeit, das eigene fachliche Wissen ständig zu erweitern sowie die Häufigkeit von Veränderungen des Arbeitsumfelds – sowohl in organisatorischer, technischer als auch in räumlicher Hinsicht.

Die Dimension **Komplexität** der Arbeit wird als Index aus vier Arbeitsmerkmalen abgebildet. Eine hohe Arbeitskomplexität ist demnach durch einen hohen Grad an vielfältigen und schwierigen Arbeitsaufgaben, ein hohes Ausmaß an Koordination und verantwortlicher Abstimmung komplexer Projekte sowie durch hohe Anforderungen an die Kommunikations- und Kooperationsfähigkeit insgesamt gekennzeichnet.

Neben dem Grad der eigenständigen Einflussnahme auf Arbeitsinhalte und Arbeitsabläufe greift der Index für die Dimension **Autonomie** zudem die Merkmalsausprägungen in Bezug auf die örtliche/räumliche Mobilität sowie die Flexibilität der Arbeitszeiten insgesamt auf. Es geht dabei also im weitesten Sinne um den Grad der autonomen Gestaltung von Arbeit.

ABB. 15 *Zusammensetzung der Indizes zur Beschreibung der Grunddimensionen von Wissensarbeit*



Eine Gesamtauswertung dieser drei arbeitsbezogenen Indizes über alle Teilnehmer hinweg zeigt, dass diesbezügliche Anforderungen generell eher hoch eingeschätzt werden (Abb. 16). Alle Indizes weisen insgesamt betrachtet jeweils überdurchschnittliche Ausprägungen und Mittelwerte auf (MW=4,77 - 5,48).

Büroarbeit – respektive Wissensarbeit – wird insgesamt also eher als abwechslungsreiche, deutlich komplexe und relativ autonome Arbeitsform erlebt.

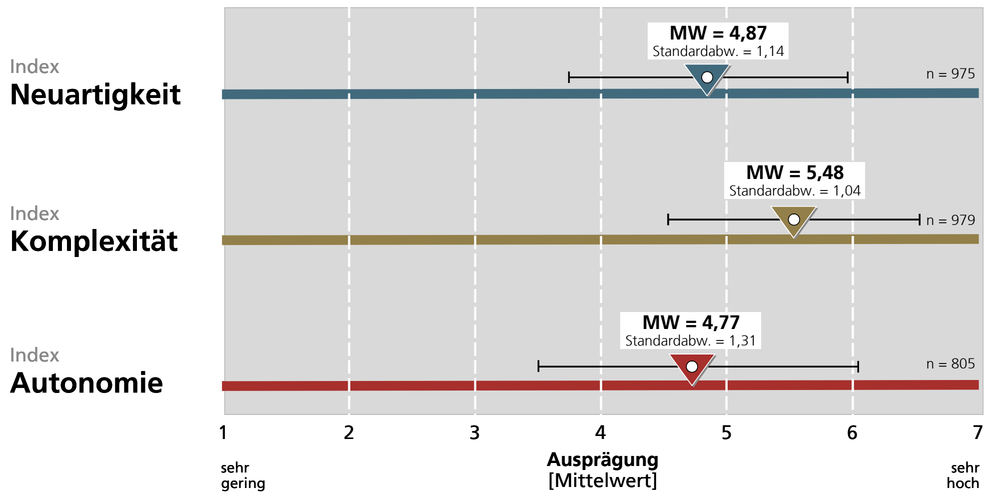


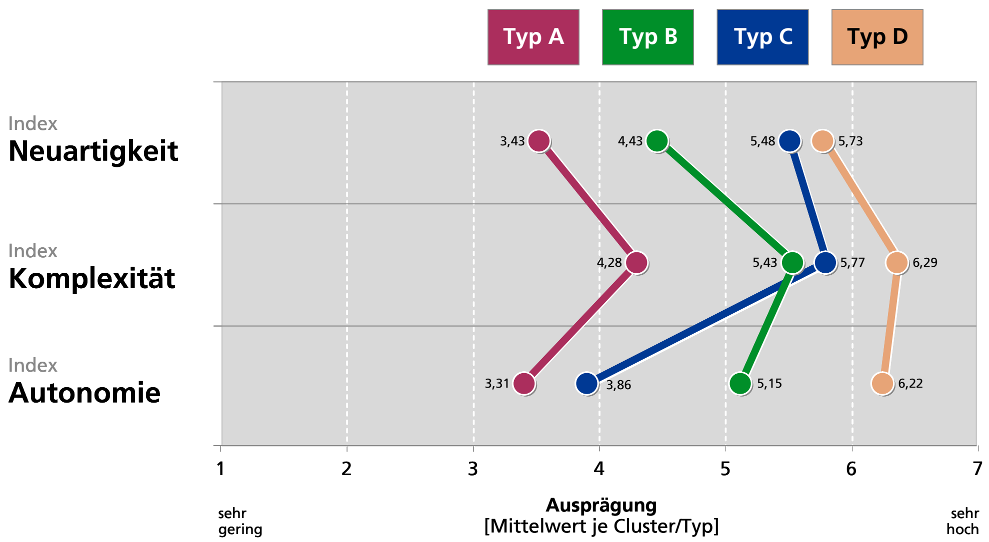
ABB. 16 *Gesamtergebnis in Bezug auf die Grunddimensionen von Wissensarbeit entlang der drei arbeitsbezogenen Indizes (Mittelwerte und Standardabweichungen)*

3.4 MERKMALE DER VIER UNTERSCHIED LICHEN WISSENSARBEITSTYPEN

Zur weiteren Differenzierung konnten auf Basis einer vertiefenden Clusteranalyse vier unterschiedliche Typen von Wissensarbeit identifiziert werden.

Die jeweils spezifischen Ausprägungen zu den betrachteten Arbeitsdimensionen Neuartigkeit, Komplexität und Autonomie sowie deren Kombination in Form von Mittelwertprofilen sind charakteristisch für die vier identifizierten Wissensarbeits-Typen (Abb. 17).

ABB. 17 *Mittelwertprofile der vier unterschiedlichen Wissensarbeits-Typen zu den betrachteten Grunddimensionen im Vergleich*



Ein erstes Cluster – im Folgenden als **Typ A** bezeichnet – ist durch mittlere Werte in allen drei Arbeitsdimensionen gekennzeichnet (MW=3,31 - 4,28). Dieses Profil beschreibt damit Formen »wissensbasierter Arbeit«, also Tätigkeiten, bei denen Erfahrung und Wissen durchaus eine wichtige Rolle spielen, Entscheidungsspielräume aber eher begrenzt sind und im Wesentlichen bekannte und standardisierte Prozesse mit merklichen Routineanteilen vorherrschen (z. B. Assistenzaufgaben).

Der **Typ B** steht für einen Personenkreis, der Informations- und Wissensarbeit normalerweise mit relativ großer Autonomie bewältigen kann und dessen Aufgabenspektrum sich durch eine hohe Komplexität bei mittlerem Neuartigkeitsgrad auszeichnet (MW=4,43 - 5,43). Dieses Profil beschreibt damit Formen »wissensintensiver Arbeit«, also Tätigkeiten, die i. d. R. eine umfassende Ausbildung bzw. langjährige Erfahrung in einem bestimmten Fachgebiet voraussetzen (z. B. Fachtätigkeiten, Spezialisten).

Kennzeichnend für Wissensarbeiter des **Typ C** ist sowohl ein recht hoher »Schwierigkeitsgrad« bezüglich der Neuartigkeit als auch der Komplexität ihrer Aufgaben (MW=5,48 - 5,77). Auch dieses Profil beschreibt damit ausgeprägte Formen »wissensintensiver Arbeit«. Andererseits verfügt dieser Typus jedoch nur über einen beschränkten Freiraum und Autonomie (MW=3,86), um eigenständig über das »wie, wann und wo« zur Bewältigung dieser Aufgaben zu entscheiden (z. B. Mitarbeiter in Laboren oder an Versuchsständen, die an definierte Abläufe gebunden sind).

Dem **Typ D** lassen sich Personen zuordnen, deren Tätigkeitsspektrum sich durch eine sehr häufige Auseinandersetzung mit neuartigen und komplexen Aufgaben beschreiben lässt und die gleichermaßen über ein hohes Maß an Autonomie verfügen (MW=5,73 - 6,29). Dieses Profil beschreibt damit Formen von »Wissensarbeit im engeren Sinne«, also Tätigkeiten, bei denen das einmal erworbene Fachwissen nicht ausreicht, sondern die es erforderlich machen, das vorhandene Wissen auch stets zu revidieren, zu verbessern und zu erneuern, um Problemlösungen zu finden (z. B. Consulting, F&E, Wissenschaft).

Die vier Wissensarbeits-Typen sind relativ gleichmäßig in der untersuchten Stichprobe der Studie verteilt. Der Typ A wird von 17,1% der Teilnehmer repräsentiert; auf die weiteren drei Typen B, C und D entfallen jeweils knapp 27% (Abb. 18).

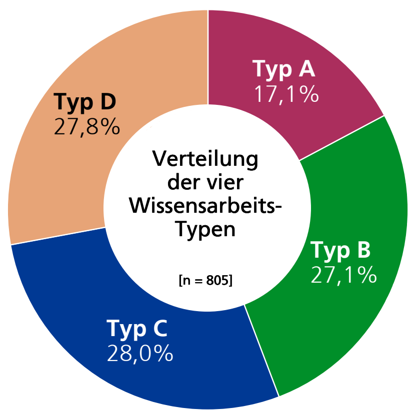


ABB. 18 *Verteilung der identifizierten vier Wissensarbeits-Typen in der untersuchten Stichprobe der Studie*

Inwieweit sich die vier Wissensarbeits-Typen hinsichtlich weiterer Kennwerte und Indizes, z. B. in Bezug auf Wohlbefinden, Zufriedenheit und Prozess-Performance unterscheiden, welche IuK-Ausstattungsprofile jeweils typisch sind bzw. welche Nutzenpotenziale sich aus den bevorzugt eingesetzten IT-Anwendungen ableiten lassen, wird nachfolgend diskutiert.

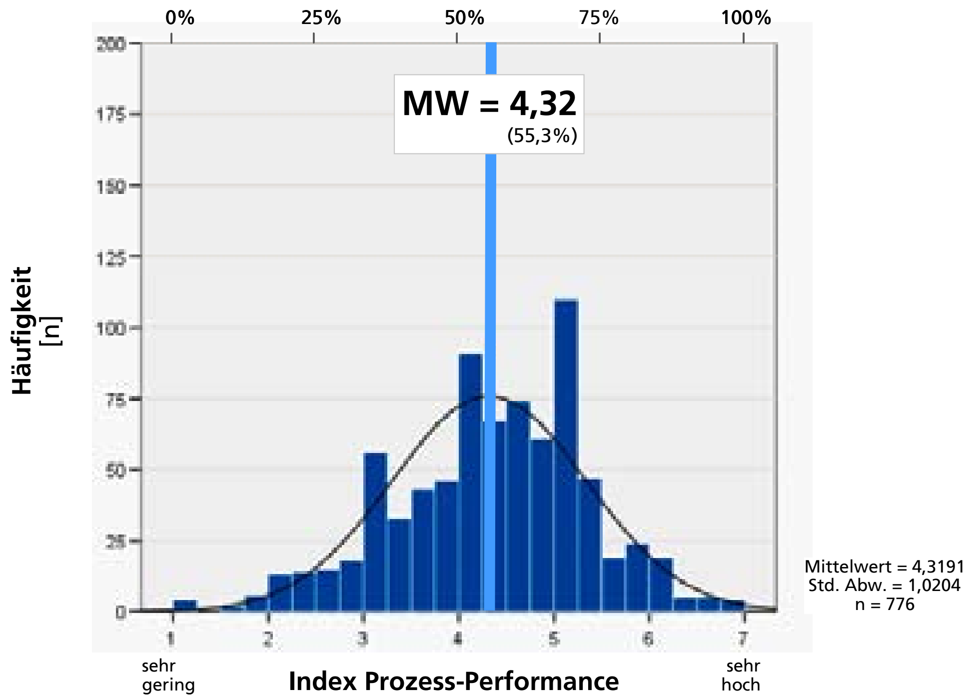
3.4.1 Prozess-Performance

Der Index »Prozess-Performance« umfasst fünf Bewertungen zu unterschiedlichen Prozess-, Effizienz- und Effektivitätsmerkmalen, nämlich

- Nutzung der richtigen Mittel und Wege zur Zielerreichung,
- Aufwand zur Zielerreichung bzw. zur Erledigung der eigenen Aufgaben,
- Bewertung der arbeitsbezogenen Kommunikation mit den Kollegen,
- Qualität der internen/teambezogenen Zusammenarbeit insgesamt,
- Abstimmung allgemeiner organisatorischer Prozesse.

Der aus den Angaben aller Teilnehmer ermittelte Gesamtwert zur Ermittlung einer spezifischen Kennzahl für die Prozess-Performance von Wissensarbeitern (»Index Prozess-Performance«) liegt im Durchschnitt aktuell bei einem Mittelwert von MWiPerf = 4,32 (Standardabweichung 1.02). Die Verteilung über die einzelnen Wertebereiche dieses Index ist in Abb. 19 dargestellt.

ABB. 19 *Gesamtverteilung des Index Prozess-Performance über einzelne Werteklassen und Mittelwert insgesamt*



Überträgt man diesen Mittelwert von der hinterlegten 7er-Skala auf eine Prozentskala (Wert 1 = 0%; Wert 7 = 100%), so entspricht dies – bezogen auf den möglichen Maximalwert – einem Gesamtniveau der Prozess-Performance bei Wissensarbeitern von nur 55,3%.

Der insgesamt recht niedrige Kennwert weist deutlich auf ein nach wie vor vorhandenes, aber offensichtlich häufig ungenutzt schlummerndes Performance- und Produktivitätspotenzial bei Büro- und Wissensarbeit hin.

Die in Abb. 20 dargestellte Übersicht zeigt, dass sich der Aspekt »Aufwand zur Zielerreichung« als die am schlechtesten bewertete Komponente des Index erweist. Im Durchschnitt konnten nur gut 50% aller Arbeitsziele mit dem veranschlagten bzw. geringst möglichen Aufwand erreicht werden. In allen anderen Fällen musste dagegen aus Sicht der Befragten mehr Aufwand betrieben werden, als der Sache eigentlich angemessen gewesen wäre.

Des Weiteren fällt auf, dass in der Tendenz häufig zwar die richtigen Mittel und Wege zur Zielerreichung genutzt werden (MW=4,56). Dennoch lässt sich bei allen Teilaspekten des Index ein deutlicher Verbesserungsbedarf erkennen. Nicht nur die Qualität der Arbeitsprozesse und die Zusammenarbeit innerhalb von Teams (MW=4,24), auch die Abstimmung allgemeiner organisatorischer, administrativer und formaler Prozesse (MW=4,29) oder die direkte, arbeitsbezogene Kommunikation mit Kollegen (MW=4,45) funktionieren selten wirklich gut.

Wenn die Performance von Wissensarbeitern also signifikant erhöht werden soll, gilt es zunächst insbesondere räumliche, technische und organisatorische Hemmnisse zu beseitigen bzw. diesbezügliche Verbesserungspotenziale zu identifizieren.

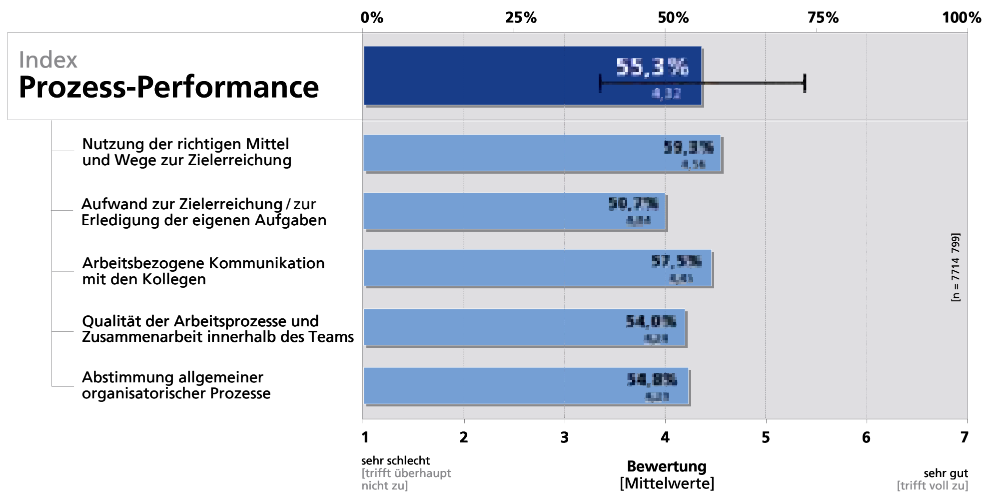
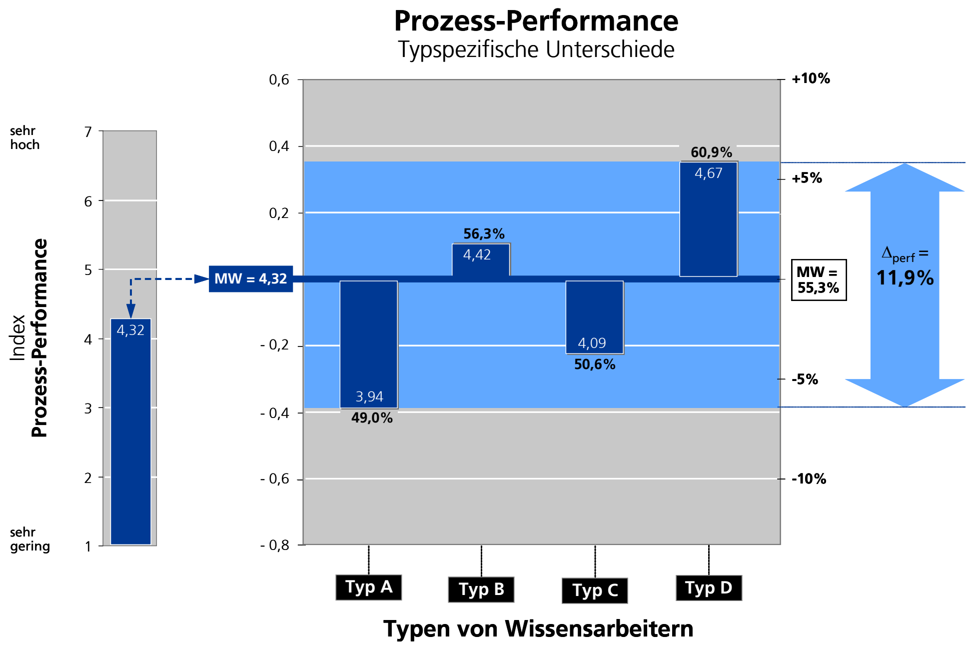


ABB. 20 *Zusammensetzung und Mittelwerts-Übersicht des Index Prozess-Performance*

Bei einer vertiefenden Betrachtung des Index Prozess-Performance in Abhängigkeit der vier identifizierten, unterschiedlichen Wissensarbeits-Typen (Abb. 21) zeigt sich, dass eine deutliche Spreizung zwischen dem höchsten Wert bei Typ D und dem niedrigsten Wert bei Typ A vorliegt (Δ=11,9%). In diesem Vergleich und in Relation zum Gesamt-Mittelwert zeichnet sich neben Typ D (+5,5%) auch der Typ B durch eine leicht überdurchschnittliche Prozess-Performance aus (+1,0%).

ABB. 21 *Index Prozess-Performance der vier Wissensarbeits-Typen im Vergleich*



Mögliche Ursachen dieser unterschiedlichen Performance-Werte lassen sich mit Hilfe der nachfolgenden Analysen weiter eingrenzen.

3.4.2 IuK-Qualität

Der Index »IuK-Qualität« umfasst die in Abb. 22 dargestellten fünf Bewertungen in Bezug auf die Qualität und Leistungsfähigkeit der informations- und kommunikationstechnischen Ausstattung:

- Möglichkeiten des schnellen und unkomplizierten Zugriffs auf benötigte Nachrichten, Daten, Unterlagen,
- Zuverlässigkeit und Stabilität der IuK-Technik,
- Vermeidung von Medienbrüchen,
- Erfüllung persönlicher Bedürfnisse durch die zur Verfügung stehende Hard-/ Software,
- Ausstattungsniveau des Arbeitsplatzes mit geeigneter IuK-Technik insgesamt.

Insgesamt betrachtet, ergibt sich für den Index IuK-Qualität ein Mittelwert von MWiIuK = 4,84 (Standardabweichung 0.99). Dies entspricht einem durchaus noch ausbaufähigen Qualitätsniveau von insgesamt 64,0%. Auffällig ist, dass insbesondere die an vielen Stellen zu verzeichnenden Medienbrüche und die fehlende Durchgängigkeit in der Verarbeitung und Nutzung vorhandener Daten (MW=4,11) sich in diesem Zusammenhang als Qualitätsmangel erweisen.

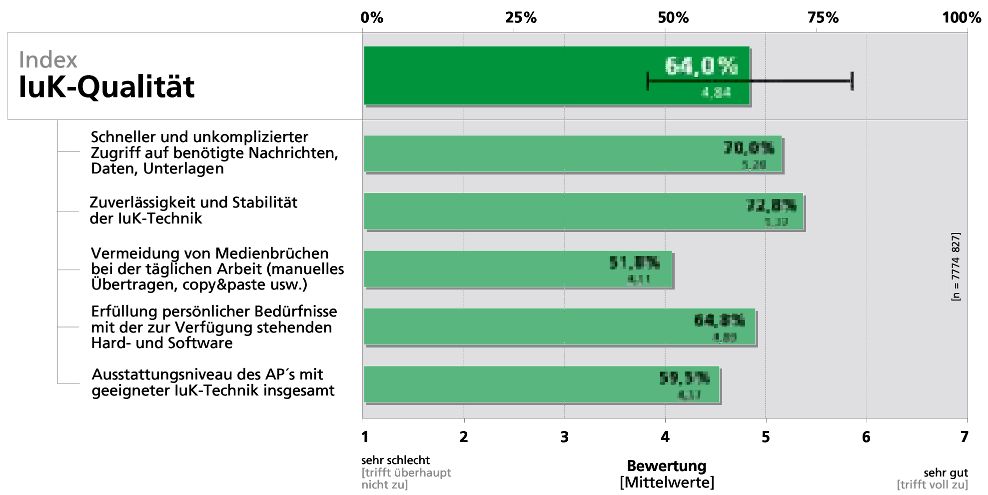
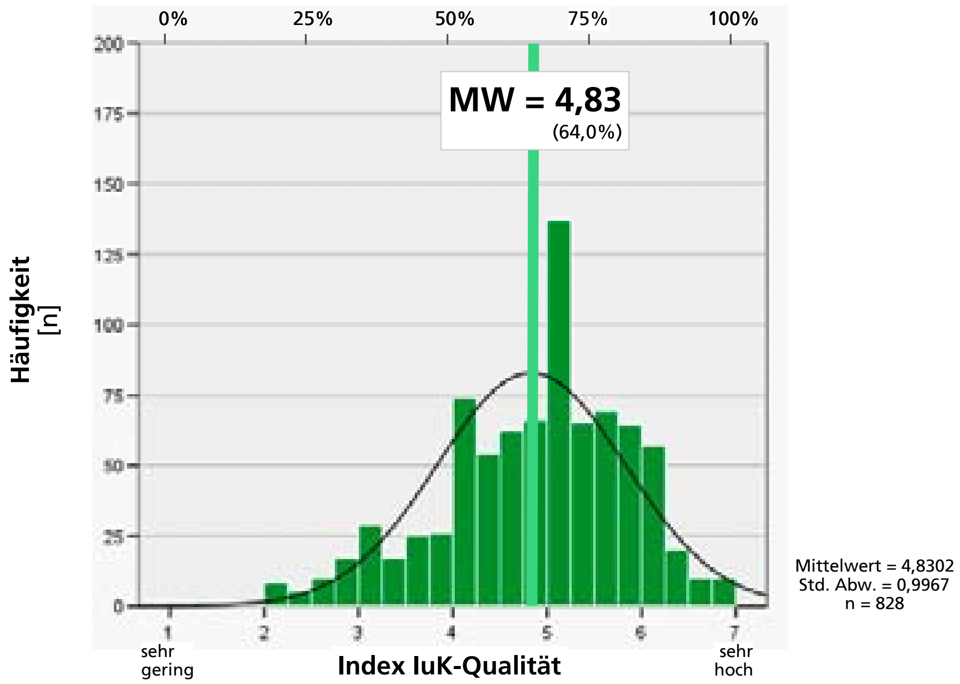


ABB. 22 *Zusammensetzung und Mittelwerts-Übersicht des Index IuK-Qualität*

Die Verteilung über einzelne Werteklassen des Index IuK-Qualität ist in Abb. 23 dargestellt.

ABB. 23 *Gesamtverteilung des Index IuK-Qualität über einzelne Werteklassen und Mittelwert insgesamt*



Auch hierzu zeigt die vertiefende Analyse dieses Index im Vergleich der vier unterschiedlichen Wissensarbeits-Typen einen in Bezug auf den Mittelwert auffälligen Sprungverlauf mit einer Gesamtspreizung von Δ_{IuK} = 12,7% (Abb. 24). Vertreter des Typ D können im Allgemeinen über die beste IuK-Qualität verfügen (69,7%), während sich das Cluster bei Typ A im Normalfall mit dem niedrigsten IuK-Qualitätsniveau begnügen muss (57,0%).

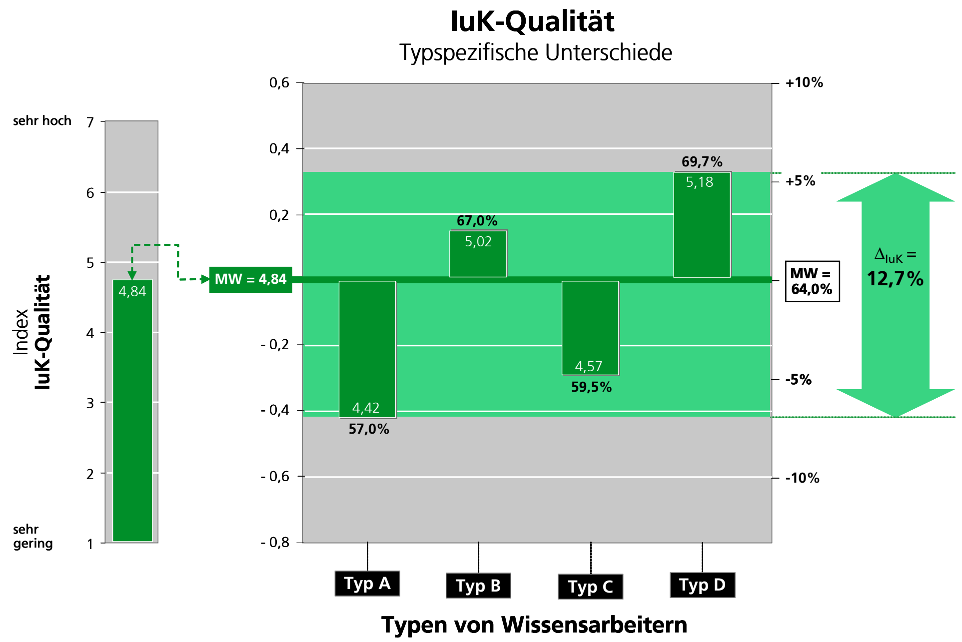


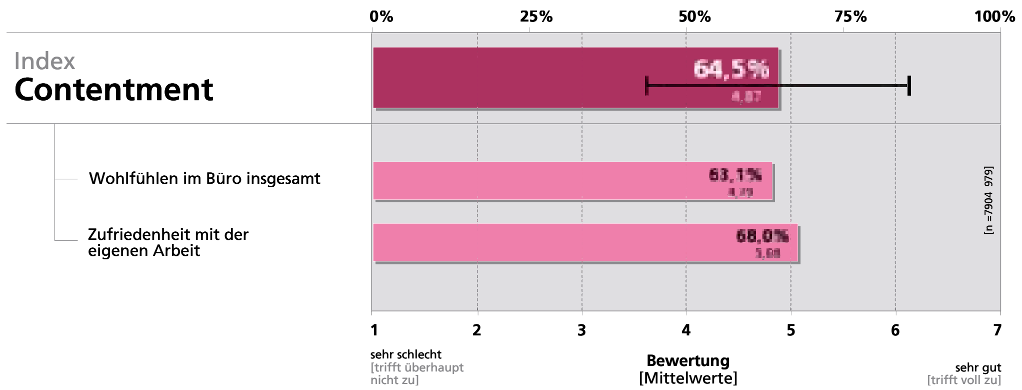
ABB. 24 *Index IuK-Qualität der vier Wissensarbeits-Typen im Vergleich*

3.4.3 Contentment

Der Grad an Zufriedenheit mit der eigenen Arbeit einerseits und dem Wohlbefinden im Büro andererseits wird mit Hilfe des Index »Contentment« beschrieben.

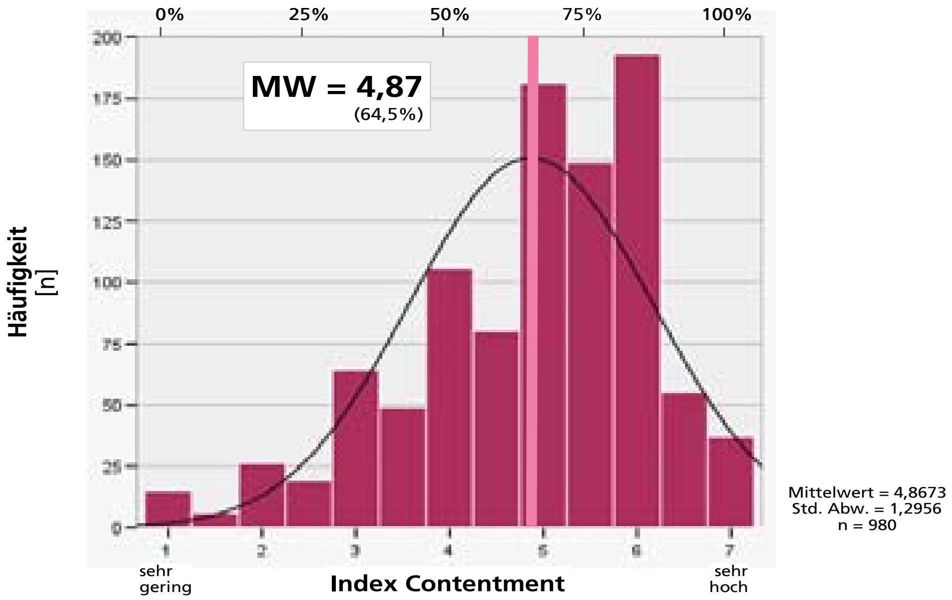
Über alle Teilnehmer hinweg betrachtet, liegt der Gesamtwert des Index Contentment bei einem Mittelwert von MW_{iCont} = 4,87 (Standardabweichung 1.29). Übertragen auf die entsprechende Prozentskala entspricht dies einem Contentment-Gesamtniveau von 64,5% (Abb. 25). Auffällig hierbei ist, dass die Wohlfühl-Qualität im Büro insgesamt (63,1%) den Zufriedenheitswerten mit der eigenen Arbeit (68,0%) hinterher hinkt. Dies kann als Signal auch an Planer und Innenarchitekten verstanden werden, das Arbeits- und Büroambiente insgesamt attraktiver, wertiger und anregender zu gestalten.

ABB. 25 *Zusammensetzung und Mittelwerts-Übersicht des Contentment-Index*



Die Verteilung über einzelne Werteklassen dieses Index ist in Abb. 26 dargestellt.

ABB. 26 *Gesamtverteilung des Index Contentment über einzelne Werteklassen*



Bei einer differenzierten Analyse des Contentment-Index im Vergleich der unterschiedlichen Wissensarbeits-Typen zueinander, zeigt sich hierbei eine noch größere Spreizung als bei den beiden vorgenannten Indizes ($\Delta_{Cont.}$=20,4%). Während sowohl Typ A (54,6%) als auch Typ C (62,0%) unterduchschnittliche Werte aufweisen, kann hingegen insbesondere bei Vertretern des Typ D (75,0%) aber auch bei Typ B (68,5%) ein weitaus größeres Maß an Zufriedenheit und Wohlbefinden festgestellt werden (Abb. 27).

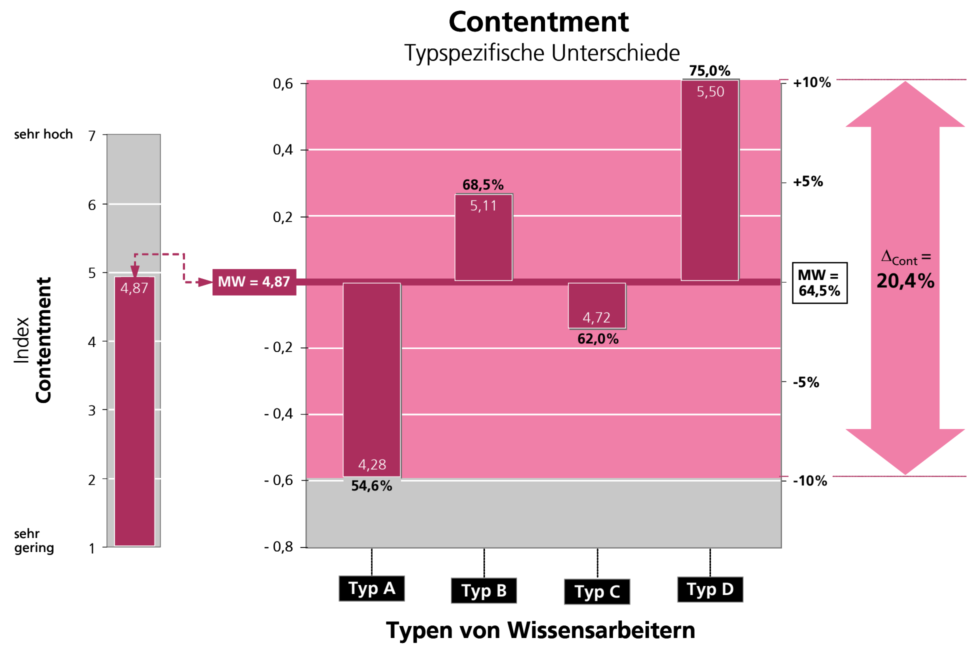


ABB. 27 *Contentment-Index der vier Wissensarbeits-Typen im Vergleich*

Die Vermutung und Hypothese, dass die überdurchschnittlich hohen Contentment-Werte bei Typ D und Typ B bzw. die deutlich geringeren Werte bei Typ A und Typ C mit dem in Abb. 17 ausgewiesenen Ranking in Bezug auf die Autonomie-Werte der jeweiligen Wissensarbeits-Typen in engerem Zusammenhang stehen, wird nachfolgend diskutiert.

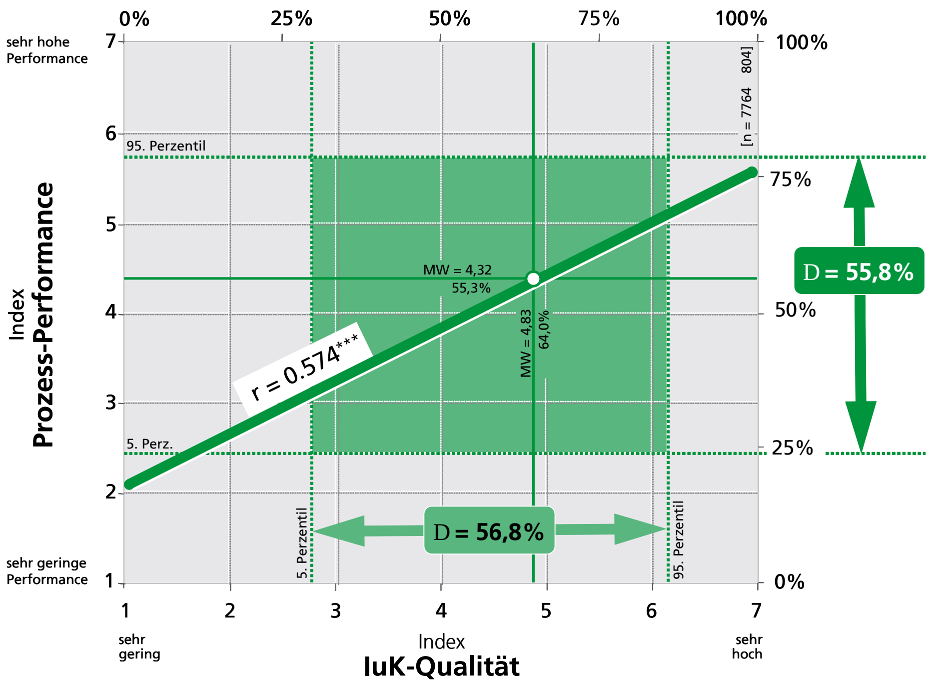
3.4.4 Performance-orientiertes Wirkungsmodell für Wissensarbeit

Auf Basis unterschiedlicher Korrelationsanalysen, mittels derer die Signifikanz möglicher Zusammenhänge zwischen den zentralen, o. g. Indizes untersucht und geprüft wurde, läßt sich ein grundlegendes performance-orientiertes Wirkungsmodell für Wissensarbeit beschreiben.

Zunächst kann ein hochsignifikanter und stark positiver Zusammenhang (Korrelation r = 0.574***, p < 1%) zwischen den beiden Indizes »Prozess-Performance« und »IuK-Qualität« aufgezeigt werden (Abb. 28).

Dies bedeutet: eine geringe IuK-Qualität korreliert in sehr hohem Maße mit einer geringen Prozess-Performance, oder je höherwertig die Qualität der zur Verfügung stehenden IuK-Ausstattung ist, desto höher ist im allgemeinen auch die Prozess-Performance der jeweiligen Wissensarbeiter.

ABB. 28 *Korrelations-Portfolio der Indizes Prozess-Performance vs. IuK-Qualität mit den jeweiligen Grenzwerten für das 5./95. Perzentil*



Betrachtet man neben den Mittelwerten auch die weitere Verteilung über die beiden Indizes, so lassen sich massive Unterschiede feststellen. Unter Vernachlässigung von Extremwerten am unteren und oberen Ende (5. - 95. Perzentil) weisen die Abweichungen in der Breite über alle Teilnehmer bezogen auf deren IuK-Qualität mit Δ_{IuK} = 56,8% bzw. deren Prozess-Performance mit $\Delta_{Perf.}$ = 55,8% eine insgesamt recht große Streuung auf (Abb. 28).

Dies ist ein deutliches Indiz dafür, dass die IuK-Qualität aus Sicht der Befragten in vielen Fällen offensichtlich nicht (mehr) zeitgemäß ist und einen hohen Verbesserungsbedarf aufweist. Ein überaus lohnender Ansatz also, wenn es darum geht, Qualitätsstandards in Bezug auf die technische Ausstattung von Informations- und Wissensarbeitern zu überprüfen und Investitionen zu planen.

Vergleicht man in diesem Kontext die vier unterschiedlichen Wissensarbeits-Typen miteinander, so lassen sich diese gut entlang der Regressionsgeraden einordnen (Abb. 29). In Bezug auf deren typspezifische IuK-Qualität beträgt die maximale Differenz zwischen Typ A (57,0%) und Typ D (69,7%)

$\Delta_{abs.}$ = 12,7%.

Die Prozess-Performance der vier Typen unterscheidet sich auf der Absolutskala zwischen Typ A (49,0%) und Typ D (60,9%) um bis zu

$\Delta_{abs.}$ = 11,9%.

Dies sind absolut betrachtet gewaltige Unterschiede bzw. Potenziale, die im Sinne einer betriebswirtschaftlichen Investitions- und Renditebetrachtung ein Ergebnis im 2-stelligen Vor-Komma-Bereich erwarten lassen.

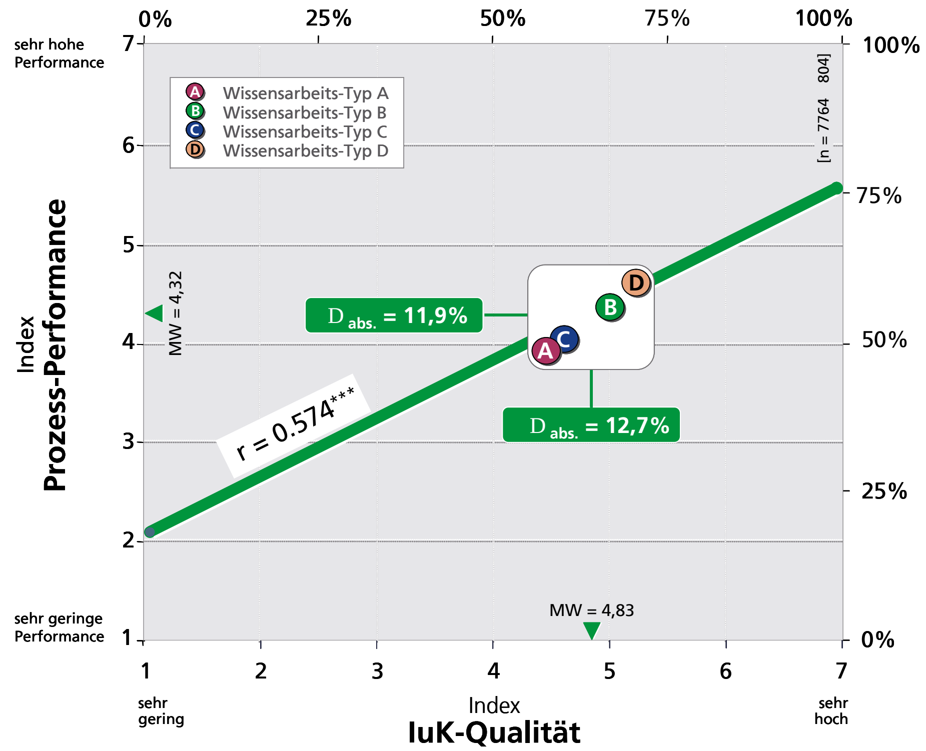
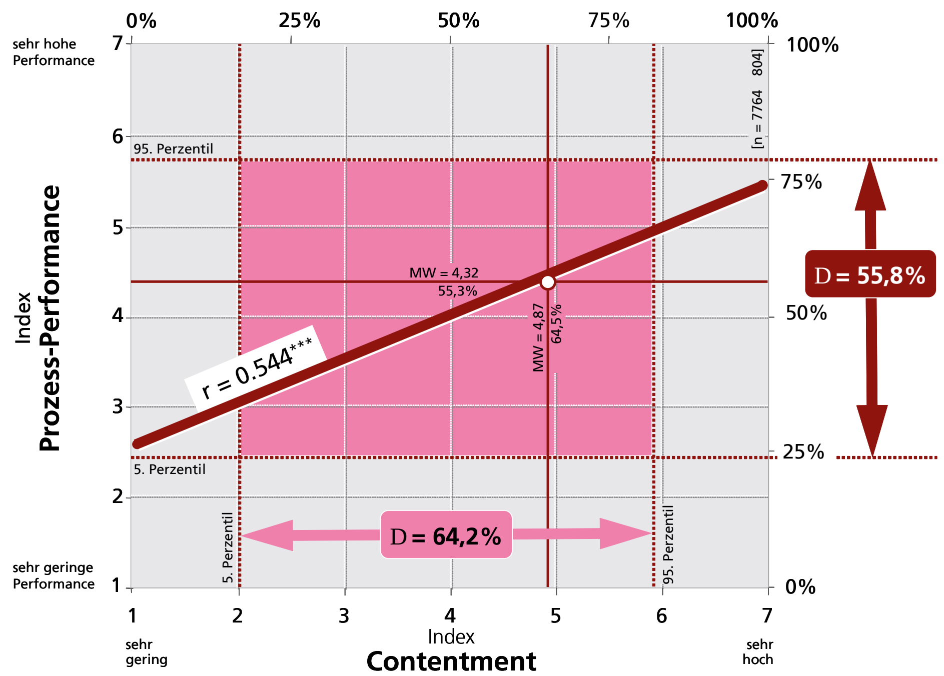


ABB. 29 *Einordnung der Wissensarbeits-Typen in das Portfolio Prozess-Performance vs. IuK-Qualität*

Ein ebenfalls hochsignifikanter und stark positiver Zusammenhang ist zwischen den beiden Indizes »Prozess-Performance« und »Contentment« feststellbar (Korrelation r = 0.544***, p < 1%). Wie aus Abb. 30 ersichtlich ist, korreliert ein geringer Contentment-Wert in sehr hohem Maße auch mit einer geringen Prozess-Performance und umgekehrt.

ABB. 30 *Korrelations-Portfolio der Indizes Prozess-Performance vs. Contentment mit den jeweiligen Grenzwerten für das 5./95. Perzentil*



Betrachtet man neben den Mittelwerten auch die weitere Verteilung über die beiden Indizes, so lassen sich ebenfalls massive Unterschiede feststellen. Unter Vernachlässigung von Extremwerten am unteren und oberen Ende (5. - 95. Perzentil) liegen die Abweichungen in der Breite über alle Teilnehmer bezogen auf deren Contentment-Werte bereits bei $\Delta_{cont.}$ = 64,2% bzw. deren Prozess-Performance bei $\Delta_{Perf.}$ = 55,8% (Abb. 30).

Dieses Ergebnis weist bezüglich Zufriedenheit und Wohlbefinden bei vielen Büro- und Wissensarbeitern offensichtlich auf erhebliche Unterschiede bzw. Defizite hin.

Bei der Ursachenforschung sollten neben unangemessenen Arbeitsinhalten insbesondere ein Übermaß nicht beherrschbarer Stressfaktoren, Spannungen im kollegialen Umfeld und mit Vorgesetzten, belastende Faktoren der Arbeitsumgebung oder ein nachlässig gestaltetes, unattraktives Büroambiente einer kritischen Überprüfung unterzogen werden. Im Umkehrschluß sind diesbezüglich positiv erlebte Ansätze und überzeugende Gestaltungskonzepte in hohem Maße für Zufriedenheit und Wohlbefinden der Wissensarbeiter verantwortlich – und damit auch entscheidend für deren hohe Prozess-Performance.

Vergleicht man in diesem Zusammenhang die vier unterschiedlichen Wissensarbeits-Typen miteinander, so reihen sich diese wiederum in der bekannten Abfolge und Spreizung entlang der Regressionsgeraden auf (Abb. 31). Die typspezifischen Contentment-Werte unterscheiden sich zwischen Typ A (57,0%) und Typ D (69,7%) um bis zu

$\Delta_{abs.}$ = 20,4%.

Die Prozess-Performance der vier Typen unterscheidet sich dabei zwischen Typ A (49,0%) und Typ D (60,9%) wie bereits oben ausgeführt, um bis zu

$\Delta_{abs.}$ = 11,9%.

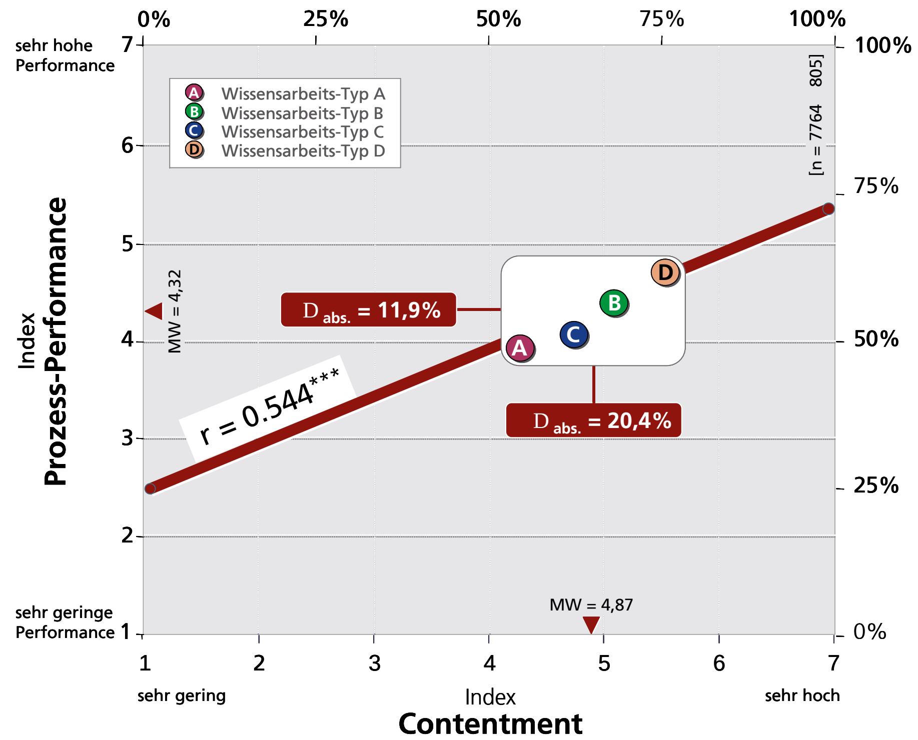


ABB. 31 *Einordnung der Wissensarbeits-Typen in das Portfolio Prozess-Performance vs. Contentment*

Die Analyse zeigt, das sowohl Contentment als auch Prozess-Performance stark mit dem Typus variieren. Autonom arbeitende Personen mit anspruchsvollen und abwechslungsreichen Aufgabenstellungen sind deutlich »performanter« (vgl. Typ D und Typ B). Ebenso weisen sie deutlich höhere Werte in Bezug auf Arbeitszufriedenheit und Wohlbefinden auf als beispielsweise die weniger autonom agierenden Typen A und C. Dies legt den berechtigten Schluss nahe, dass bei Wissensarbeitern die Erhöhung der (Arbeits-)Autonomie, also das »einräumen und nutzen können« sowohl inhaltlicher und zeitlicher aber auch örtlich/räumlicher Freiheitsgrade und Wahlmöglichkeiten nicht nur die Zufriedenheit und das Wohlbefinden fördert, sondern auch die Performanz steigert.

Die abschließende Korrelationsanalyse über die beiden Indizes IuK-Qualität und Contentment zeigt ebenfalls einen hochsignifikanten und stark positiven Zusammenhang auf (Korrelation r =0.484***, p < 1%).

Auch hier ist eine Verbesserung des einen Kennwertes jeweils auch mit der Verbesserung des anderen Kennwertes gekoppelt bzw. umgekehrt (Abb. 32).

Die Streuung innerhalb der Grenzwerte für das 5. bzw. 95. Perzentil bewegen sich in den bekannten Bereichen ($\Delta_{Cont.}$ = 64,2% bzw. Δ_{IuK} = 56,8%).

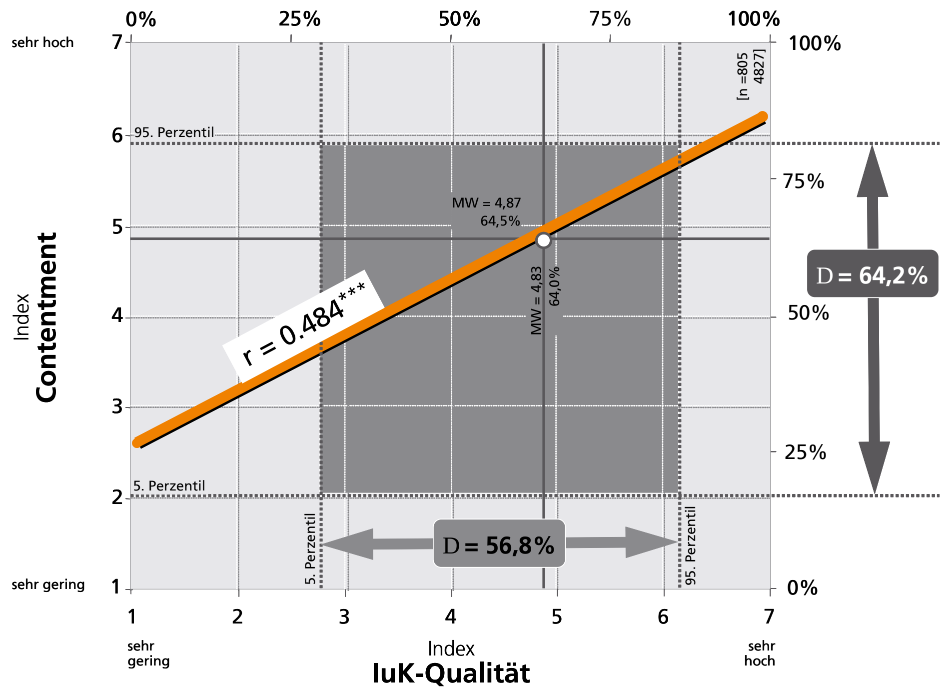


ABB. 32 *Korrelations-Portfolio der Indizes IuK-Qualität vs. Contentment mit den jeweiligen Grenzwerten für das 5./95. Perzentil*

Die vier Wissensarbeits-Typen lassen sich in dem Portfolio Contentment vs. IuK-Qualität wie in Abb. 34 dargestellt, platzieren. Die größere vertikale Streuung um die Regressionsgerade weist neben dem erkennbaren Einfluß der IuK-Qualität auch auf bereits oben diskutierte, überlagernde Einflußgrößen (z. B. Autonomiegrad) bezüglich des Contentment-Wertes hin.

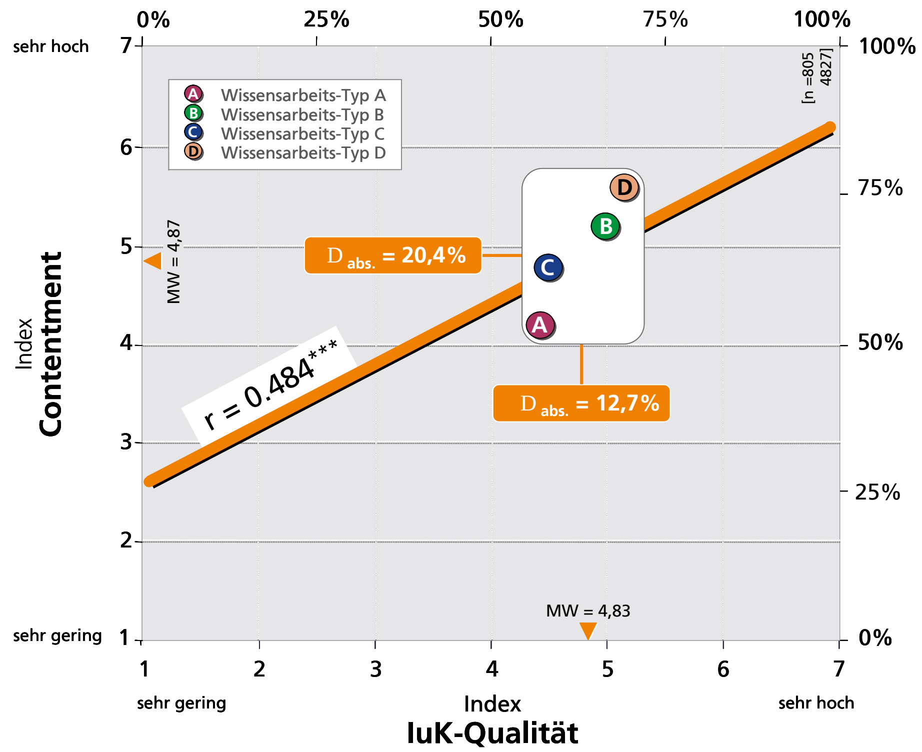


ABB. 33 *Einordnung der Wissensarbeits-Typen in das Portfolio Contentment vs. IuK-Qualität*

Da ein signifikanter Zusammenhang zwischen den beiden Indizes Contentment und IuK-Qualität nachweisbar ist und beide wiederum auch in direktem Zusammenhang mit der Prozess-Performance von Wissensarbeitern stehen, lassen sich die gezeigten Abhängigkeiten in Form des in Abb. 34 dargestellten performance-orientierten Wirkungsmodells für Wissensarbeit zusammenfassen.

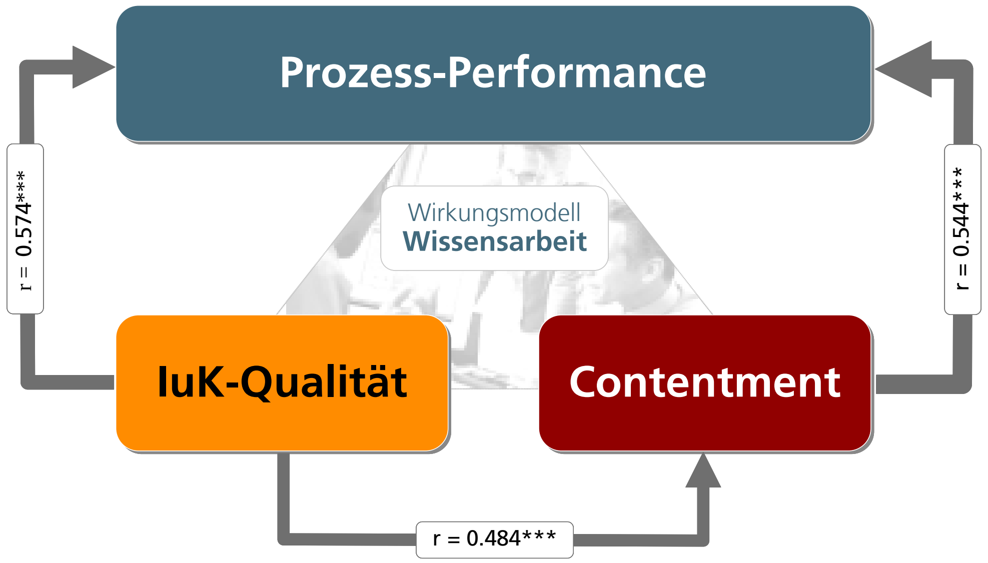


ABB. 34 *Korrelationen eines performance-orientierten Wirkungsmodells für Wissensarbeit auf Basis der drei LeitIndizes*

Damit konnten auf Basis der Studie mächtige Stellhebel identifiziert und quantifizierbar gemacht werden.

Oder anders formuliert: Um die Prozess-Performance bei Wissensarbeit positiv zu beeinflussen bzw. zu optimieren, gilt es gleichermaßen – und jeweils auch über alle genannten Teilaspekte hinweg (vgl. Kap. 3.4.2, Kap. 3.4.3) – sowohl die IuK-Qualität als auch Arbeitszufriedenheit und Wohlbefinden der Wissensarbeiter kontinuierlich zu hinterfragen, zu fördern und hoch zu halten.

3.4.5 Technologieprofile

Neben der allgemeinen qualitativen Einordnung und Bewertung der von den unterschiedlichen Wissensarbeitern genutzten informations- und kommunikationstechnischen Ausstattung wird nachfolgend insbesondere die Frage untersucht, inwieweit es bezüglich der zur Verfügung stehenden IT-/Technologiekomponenten Unterschiede zwischen den einzelnen Wissensarbeits-Typen gibt und wie diese sich im Anwendungs- und Nutzungsverhalten bei der täglichen Arbeit unterscheiden.

Dazu sind für die vier unterschiedlichen Wissensarbeits-Typen typische Technologieprofile sowohl in Bezug auf den Anteil der tatsächlicher Nutzer als auch der jeweiligen Nutzungshäufigkeit einzelner IT-/Technologiekomponenten ausgewertet und in Form von Portfolios dargestellt.

Für ca. 90% aller Vertreter des Typ A sind ein stationärer PC und ein tischgebundenes Telefon nicht nur die gängigsten, sondern auch die am häufigsten genutzten technischen Arbeitsmittel (Abb. 35). Ergänzend dazu kommen bei ca. 55-40% der A-Typen insbesondere noch Notebook, Mobiltelefon und WLAN zum Einsatz. Da für den Typ A im allgemeinen ein niedriger Autonomiegrad bei starker örtlicher bzw. räumlicher Bindung an den eigenen Arbeitsplatz kennzeichnend ist, lässt sich dieses Technikprofil als »konservative Grundausstattung« beschreiben.

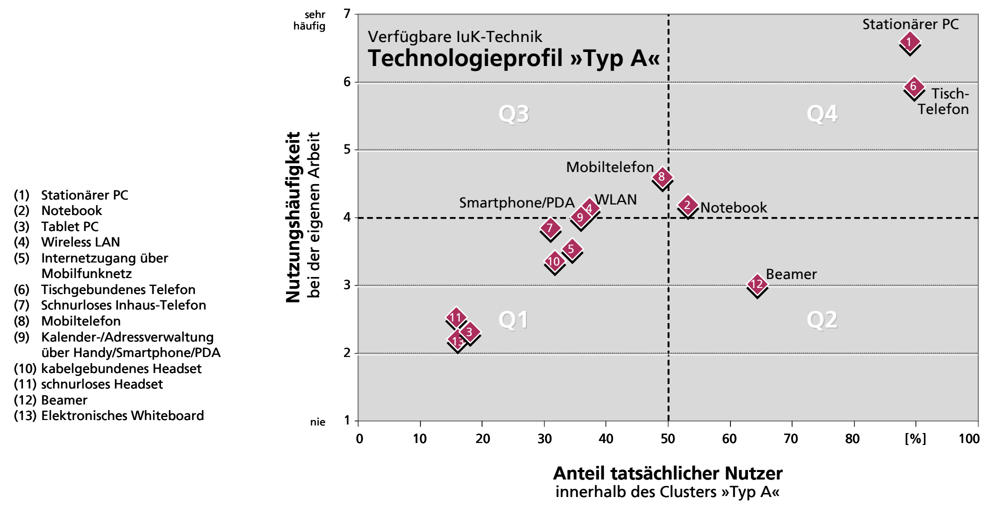
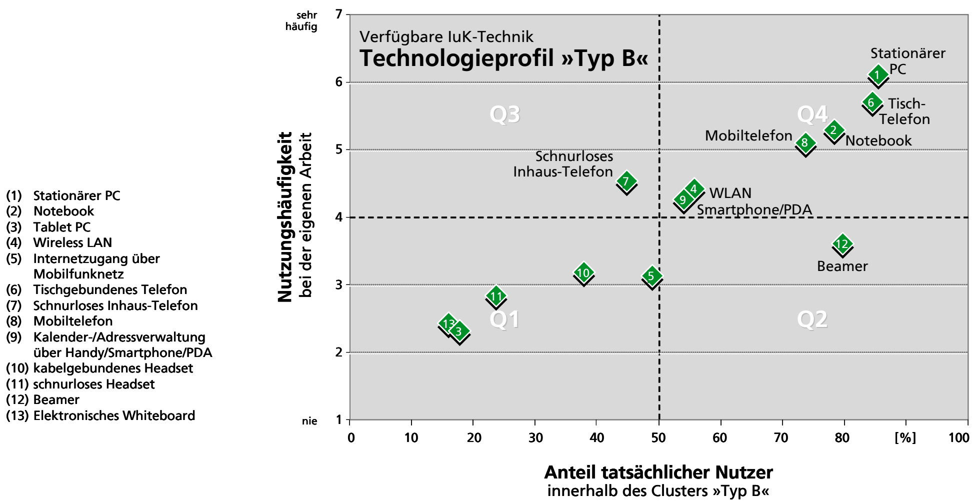


ABB. 35 *Verfügbare IuK-Technik und deren Nutzungshäufigkeit bei Typ A (Technologieprofil Typ A)*

Wissensarbeitern des Typ B steht normalerweise ein Technikpaket zur Verfügung, das bei ca. 90% dieses Clusters neben örtlich stationären Komponenten (PC, Tischtelefon) für weitere knapp 80% auch mobile Komponenten (Notebook, Mobiltelefon) enthält (Abb. 36). Zur weiteren Unterstützung mobiler und flexibler Arbeitsweisen sind zumindest für ca. 40-60% der Befragten auch WLAN, Smartphones und schnurlose Inhouse-Telefone verfügbar.

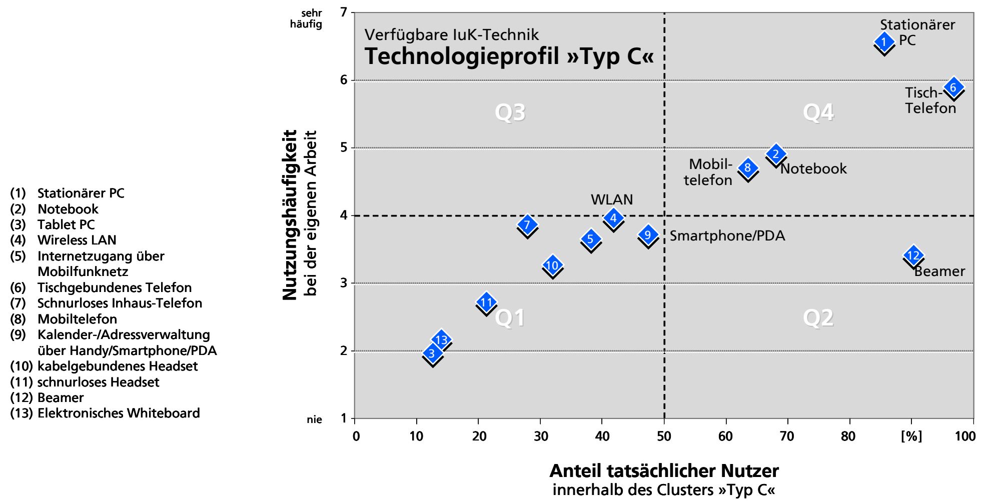
Insgesamt betrachtet befinden sich bei Vertretern des Typ B die Mehrzahl der aufgeführten technischen Arbeitsmittel in der rechten Hälfte des Portfolios (Q2/Q4), d. h . der Anteil tatsächlicher Nutzer dieser Ausstattung ist recht hoch (Nutzeranteil zwischen 50-85%).

ABB. 36 *Verfügbare IuK-Technik und deren Nutzungshäufigkeit bei Typ B (Technologieprofil Typ B)*



Für Wissensarbeiter des Typ C ist ein typisches und intensiv genutztes IuK-Technikpaket weniger umfangreich und besteht im wesentlichen aus Tischtelefon (94%) und örtlich stationärem PC (83%). Für ca. 60-70% der Befragten auch aus Notebook und Mobiltelefon, die jedoch merklich weniger intensiv genutzt werden. WLAN ist bei einer mittleren Nutzungshäufigkeit im Durchschnitt für ca. 40% dieses Typs verfügbar (Abb. 37). Die meisten der aufgeführten technischen Arbeitsmittel liegen jedoch in dem unteren linken Quadranten des Portfolios (Q1), d. h. diese sind nur für relativ wenige Vertreter dieses Typus verfügbar und werden auch relativ selten bei der eigenen Arbeit genutzt.

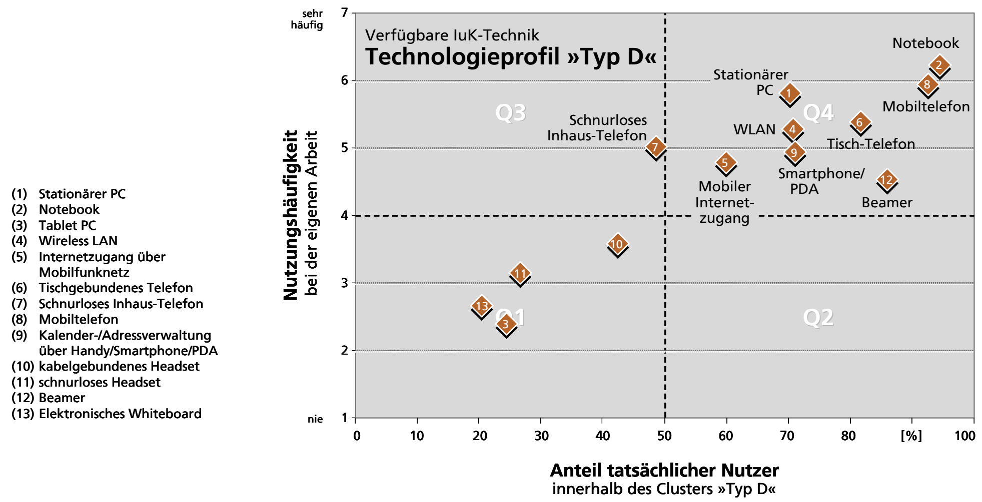
ABB. 37 *Verfügbare IuK-Technik und deren Nutzungshäufigkeit bei Typ C (Technologieprofil Typ C)*



Bei Typ D ist auffällig, dass im Vergleich zu den anderen bislang diskutierten Wissensarbeits-Typen grundsätzlich sowohl in der Breite als auch in der täglichen Anwendung und Nutzung deutlich mehr IuK-Komponenten zur Verfügung stehen und diese Arbeitsmittel dann auch im Einsatz bei der eigenen Arbeit genutzt werden.

Das sehr umfangreiche Technologieprofil bei Typ D mit 9 »Hauptkomponenten« im rechten oberen Quadranten Q4 ist in Abb. 38 dargestellt. Notebook und Mobiltelefon sind nahezu bei allen Vertretern des Typ D das zentrale Arbeitsmittel (Nutzeranteil ca. 95%). Auf einen stationären PC greifen im Vergleich zu den anderen Typen hier deutlich weniger Nutzer zurück (ca. 70%). Ebenso gehören Beamer, Tischtelefone, PC, WLAN, Smartphone, mobiler Internetzugang und schnurlose Inhaus-Telefone zu den weitverbreiteten und intensiv genutzen Standards bei diesem Typus (Nutzeranteil ca. 85-50%).

ABB. 38 *Verfügbare IuK-Technik und deren Nutzungshäufigkeit bei Typ D (Technologieprofil Typ D)*



Die gerade für den Wissensarbeits-Typ D sehr stark ausgeprägte Arbeitsautonomie wird demzufolge auch konsequent mit einer sehr umfangreichen technologischen Grundausstattung zur Unterstützung mobiler und flexibler Arbeitsweisen gekoppelt. Offensichtlich ist, dass Vertreter dieses Typus gewohnt und bestrebt sind, auf ein sehr umfangreiches und »modernes« Technologiepaket zurückgreifen zu können und dieses auch bei der täglichen Arbeit einsetzen.

Insgesamt betrachtet – und nahezu deckungsgleich über alle Wissensarbeits-Typen – zeigt sich, dass elektronische Whiteboards, Tablet-PC´s und schnurlose Headsets bislang erst relativ selten eingesetzt werden bzw. diese Komponenten grundsätzlich noch größere Entwicklungspotenziale aufweisen.

3.4.6 Typische IT-Anwendungen und deren (Markt-)Potenzial

Die subjektive Nutzeneinschätzung zu einzelnen IT-Anwendungen in Kombination mit dem aktuellen Anteil an Wissensarbeitern, die über bestimmte IT-Anwendungen verfügen und diese auch tatsächlich nutzen, sind wichtige Indikatoren nicht nur für die Marktdurchdringung, sondern auch für das (Markt-)Potenzial dieser IT-Anwendungen in den unterschiedlichen Zielgruppen. Demzufolge kann man bei denjenigen IT-Anwendungen von einem hohen (Markt-)Potenzial ausgehen, denen seitens der Befragungsteilnehmer zum einen ein relativ hoher Nutzen attestiert wird, die sich aber zugleich durch eine bislang noch eher geringe Marktdurchdringung auszeichnen, d. h. aktuell erst von wenigen Teilnehmern genutzt werden. Dieses Szenario wird durch das linke, obere »SOLL-Feld« des in Abb. 39 dargestellten Potenzial-Portfolios beschrieben.

Dem »MUSS-Feld« – in diesem Portfolio rechts oben angesiedelt – lassen sich dagegen Technologien mit einem relativ hohen Nutzeranteil bei gleichzeitig hoher Nutzeneinschätzung zuordnen. Hier finden sich sinnvolle und bewährte Standardtechnologien mit hoher Marktdurchdringung wieder.

Ein geringer Anteil tatsächlicher Nutzer und geringe Nutzeneinschätzung – im dritten »TEST-Feld« links unten angesiedelt – lassen dagegen auf Technologien mit (noch) geringer Marktreife schliessen, die z. B. noch weiterer Verbesserungen oder Tests bedürfen, deren Business-Modelle zu überprüfen oder deren Marketingstrategien zu optimieren sind.

Am problematischsten ist sicher das »PRÜFEN-Feld«, rechts unten im Portfolio. Bei Technologien mit hohem Nutzeranteil aber geringer Nutzeneinschätzung ist zu hinterfragen, ob deren Anwendungen im jeweiligen Fall sinnvoll und notwendig sind.

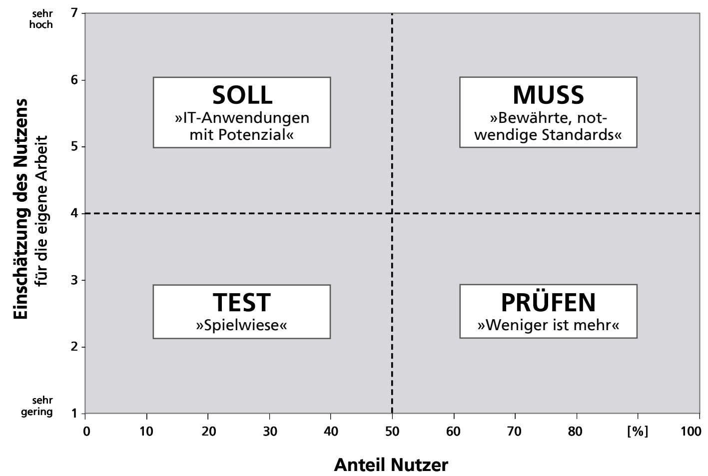


ABB. 39 *Segmente des Potenzial-Portfolios typischer IT-Anwendungen [Prinzipdarstellung]*

Betrachtet man die bevorzugten Technologieanwendungen insgesamt über alle Teilnehmer hinweg, dann gilt zunächst für alle vier Wissensarbeits-Typen generell, dass Teamserver und Teamkalender deutlich dem »MUSS-Feld« der bewährten Standardtechnologien zuzuordnen sind (Nutzeranteil ca. 65-90%).

Eine weiteres Cluster bilden Groupware, Dokumenten-Managementsysteme und der Datenzugang über Web-Portale. Auch diese Technologien werden insgesamt betrachtet weitestgehend positiv beurteilt und verfügen über einen mittleren Nutzeranteil (ca. 40-60%).

Bei den beiden relativ »immobilen« Wissensarbeits-Typen A und C fällt auf, dass sich neben den zwei IT-Anwendungen (Teamserver, Teamkalender) keine weiteren Technologien in dem rechten oberen »MUSS-Feld« ihres Portfolios befinden (Abb. 40; Abb. 42). Sehr viele der abgefragten Technologien befinden sich im »TEST-Feld« mit geringen Nutzeranteilen (ca. 15-30%) und verhaltenem Nutzenurteil.

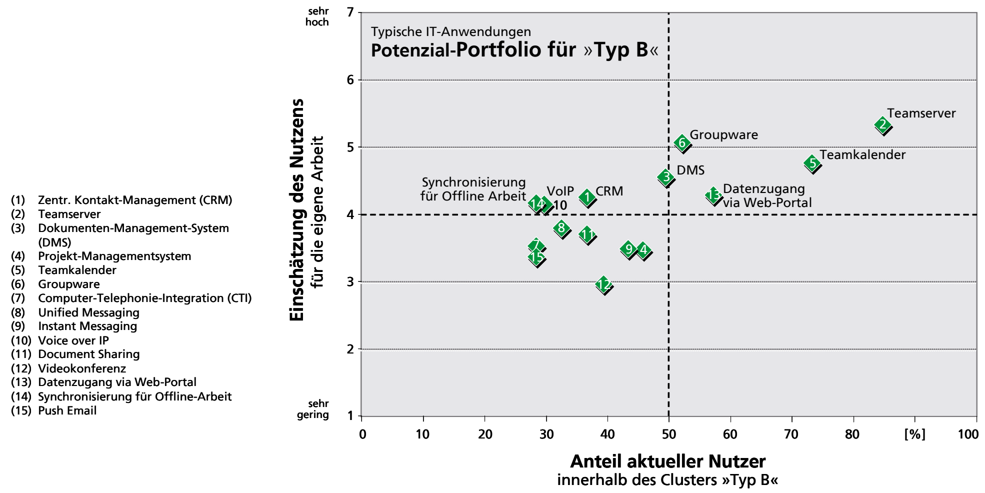
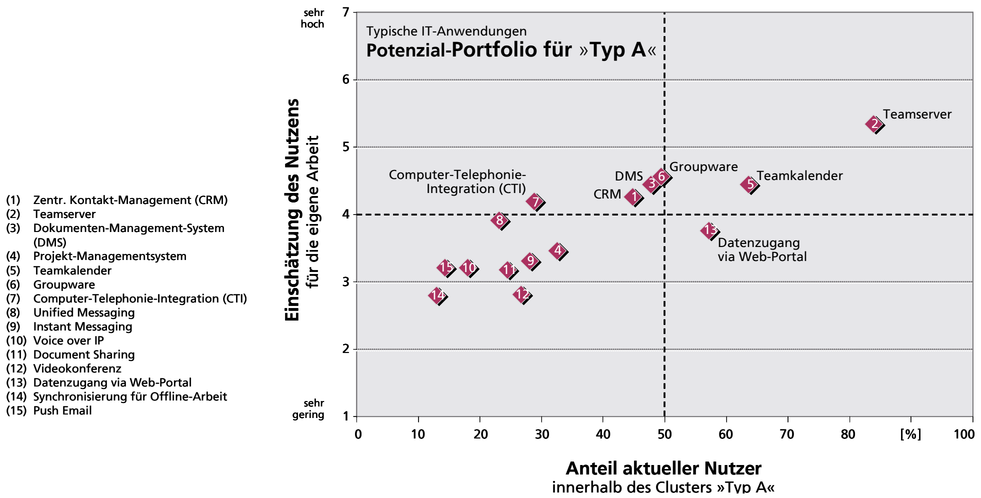


ABB. 41 *Potenzial-Portfolio typischer IT-Anwendungen für Typ B*

ABB. 40 *Potenzial-Portfolio typischer IT-Anwendungen für Typ A*



Der Wissensarbeits-Typ C ist bezüglich der derzeitigen und tatsächlichen Anwendung und Einschätzung der aufgeführten Technologien ähnlich aufgestellt wie Typ A. Die meisten Anwendungen finden sich im »TEST-Feld« des Portfolios (Abb. 42).

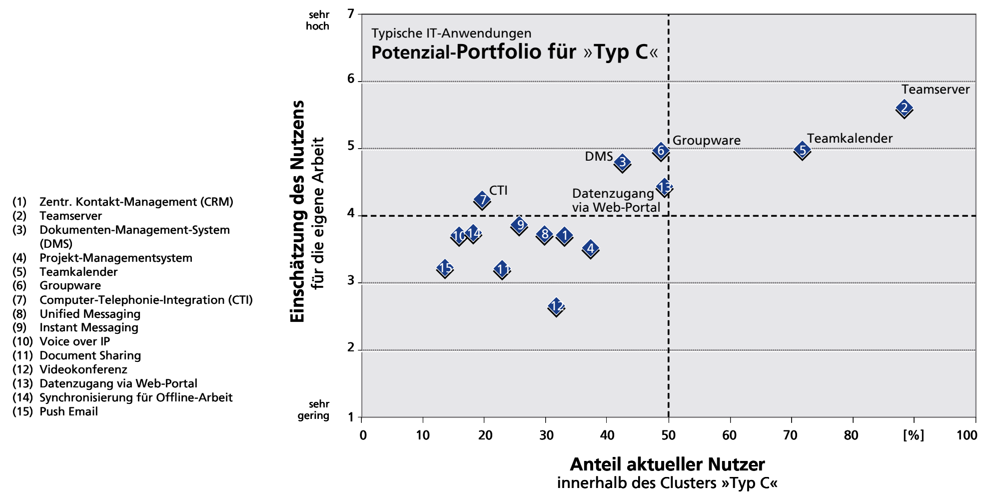


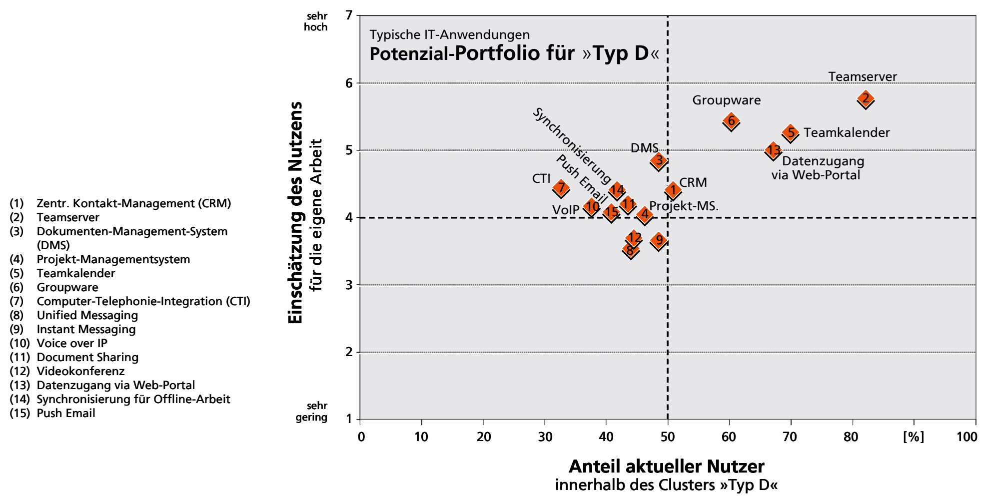
ABB. 42 *Potenzial-Portfolio typischer IT-Anwendungen für Typ C*

Für Vertreter des Typ B gehören neben Teamserver und Teamkalender speziell auch noch Groupware und der Datenzugang über Webportale zur technologischen Basisausstattung (Abb. 41). Außer Dokumenten- und Kontaktmanagement-Systemen (DMS und CRM), deuten hier noch am ehesten Voice-over-IP-Anwendungen und die Synchronisierung benötigter Dateien für Offline-Arbeit auf weitere Markterschließungs-Potenziale hin.

In auffälliger Abweichung zu den bisherigen Verteilungen bewerten Vertreter des Wissensarbeits-Typ D einen Großteil der aufgeführten Technologien hinsichtlich des tatsächlichen Nutzens für ihre eigene Arbeit insgesamt deutlich positiver.

Aus der Perspektive der D-Typen ist demnach der überwiegende Teil der abgefragten Technologien bei mobilen Arbeitsweisen und komplexen Arbeitsinhalten besonders nützlich und wird auch aktuell bereits von vielen eingesetzt. Nur drei der zahlreichen Technologien befinden sich knapp im »TEST-Feld« mit leicht unterdurchschnittlichem Nutzeranteil und Nutzenurteil.

ABB. 43 *Potenzial-Portfolio typischer IT-Anwendungen für Typ D*



Dieser Sachverhalt kann als weitere Bestätigung der oben bereits ausgeführten Erkenntnisse gewertet werden, dass der gezielte und umfassende Einsatz unterstützender IT-Technologien und Anwendungen nutzbringend für die Arbeit ist und die Performance von Wissensarbeitern entscheidend erhöht.

3.5 WEITERE ANALYSEN ZUM ARBEITSUMFELD VON WISSENSARBEITERN

Neben den teilweise recht massiven Unterschieden, die aufgrund der vorangegangen Analysen zwischen den einzelnen Wissensarbeits-Typen aufgezeigt werden konnten, soll im Rahmen der OFFICE 21®-Studie »Information Work 2009« auch das weitere Arbeitsumfeld von Wissensarbeitern nachfolgenden noch etwas genauer betrachtet werden.

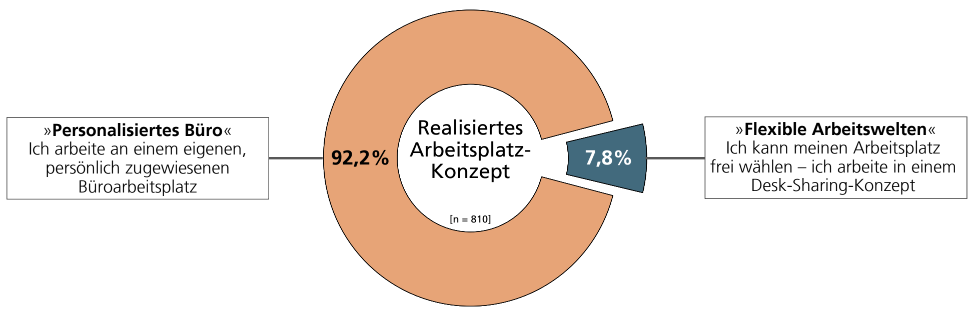
3.5.1 Arbeitsplatz-Konzept

Bei der Frage, welche Arbeitsplatz-Konzepte aktuell bei Wissensarbeitern realisiert sind, geben 92,2% der Teilnehmer an, über einen eigenen, persönlich zugeordneten Büroarbeitsplatz zu verfügen (Abb. 44).

Knapp 8% der Befragten arbeiten in »Flexiblen Arbeitswelten« bzw. auf Basis eines Desk-Sharing-Konzeptes, bei dem die Mitarbeiter in der Regel über keinen dauerhaft zugewiesenen Arbeitsplatz verfügen. Vielmehr werden bedarfsorientiert unterschiedliche bzw. wechselnde Arbeitsplätze genutzt.

Erwartungsgemäß sind – bedingt sowohl durch deren höhere örtliche und räumliche Mobilität als auch autonomere Arbeitsweisen – bei Vertretern des Typs D »Flexible Arbeitswelten« überproportional häufig realisiert (13,1%).

ABB. 44 *Realisierte Arbeitsplatz-Konzepte*



Untersucht man nun wiederum nur die Vertreter des Typ D hinsichtlich deren Arbeitsplatz-Konzept, so ist - wenn auch bei kleineren Gruppengrößen - auffällig, dass D-Vertreter aus »Flexiblen Arbeitswelten« im Vergleich zu D-Vertretern, die an einem fixen Platz in einem personalisierten Büro arbeiten, eine nochmals um 6% höhere Prozess-Performance aufweisen.

Eine strategische Überprüfung des geeigneten Arbeitsplatz-Konzeptes sollte demzufolge insbesondere bei Wissensarbeitern des Typ D auf jeden Fall vorgenommen werden. Auch wenn grundsätzlich jeder Typus dafür in Frage kommen kann, bei der Einführung von »Flexiblen Arbeitswelten« sind noch eine Reihe weiterer, spezifischer Punkte zu prüfen und zu bedenken.

3.5.2 Bildschirmfläche

Neuere Studien – u. a. auch im Rahmen einer vorangegangenen OFFICE 21®-Laborstudie – haben gezeigt, dass Mehrflächendisplays die Produktivität bei Bildschirmarbeit deutlich steigern können.

Im Rahmen der vorliegenden Studie wurden daher auch die aktuell zur Verfügung stehenden Monitorgößen untersucht und analysiert. Wie aus Abb. 45 ersichtlich, überwiegen derzeit 19-Zoll Monitore (33%) und 17-Zoll Monitore (29,1%). Kleinere Formate (14-/15-Zoll) werden von ca. 20% der Befragten verwendet. Auf Monitorgrößen ab 20 Zoll aufwärts entfällt ein Anteil von insgesamt 18%.

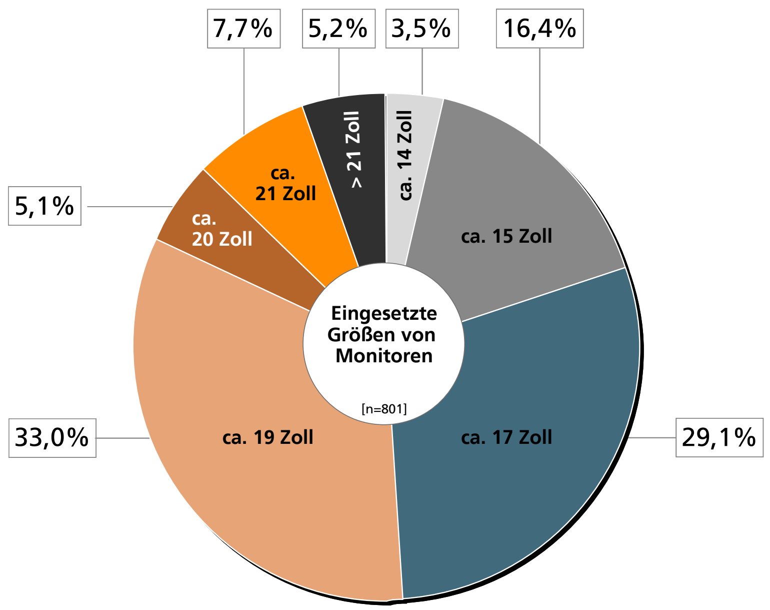


ABB. 45 *Eingesetzte Monitorgrößen (Häufigkeitsverteilung)*

Bei der Frage, wie zufrieden die Nutzer mit der jeweils eigenen Monitorgröße sind, ergibt sich ein klares und eindeutiges Meinungsbild. Je größer die zur Verfügung stehende Bildschirmfläche ausfällt, desto zufriedener sind die Nutzer (Abb. 46). Dieser Zusammenhang ist hochsignifikant (Rangkorrelation r = 0,390***).

Während die Gruppe der Nutzer von 14-Zoll Monitoren diese deutlich unterdurchschnittlich beurteilt ($MW_{14"}$ = 3,57), bewegen sich die Zufriedenheitswerte bei Monitoren mit einer Bildschirmdiagonalen von über 21-Zoll ($MW_{>21"}$ = 6,34) nahe am Maximalwert.

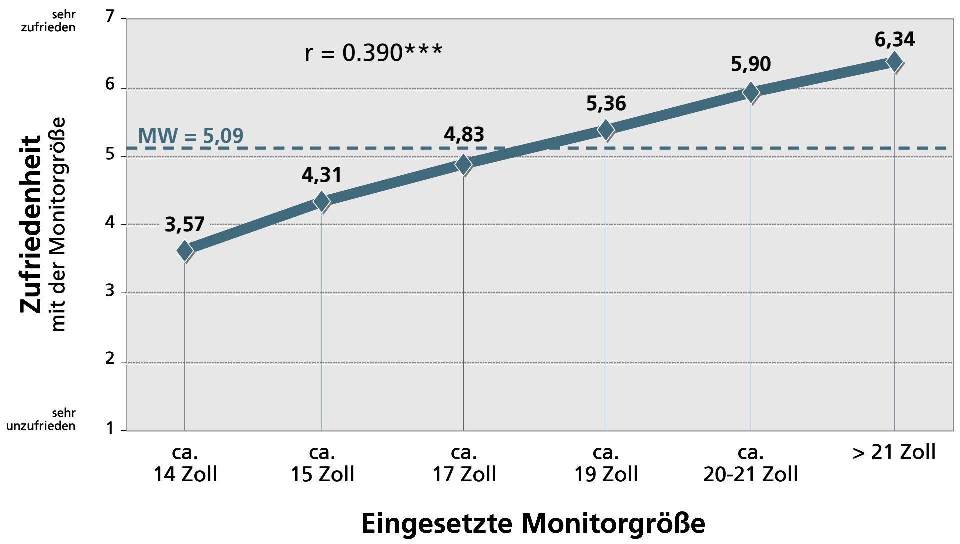


ABB. 46 *Eingesetzte Monitorgrößen und deren Zufriedenheitswerte*

3.5.3 Sicherheit und Datenschutz

Das Bewusstsein für die Bedeutung von Sicherheitsmaßnahmen in der IT ist bei den Befragten grundsätzlich vorhanden und geschärft. Generell sind für alle Teilnehmer sicherheitsspezifische IT-Themen, wie z. B. Virenschutz-Software, Datenverschlüsselung und sichere Authentifizierung von hoher Relevanz. Wie Abb. 47 bestätigt, wird diesen Themen quer über alle Typen eine sehr hohe Bedeutung für die eigene Arbeit zugemessen (MW = 5,88).

Die aktuelle und konkrete Gefährdungslage in puncto IT-Sicherheit wird aus Sicht der Befragten als nicht sehr bedrohlich eingeschätzt. Mit einem Durchschnittswert von MW = 5,53 sehen sich die Teilnehmer insgesamt sehr gut vor Datenverlust oder unerlaubtem Zugriff durch IT-bezogene Sicherheitsmaßnahmen des Unternehmens geschützt.

Im Vergleich der vier Wissensarbeits-Typen untereinander fühlt sich der Typ A diesbezüglich insgesamt betrachtet jedoch etwas weniger gut geschützt ($MW_{IT\text{-}Schutz\_A}$ = 4,99) als die Vertreter der drei anderen Typen. Typ A misst dem Thema IT-Sicherheit jedoch auch eine etwas geringere Bedeutung bei.

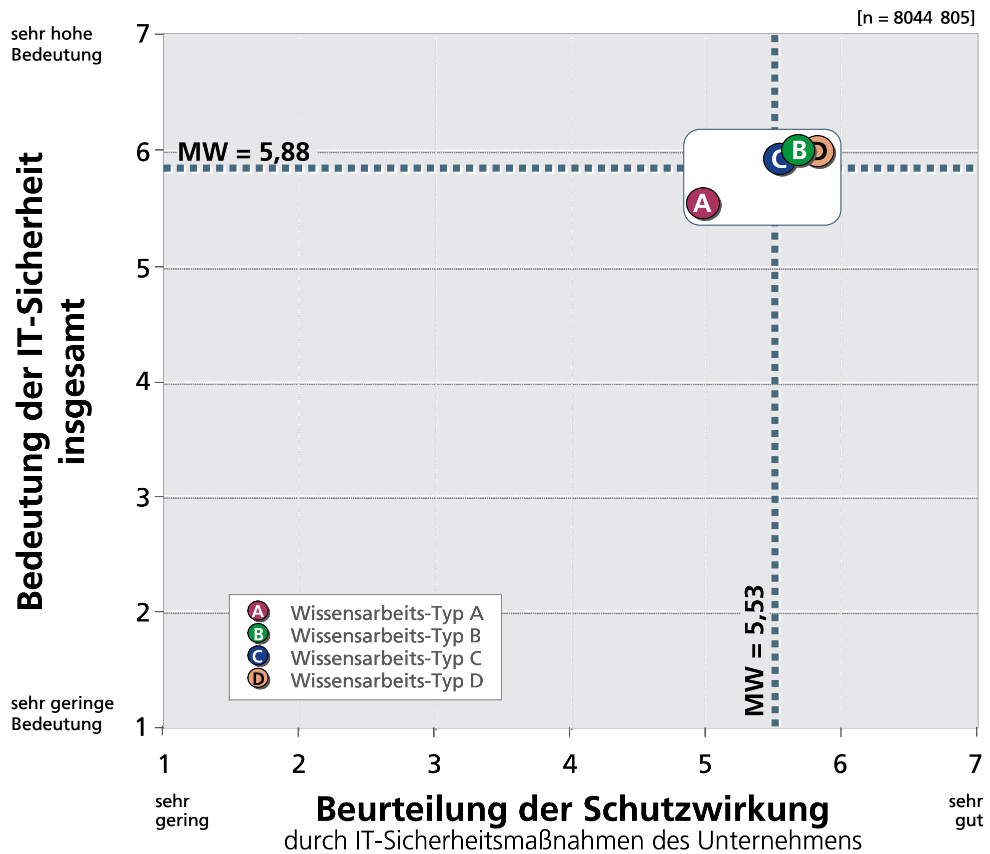


ABB. 47 *Bedeutung der IT-Sicherheit vs. Beurteilung der Schutzwirkung durch IT-bezogene Sicherheitsmaßnahmen des Unternehmens [Mittelwerte insgesamt und spezifische Werte der vier Wissensarbeits-Typen]*

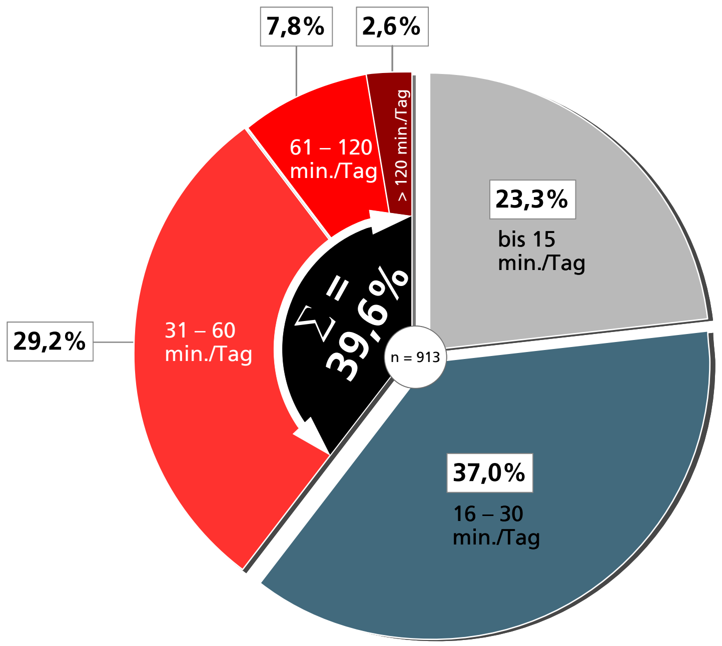
3.5.4 Suchzeiten

Als besonderer »Performance-Killer« hat sich die tägliche Suche nach Dokumenten und Unterlagen im Büro erwiesen. Von allen befragten Teilnehmern benötigen knapp 40% täglich mehr als eine halbe Stunde allein für die Suche nach benötigten, arbeitsrelevanten Dokumenten, Daten, Nachrichten und Unterlagen (Abb. 48).

Alarmierend dabei ist, dass 10,4% der Befragten täglich sogar deutlich mehr als 1 Stunde, teilweise sogar mehr als 2 Stunden für derartige Suchaktionen aufwenden. Untersucht man diese Teilstichprobe mit den hohen Suchzeiten weiter in Abhängigkeit einzelner Wissensarbeits-Typen, so fällt auf, dass insbesondere Vertreter des Typ C hier überdurchschnittlich oft vertreten sind (relativer Anteil 31,6%). Der Typ D hingegen taucht am wenigsten oft auf (relativer Anteil 14,7%) und hat seine Daten und Unterlagen - relativ gesehen - am besten im Griff.

Der bei Typ D festgestellte umfangreichere Technologieeinsatz und die bessere IuK-Qualität machen sich demzufolge auch auf diese indirekte Art und Weise postiv bemerkbar bzw. »bezahlt«.

ABB. 48 *Suchzeiten je Tag nach selbst abgelegten bzw. intern bereits vorhandenen Dokumenten, Informationen oder Dateien*



4 ZUSAMMENFASSUNG UND AUSBLICK


Bild © Nmedia - Fotolia.com

Die Kreativität und das Engagement der Mitarbeiter stärken und zur Sicherung des wirtschaftlichen Erfolgs nutzen – für innovative Unternehmen ist dies ein wesentlicher Erfolgsfaktor.

Als Beitrag zur Identifikation nutzbarer Produktivitätspotentiale im Büro beschreibt die Studie aktuelle Trends und untersucht, wie sich Wissensarbeit in Abhängigkeit eines spezifischen Technologie- und Arbeitsumfeldes ganz unterschiedlich entwickeln kann. Grundlage dazu bildet eine empirische Erhebung von Fraunhofer IAO, die im Rahmen des Verbundforschungsprojektes OFFICE 21® durchgeführt wurde.

Es konnte aufgezeigt werden, dass insgesamt betrachtet Wissensarbeit noch längst nicht auf einem akzeptablen Produktivitäts- und Performance-Niveau angelangt ist. Der recht niedrige Mittelwert von MW= 4,32 entspricht nur einem Durchschnitts-Niveau von 55,3% und weist nachdrücklich auf nach wie vor vorhandene, aber offensichtlich häufig ungenutzt schlummernde Performance- und Produktivitätspotenziale bei Büro- und Wissensarbeit hin.

Für eine differenzierte Betrachtung in Abhängigkeit spezifischer Merkmalsausprägungen zu den Aspekten Neuartigkeit, Komplexität sowie Autonomie- und Freiheitsgrade bei der Arbeit, konnten vier wesentliche Typen von Wissensarbeit identifiziert und beschrieben werden. Deren jeweils spezifisches Arbeitsumfeld – sowohl in technologischer als in räumlich-organisatorischer Hinsicht – zeigt eine Reihe signifikanter Unterschiede auf.

Die vier identifizierten Wissensarbeits-Typen unterscheiden sich deutlich, nicht nur in Bezug auf deren jeweiliges Aufgaben- und Tätigkeitsprofil, sondern auch hinsichtlich:

- der IuK-Ausstattung (Art, Umfang) und deren Qualität (Δ_{IuK} = 12,7%),
- der durchschnittlichen Prozess-Performance ($\Delta_{Perf.}$ = 11,9%)
- sowie in ihren Beurteilungen zu Arbeitszufriedenheit und Wohlbefinden ($\Delta_{Cont.}$ = 20,4%).

Im Sinne einer betriebswirtschaftlichen Investitions- und Renditebetrachtung in Bezug auf die Verbesserung der Prozess-Performance sind demzufolge Ergebnisse im 2-stelligen Bereich durchaus erwartbar und realistisch.

Die hochsignifikanten Korrelationswerte der einzelnen Indizes zeigen, dass sowohl die Qualität der IuK-Ausstattung als auch Contentment- und Autonomieaspekte geeignete Stellhebel zur Performance-Optimierung bieten und sich hierüber gewichtige Performance-Potenziale erschliessen lassen.

Die Erhöhung der (Arbeits-)Autonomie, also das »einräumen und nutzen können« sowohl inhaltlicher und zeitlicher aber auch örtlich/räumlicher Freiheitsgrade und Wahlmöglichkeiten trägt bei Wissensarbeitern nicht nur zur Förderung von Zufriedenheit und Wohlbefinden bei, sondern steigert auch deren Performance.

Zudem gilt: je höherwertiger die Qualität der zur Verfügung stehenden IuK-Ausstattung ist, desto besser ist im allgemeinen auch die Prozess-Performance der Wissensarbeiter.

Es ist auffällig, dass in der Breite über alle Teilnehmer hinweg (5.-95. Perzentil), eine sehr große Streuung bezüglich der zur Verfügung stehenden IuK-Qualität auszumachen ist (Δ_{IuK} = 55,8%). Dies ist ein deutliches Indiz dafür, dass die IuK-Qualität aus Sicht der Befragten in vielen Fällen offensichtlich nicht (mehr) zeitgemäß ist und einen hohen Verbesserungsbedarf aufweist. Ein überaus lohnender Ansatz also, wenn es darum geht, Qualitätsstandards in Bezug auf die technische Ausstattung von Informations- und Wissensarbeitern zu überprüfen und Investitionen zu planen.

In Bezug auf die zur Verfügung stehende und für die tägliche Arbeit genutzte IuK-Ausstattung ergibt sich ein sehr heterogenes Bild. Auffällig dabei ist, dass insbesondere Vertretern des Wissensarbeits-Typ D ein sehr umfangreiches IuK- und Technologiepaket zur Verfügung steht und dies auch »performancewirksam« genutzt werden kann.

Bei der Betrachtung zahlreicher wichtiger Schlüsseltechnologien der Informations- und Kommunikationstechnik bieten nicht nur das aktuelle Anwenderverhalten, sondern auch die Einschätzungen der Befragten in Bezug auf den Nutzen dieser Technologien in Abhängigkeit von dem jeweiligen Typus reichlich Ansatzpunkte zur weiteren Bearbeitung des Marktes.

Wenn die Performance von Wissensarbeitern signifikant erhöht werden soll, gilt es gleichermaßen sowohl organisatorische, technische als auch räumliche Hemmnisse zu beseitigen. Investitionen in eine qualitätsvolle, hochwertige und moderne Technologie- und Technikausstattung sind kein Selbstzweck, sondern verbessern die Performance und Produktivität bei Büro- und Wissensarbeitern entscheidend.

Schon heute darf man gespannt darauf sein, was sich bis zur nächsten geplanten Auswertung zur Fortsetzung der Studienreihe im Jahre 2010 verändern wird. Insbesondere auch deshalb, weil mittlerweile eine englisch-sprachige Version des »Information-Worker-Check« über das auch weiterhin frei zugänglich Portal verfügbar ist.

5 ANHANG

5.1 ABKÜRZUNGSVERZEICHNIS

AV	Arbeitsvorbereitung
CRM	Customer relations management (zentrales Kontakt-Management)
CTI	Computer-Telephonie-Integration
DMS	Dokumenten-Managementsystem
FM	Facility Management
IP	Internet Protokoll
IT	Informationstechnologie
IuK	Informations- und Kommunikations(-Technologie)
ICont.	Index Contentment
IIuK.	Index Iuk-Qualität
IPerf.	Index Prozess-Performance
IWC	Information-Worker-Check
MW	Mittelwert
PC	Personal Computer
QM	Qualitätsmangement
WLAN	Wireless local area network

5.2 PROJEKTPARTNER OFFICE 21®

Unter der wissenschaftlichen Leitung durch das Fraunhofer IAO bündeln im Verbundforschungsprojekt OFFICE 21® derzeit 20 Partner aus unterschiedlichen Branchen ihre Kompetenzen zu Zukunftsfragen innovativer Arbeits- und Bürowelten.

Ziel des Verbundforschungsprojekts OFFICE 21® ist es, durch eine enge Kooperation von Anwendern, Herstellern und angewandter Forschung zielgerichtete Akzente bei der ganzheitlichen Durchdringung des gesamten Themenfeldes »Büro« zu setzen und damit notwendige Innovationsprozesse zu initiieren.

Detaillierte Informationen zum Verbundforschungsprojekt OFFICE 21®, den Partnern sowie zu bisherigen und aktuellen Forschungsschwerpunkten sind im Internet unter www.office21.de abrufbar.

Als Beitrag zur Identifikation nutzbarer Produktivitätspotentiale im Büro beschreibt die Studie aktuelle Trends und zeigt auf, wie sich stationäre bzw. mobile Wissensarbeit in einem spezifischen Technologie- und Arbeitsumfeld ganz unterschiedlich entwickelt. Grundlage dazu bildet eine empirischen Erhebung von Fraunhofer IAO, die im Rahmen des Verbundforschungsprojektes OFFICE 21® durchgeführt wurde.

Neben zahlreichen explorativen Gesamtauswertungen zum allgemeinen Arbeitsumfeld von Bürobeschäftigten werden insbesondere die Unterschiede bei vier speziellen Typen von Wissensarbeitern näher untersucht und beschrieben. Und zwar sowohl im Hinblick auf typische technologische Ausstattungs- und Qualitätsmerkmale als auch in Bezug auf damit verbundene Kennwerte für deren Prozess-Performance, Arbeitszufriedenheit und Wohlbefinden.

Die Studie wurde gefördert und durchgeführt im Rahmen des Verbundforschungsprojektes

OFFICE 21®
Zukunft der Arbeit

Fraunhofer IRB Verlag
ISBN 978-3-8396-0007-8

Anlage 8

Kaya, Naz: Relationship between color and emotion: a study of college students

Relationship between color and emotion: a study of college students

Naz Kaya

Ninety-eight college students were asked to indicate their emotional responses to five principle hues (i.e., red, yellow, green, blue, purple), five intermediate hues (i.e., yellow-red, green-yellow, blue-green, purple-blue, and red-purple), and three achromatic colors (white, gray, and black) and the reasons for their choices. The color stimuli were referenced from the Munsell Color System. The results revealed that the principle hues comprised the highest number of positive emotional responses, followed by the intermediate hues and the achromatic colors. The color green evoked mainly positive emotions such as relaxation and comfort because it reminded most of the respondents of nature. The color green-yellow had the lowest number of positive responses because it was associated with vomit and elicited the feelings of sickness and disgust. For the achromatic colors, white attained a large number of positive responses, followed by the colors black and gray. The reasons for the color-emotion associations are discussed and future research areas are suggested.

Introduction

Color is an inseparable part of our everyday lives and its presence is evident in everything that we perceive. It is widely recognized that colors have also a strong impact on our emotions and feelings (Hemphill, 1996; Lang, 1993; Mahnke, 1996). For instance, the color red has been associated with excitement, orange has been perceived as distressing and upsetting, purple as dignified and stately, yellow as cheerful, and blue has been associated with comfort and security (Ballast, 2002; Wexner, 1982). Moreover, some colors may be associated with several different emotions and some emotions are associated with more than one color (Linton, 1999, Saito, 1996). Red, symbolically known as a dominant and dynamic color, has an exciting and stimulating hue effect. It has both positive and negative impressions such as active, strong, passionate, warm, but on the other hand aggressive, bloody, raging and intense. Green has been found to have a retiring and relaxing effect. It too has both positive and negative impressions such as refreshment, quietness, naturalness, and conversely tiredness and guilt (Davey, 1998, Mahnke, 1996, Saito, 1996).

The relationship between color and emotion is closely tied to color preferences. In particular, color preferences are associated with whether a color elicits positive or negative feelings. While particular colors have been found to be highly preferred regardless of age, racial group, or culture (Adams & Osgood, 1973, Eysenck, 1941), there is some evidence that color preference may be culturally-based. For example, Choungourian (1968) found that the colors red and blue were the most preferred colors among American subjects, but were less preferred in other cultures. In a comparision of Japanese and Korean subjects, Saito (1996) found unique color preference tendencies between the two countries, and also with respect to age, gender, and geographical region within the individual country.

In an investigation of children's emotional associations with colors, Boyatzis and Varghese (1994) found that light colors (e.g., yellow, blue) are associated with positive emotions (e.g., happy, strong) and dark colors (e.g., black, gray) with negative emotions (e.g., sad, angry). In a study examining color-emotion associations among college students in Australia, Hemphill (1996) also found that bright colors elicited mainly positive emotional associations, while dark colors elicited negative emotional associations, confirming the results obtained by Boyatzis and Varghese (1994). However, Saito (1996) found that the color black elicited both negative and positive responses among Japanese subjects, and that black was often a preferred color among young people.

Colors can also be described in temperature terms, such as "warm" or "cool"as related to the dominant wavelength of the color. The cool colors (e.g., blue, green, purple) are generally considered to be restful and quiet, while the warm colors (e.g., red, yellow, orange) are seen as active and stimulating (Ballast, 2002). As cited in Lang (1993), Grandjean made observations about the effects of color on perceptions of room size and psychological response noting that cool colors such as blue and green make a space restful and increase spaciousness; however warm colors such as red, orange, and yellow make a space less spacious, while increasing stimulation. Furthermore, people exposed to red and yellow colors reported higher levels of anxiety than did people exposed to cool blue and green colors (Kwallek, Lewis, & Robbins, 1988; Mahnke & Mahnke, 1993). However, in other studies, no relationships have been found between the individuals' mood states and colors (Ainsworth, Simpson, & Cassell, 1993; Kwallek, Lewis, Lin-Hsiao, & Woodson, 1996).

Of the numerous color systems that exist (see Jacobson & Bender, 1996 for discussion), one color system noted internationally for its precise identification process is the Munsell Color System (Ballast, 2002; Valdez & Mehrabian, 1994). According to this system, each color has three basic attributes: hue, value (brightness), and chroma (saturation). Hue is the first attribute of a color by which we distinguish one color from another (e.g., blue from red, green from yellow). There are 10 hues, five of which are identified as principal hues (i.e., red, yellow, green, blue, and purple) and the other five are intermediate hues (i.e., yellow-red, green-yellow, blue-green, purple-blue, and red-purple). Value, the second attribute of color, describes the degree of lightness or darkness of a color in relation to white and black. Black, white and the shades of gray are called neutral (achromatic) colors. The third attribute of a color is chroma, which is the degree of purity or vividness of the hue (i.e., with high saturated colors containing less gray) when compared with a neutral gray of the same value (Ballast, 2002).

Despite a rapidly growing literature on the impact of color on our emotions and considerable interest in this research area, many studies have failed to use color samples from a standardized system of color notation (Boyatzis & Varghese, 1994; Hemphill, 1996; Terwogt & Hoeksma, 1995), while others elicited individuals' responses to verbal labels of color (e.g., "red", "blue") instead of using actual color stimuli (Hupka, Zaleski, Otto, Reidl, & Tarabrina, 1997). Furthermore, several studies have used color-emotion matching tasks (Zentner, 2001); matching colors (e.g., red, yellow, blue) to a certain number of emotions (e.g., happiness, sadness, anger), which results in limited assessments of reactions to colors.

The purpose of the present study was to examine college students' color-emotion associations, referencing color samples from the standardized Munsell Color System and to investigate the reasons for students' emotional reactions to each color.

Method

Participants

The sample consisted of 98 volunteered college students (44 men and 54 women) at a public institution in the southeast. The mean age was 21 years (range = 18-25 years). None of the participants had defective color vision as verified with the Ishihara Color Deficiency test (1993).

Stimuli

Ten fully saturated chromatic colors were chosen from the Munsell Color System: red, yellow, green, blue, purple, yellow-red, green-yellow, blue-green, purple-blue, and red-purple. The Munsell notations are shown in Table 1. The color samples were prepared by using Freehand 10.0 software, in which Munsell color notations were available in that computer program. Apart from these ten hue groups, three achromatic colors (white, black and middle gray) were also used.

Procedure

Participants were tested individually in an office space where they were seated at a personal computer. Each color sample (10 cm _ 12 cm) was displayed in the middle of the computer screen one at a time on a neutral gray background, Munsell N/7. Order of presentation of the color samples was randomized across participants. Participants were asked, "What emotional response do you associate with this color? How does this color make you feel?" and "Why do you feel this way?" These questions were adapted from Boyatzis and Varghese (1994) and Hemphill (1996). Students were allowed to state only one emotional response for each color. Their answers were recorded on an observation sheet. Each experimental session lasted for about ten minutes.

Results

Data were analyzed using Statistical Package for Social Sciences (SPSS) software program. Descriptive statistics were used to summarize data. Based on the results obtained from the student's responses, a total of twenty-two emotions were gathered (see Table 2). Some of the emotions had the same meaning (e.g., empty, void) and some were overlapped (e.g., happy, happiness, joy), so they were grouped under the same emotion category. There was also a category for those responses that indicated no emotional response.

Because of the low frequencies in several cells, the emotions were coded as "positive", "negative", or "no emotion" (see Table 3). Overall, 62.2% of the participants expressed positive responses to colors, 34.2% expressed negative responses, and 3.6% expressed no emotion. About 80% of the responses to the principle hues, including red, yellow, green, blue, and purple were positive, compared with only 29.2% for the achromatic colors, including white, gray, and black (see Table 3). Only 17.8% of the responses to the principle hues were negative, whereas 68.4% of the responses were negative for the achromatic colors.

As shown in Table 3, the color green attained the highest number of positive responses (95.9%), closely followed by yellow (93.9%). The majority of emotional responses for the green color indicated the feelings of relaxation and calmness, followed by happiness, comfort, peace, hope, and excitement. Green was associated with nature and trees, and thus creating feelings of comfort and soothing emotions. The color yellow was generally seen to be lively and energetic and elicited positive emotions including happiness and excitement because it was associated with the sun, blooming flowers, and summer time.

Among the principle hues, the next highest number of positive response was given for the color blue (79.6%), followed by red and purple (64.3% each). Blue revealed the feelings of relaxation and calmness, followed by happiness, comfort, peace, and hope. The negative emotions for the color blue were sadness, depression, and loneliness (see Table 2). Furthermore, the color red prompted both positive and negative emotional reactions. Red was seen to be positive because it was associated with love and romance, while the negative aspects of red included having associations with fight and blood as well as Satan and evil. Finally, the color purple elicited the feelings of relaxation and calmness, followed by happiness, sadness, tiredness, power, fear, boredom, excitement, and comfort (see Table 2). The positive aspects of purple are tended to be mainly associated with children and laughing, while reasons given for negative responses to purple consistently showed that purple was not a favorite', color.

For the intermediate hues, the majority of emotional reactions (64.5%) were positive. As shown in Table 3, blue-green elicited the highest number of positive responses (81.6%), followed by red-purple (76.5%), yellow-red (75.4%), and purple-blue (65.3%). On the contrary, the color green-yellow elicited the highest number (71.4%) of negative emotional responses because it was associated with vomit and elicited the feelings of sickness and disgust (see Table 2 and 3).

For the achromatic colors, white attained a large number of positive responses (61.2%), compared with only 19.4% and 7.1% positive responses for black and gray, respectively. White was seen to be positive and was associated with the feelings of innocence, peace, and hope because it tended to be related with purity and being simple and clean. Further, it reminded some respondents of bride, snow, dove, and cotton. Reasons given for negative emotional responses to white consistently showed that white elicited the feelings of emptiness, loneliness, and boredom. In addition, the color black was seen to evoke negative emotions such as sadness, depression, fear, and anger because it was associated with death; mourning and tragic events as well as darkness and night time. The positive aspects of black were richness, wealth, and power. It also reminded some respondents of tuxedos and formal gowns.

Finally, the color gray was mainly associated with negative emotions (89.8%); including the feelings of sadness, depression, boredom, and confusion, as well as tiredness, loneliness, anger, and fear. Reasons given for negative emotional responses to gray consistently showed that the color gray tends to make reference to bad weather, rainy, cloudy or foggy days and brings out the feelings of sadness, depression, and boredom.

Discussion

The primary goal of this study was to examine the color-emotion associations among college students, referencing color stimuli from the standardized Munsell Color System. Based on Munsell Color System, the present study used five principle (i.e., red, yellow, green, blue, and purple) and five intermediate hues (i.e., yellow-red, green-yellow, blue-green, purple-blue, and red-purple), in addition to three achromatic colors (i.e., white, gray, and black). Overall, the participants' responses of color-emotion associations for the principle hues were positive (79.6%), compared with the positive responses for the intermediate hues (64.5%) and achromatic colors (29.2%). The color green elicited mainly positive emotional responses, including the feelings of relaxation, calmness, and happiness as well as comfort, peace, and hope. This is somewhat in agreement with the findings of Saito (1996), whose subjects found green to be refreshing and beautiful. Reasons given for positive responses to green showed that green was associated with nature, grass, trees, and reminds someone of outdoors and springtime, consistent with Hemphill's (1996) findings. Similarly, Saito (1996) noted that some of the Asian subjects who preferred green indicated the positive feeling about the color because of its association with the image of a forest.

Blue elicited a high number of positive emotional responses, including the feelings of relaxation and calmness, happiness, comfort, peace, and hope, with a low number of negative responses, including sadness and depression. Reasons that blue elicited positive emotions seem to be because many participants associated the color blue with the ocean, beach, water, or the sky and thus inducing relaxing and calming effect. Blue evoked negative emotions because it was associated with the night and dark skies, thus making someone feel depressed. One respondent said blue made her sad because "it makes you feel blue". Interestingly, Saito (1996), who found that vivid blue was the preferred color among all of the Asian groups, noted only positive aspects related to the color blue, namely refreshing, beautiful, and bright.

Colors are rich with symbolism. This symbolism can be apparent in how an individual associates colors with things, objects or physical space. For instance, in the present study, the color yellow-red was associated with the color of autumn or Halloween. One respondent said that yellow-red made her happy because "it reminds me of school buses and my childhood". Furthermore, the color blue-green was associated not only with the ocean and the sky, hut also reminded some respondents of cool mints and toothpaste. Red-purple was associated with the color of red wine, plum, bridesmaid dress, or the color of a bedroom. One said red-purple makes her feel happy because "it reminds me of being in love". In addition, the color red was associated not only with love and romance, but also with evil, Satan and blood. One respondent said that the color red reminded her of Valentine's Day and the shape of heart. Another said that the color reminds her of red lingerie and Victoria's Secret. Some associated black with "power," and said it reminded them of nice sport cars. Black made some respondents feel sophisticated and reminded them of "fashion and clothing". Yet, another respondent said black made him sad and reminded him of "funerals where people wear black". Therefore, it seems that a color-related emotion is highly dependent on personal preference and one's past experience with that particular color. A replication of this study at different institutions in the United States should give us a more comprehensive understanding of the issues raised here. Cross-site studies could be conducted to identify similar or different patterns in students' emotional associations to colors.

Moreover, color conventions differ from one society to another. In Western cultures, red is supposed to be a fiery color, green is said to be soothing. Another well-known example is with the two achromatic colors, black and white. Black is accepted as the symbolism of mourning in some countries, however it symbolizes wedding in some others (Linton, 1991). Many attempts have been made to identify the impact of various hues, but it cannot be ascertained whether these reactions are innate or cultural. For example, death and mourning are associated with the color black in Western traditions, whereas in China the color of death is white. Our findings of both positive and negative feelings about the color black were in agreement with those of Saito (1996), although the specific associations differed between the two studies. In the present study, the color black was associated not only with royalty, power, and wealth, but also with death, mourning, and tragic events. Saito noted positive images of clearness, tightness, sharpness, dignity, and nobleness, but negative associations with anxiety, fear, sin, and death. Saito (1996) also found a very strong preference for the color white among the Asian groups studied, particularly the Japanese subjects. Within Saito's study, white was found to be positively associated with the feelings of being clean, pure, harmonious, refreshing, beautiful, clear, gentle, and natural. Saito (1996) further explained the possible influence of ancient Japanese religion and mythology on the Japanese preference for white. A small number of Saito's subjects in Taipei expressed a negative feeling toward white, indicating an association with the image of death. In the present study, the findings revealed that the color white was seen to be generally positive and was associated with purity and being simple and clean. Some respondents associated white with innocence and peace and said it reminded them of a bride or dove. Another said the color white reminded her of snow. However, it also evoked negative emotions and was associated with emptiness and void. Some associated white with loneliness and boredom because it reminded them of insane asylum. Cross-cultural research could shed light on these issues by determining how cultural differences vary in color-emotion associations.

In addition to cross-cultural studies and investigation to reasons for color associations, future work might also utilize rating scales for color associations, such as "beautiful-ugly", "warm-cold", etc. that have been studied by Kawamoto and Soen (1993). Also, in our study, all colors were presented on a neutral background. Future work might involve investigation of color-emotion associations in which colors are presented on different colored backgrounds. This could lead to investigations of feelings about color harmony and color associations, which have been studied by others (Sivik & Hard, 1994).

Table 1

Munsell Notations for Color Samples

Color	Hue	Value/Chroma
Red	5R	5/14
Yellow	7.5Y	9/10
Green	2.5G	5/10
Blue	10B	6/10
Purple	5P	5/10
Yellow-red	5YR	7/12
Green-yellow	2.5GY	8/10
Blue-green	5BG	7/8
Purple-blue	7.5PB	5/12
Red-purple	10RP	4/12
White	N/9	
Gray	N/5	
Black	N/1	

Table 2
Frequencies of Emotional Reactions Given to Each Color

Emotions *	Red	Yellow	Green	Blue	Purple
Angry (a)	28 (28.6)	0	0	0	0
Annoyed (a)	0	0	0	0	0
Bored (a)	0	0	0	0	5 (5.1)
Calm (b)	4 (4.1)	0	29 (29.6)	60 (61.2)	28 (28.6)
Comfortable (b)	0	0	15 (15.3)	4 (4.1)	3 (3.1)
Confused (a)	0	0	0	0	0
Depressed (a)	0	0	0	6 (6.1)	0
Disgusted (a)	0	0	0	0	0
Empty/void (a)	0	0	0	0	0
Energetic (b)	5 (5.1)	10 (10.2)	0	0	0
Excited (b)	18 (18.4)	8 (8.2)	2 (2.0)	0	4 (4.1)
Fearful (a)	0	0	0	0	5 (5.1)
Happy (b)	21 (21.4)	74 (75.5)	28 (28.6)	10 (10.2)	21 (21.4)
Hopeful (b)	0	0	8	0	0

			(8.2)		
Innocent (b)	0	0	0	0	0
Lonely (a)	0	0	0	3 (3.1)	0
Loved (b)	15 (15.3)	0	0	0	0
Peaceful (b)	0	0	12 (12.2)	4 (4.1)	0
Powerful (b)	0	0	0	0	7 (7.1)
Sad (a)	4 (4.1)	0	0	8 (8.2)	13 (13.3)
Sick (a)	0	0	0	0	0
Tired (a)	0	6 (6.1)	0	0	9 (9.2)
No emotion	3 (3.1)	0	4 (4.1)	3 (3.1)	3 (3.1)

Emotions *	Yellow-red	Green-yellow	Blue-green	Purple-blue
Angry (a)	0	0	0	0
Annoyed (a)	5 (5.1)	8 (8.2)	7 (7.1)	0
Bored (a)	4 (4.1)	2 (2.0)	0	0
Calm (b)	0	0	16 (16.3)	38 (38.8)
Comfortable (b)	3 (3.1)	7 (7.1)	7 (7.1)	0
Confused (a)	0	2 (2.0)	6 (6.1)	0
Depressed (a)	0	0	0	12 (12.2)
Disgusted (a)	9 (9.2)	26 (26.5)	2 (2.0)	0
Empty/void (a)	0	0	0	0
Energetic (b)	14 (14.3)	0	10 (10.2)	0
Excited (b)	25 (25.5)	6 (6.1)	11 (11.2)	0
Fearful (a)	0	0	0	0
Happy (b)	31 (31.6)	11 (11.2)	36 (36.7)	13 (13.3)
Hopeful (b)	0	0	0	5 (5.1)

Innocent (b)	0	0	0	0
Lonely (a)	0	0	0	3 (3.1)
Loved (b)	0	0	0	0
Peaceful (b)	0	0	0	8 (8.2)
Powerful (b)	0	0	0	0
Sad (a)	0	0	0	10 (10.2)
Sick (a)	0	32 (32.7)	0	0
Tired (a)	0	0	0	5 (5.1)
No emotion	7 (7.1)	4 (4.1)	3 (3.1)	4 (4.1)

Emotions *	Red-purple	White	Gray	Black
Angry (a)	0	0	3 (3.1)	7 (7.1)
Annoyed (a)	2 (2.0)	0	0	0
Bored (a)	2 (2.0)	5 (5.1)	14 (14.3)	0
Calm (b)	13 (13.3)	8 (8.2)	5 (5.1)	0
Comfortable (b)	0	0	0	5 (5.1)
Confused (a)	0	0	6 (6.1)	0
Depressed (a)	8 (8.2)	0	23 (23.5)	22 (22.4)
Disgusted (a)	3 (3.1)	0	0	0
Empty/void (a)	0	25 (25.5)	0	0
Energetic (b)	0	0	0	0
Excited (b)	12 (12.2)	0	0	0
Fearful (a)	0	0	3 (3.1)	17 (17.3)
Happy (b)	26 (26.5)	0	0	0
Hopeful (b)	0	6 (6.1)	0	0

Innocent (b)	0	33 (33.7)	0	0
Lonely (a)	0	6 (6.1)	4 (4.1)	0
Loved (b)	17 (17.3)	0	0	0
Peaceful (b)	0	13 (13.3)	0	0
Powerful (b)	7 (7.1)	0	2 (2.0)	14 (14.3)
Sad (a)	0	0	30 (30.6)	24 (24.5)
Sick (a)	0	0	0	0
Tired (a)	0	0	5 (5.1)	7 (7.1)
No emotion	8 (8.2)	2 (2.0)	3 (3.1)	2 (2.0)

Note. Emotions are listed in alphabetical order.

The cell numbers indicate frequencies; the percentages are listed in parentheses.

(a) Negative emotions

(b) Positive emotions

Table 3
Color and Emotion Associations

	Emotional Association		
Color	Positive	Negative	No emotion
Principle Hues			
Red	63 (64.3)	32 (32.7)	3 (3.1)
Yellow	92 (93.9)	6 (6.1)	0
Green	94 (95.9)	0	4 (4.1)
Blue	78 (79.6)	17 (17.3)	3 (3.1)
Purple	63 (64.3)	32 (32.7)	3 (3.1)
Total	390 (79.6)	87 (17.8)	13 (2.6)
Intermediate Hues			
Yellow-red	73 (74.5)	18 (18.4)	7 (7.1)
Green-yellow	24 (24.5)	70 (71.4)	4 (4.1)
Blue-green	80 (81.6)	15 (15.3)	3 (3.1)
Purple-blue	64 (65.3)	30 (30.6)	4 (4.1)

Red-purple	75 (76.5)	15 (15.3)	8 (8.2)
Total	316 (64.5)	148 (30.2)	26 (5.3)
Achromatic Colors			
White	60 (61.2)	36 (36.7)	2 (2.0)
Gray	7 (7.1)	88 (89.8)	3 (3.1)
Black	19 (19.4)	77 (78.6)	2 (2.0)
Total	86 (29.2)	201 (68.4)	7 (2.4)
Overall	792 (62.2)	436 (34.2)	46 (3.6)

Note. The cell numbers indicate frequencies; the percentages are listed in parentheses.

References

Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meaning of color. Journal of Cross-Cultural Psychology, 7, 135-157.

Ainsworth, R. A., Simpson, L, & Cassell, D. (1993). Effects of three colors in an office interior on mood and performance. Perceptual & Motor Skills, 76, 235-241.

Ballast, D. K. (2002). Interior design reference manual. Professional Pub. Inc.: Belmont, CA.

Boyatzis, C. J., & Varghese, R. (1994). Children's emotional associations with colors. Journal of Genetic Psychology, 155, 77-85.

Choungourian, A. (1968). Color preference and cultural variation. Perceptual & Motor Skills, 26, 1203-1206.

Davey, P. (1998). True colors: The glorious polychromy of the past suggests a strong historical need for color, despite current reductive fashions. The Architectural Review, 204, 34-36.

Eysenck, H. J. (1941). A critical and experimental study of color-preferences. American Journal of Psychology, 54, 385-394.

Hemphill, M. (1996). A note on adults' color-emotion associations. Journal of Genetic Psychology, 157, 275-281.

Hupka, R. B., Zaleski, Z., Otto, J., Reidl, L., & Tarabrina, N. V. (1997). The colors of anger, envy, fear, and jealousy: A Cross-cultural study. Journal of Cross-Cultural Psychology 28(2), 156-171.

Ishihara, S. (1993). Ishihara's test for color-blindness. Tokyo, Japan: Kanehara & Co. Ltd.

Jacobson, N., & Bender, W. (1996). Color as a determined communication. IBM Systems Journal, 35(3,4), 526-538.

Kawamoto, N., & Soen, T. (1993). Objective evaluation of color design II. Color Research and Application, 4, 260-266.

Kwallek, N., Lewis, C. M., & Robbins, A. S. (1988). Effects of office interior color on workers' mood and productivity. Perceptual & Motor Skills, 66, 123-128.

Kwallek, N., Lewis, C. M., Lin-Hsiao, J. W. D., & Woodson, H. (1996). Effects of nine monochromatic office interior colors on clerical tasks and worker mood. Color Research and Application, 21(6), 448-458.

Lang, J. (1993). Creating architectural theory: The role of the behavioral sciences in environmental design. New York: Van Nostrand Reinhold.

Linton, H. (1991). Color consulting: A Survey of international color design. New York: Van Nostrand Reinhold.

Linton, H. (1999). Color in architecture: Design methods for buildings, interiors and urban spaces. New York: McGraw Hill.

Mahnke F. (1996). Color, environment, human response. New York: Van Nostrand Reinhold.

Mahnke, F. H., & Mahnke, R. H. (1993). Color and light in man-made environments. New York: Van Nostrand Reinhold.

Saito, M. (1996). Comparative studies on color preference in Japan and other Asian regions, with special emphasis on the preference for white. Color Research and Application, 21 (1), 35-49.

Sivik, L., & Hard, A. (1994). Some reflections on studying color combinations. Color Research and Application, 19(4), 286-295.

Terwogt, M. M., & Hoeksma, J. B. (1995). Colors and emotions: Preferences and combinations. The Journal of General Psychology, 122(1), 5-17.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. Journal of Experimental Psychology, 123(4), 394-409.

Wexner, L. B. (1982). The degree to which colors are associated with mood-tones. Journal of Applied Psychology, 6, 432-435.

Zentner, M. R. (2001). Preferences for colors and color-emotion combinations in early childhood. Developmental Science, 4(4), 389-398.

NAZ KAYA, Ph.D.

Assistant Professor

HELEN H. EPPS, PH.D.

Professor

Dawson Hall

Department of Textiles, Merchandising and Interiors

College of Family and Consumer Sciences

The University of Georgia

Anlage 9

Boyce, Peter. R. u.a.: Lighting Quality and Office Work: A Field Simulation Study, 2003

PNNL-14506

Lighting Quality and Office Work: A Field Simulation Study

Peter R. Boyce[1]
Jennifer A. Veitch[2]
Guy R. Newsham[2]
Michael Myer[1]
Claudia Hunter[1]

Judith H Heerwagen, Research Manager
Carol C. Jones, Project Manager

December 2003

Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RL01830

---

[1] Lighting Research Center, Rensselaer Polytechnic Institute
RPI # A11106
[2] National Research Council of Canada, Institute for Research in Construction
NRC # B3214.1

PNNL-14506

Lighting Quality and Office Work: A Field Simulation Study

Peter R. Boyce[1]
Jennifer A. Veitch[2]
Guy R. Newsham[2]
Michael Myer[1]
Claudia Hunter[1]

Judith H Heerwagen, Research Manager
Carol C. Jones, Project Manager

December 2003

Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

---

[1] Lighting Research Center, Rensselaer Polytechnic Institute
RPI # A11106
[2] National Research Council of Canada, Institute for Research in Construction
NRC # B3214.1

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.

Lighting Quality and Office Work: A Field Simulation Study

Peter R. Boyce[1], Jennifer A. Veitch[2], Guy R. Newsham[2],
Michael Myer[1], and Claudia Hunter[1]

1. Lighting Research Center, Rensselaer Polytechnic Institute,
21, Union Street, Troy, New York, 12180-3352
Tel: 518-687-7130; Fax: 518-687-7120
boycep@rpi.edu

2. National Research Council of Canada, Institute for Research in Construction,
Bldg M-24 1200 Montreal Road, Ottawa, Ontario KIA OR6, Canada
Tel: 613-993-9671 / 993-9607; Fax: 613-954-3733
jennifer.veitch@nrc-cnrc.gc.ca / guy.newsham@nrc-cnrc.gc.ca

A Report for the
Light Right Consortium

167 pages

September 26, 2003

PNNL RFP # 404141 / RPI # A11106 / NRC # B3214.1

Executive Summary

The question this study addressed was, "Can different forms of realistic office lighting affect the performance of office work or the health and well-being of employees?" An office was furnished as a typical open plan workplace for nine workers, with perimeter windows allowing access to a view but limited daylight penetration. Two experiments were conducted. In Experiment 1, there were provisions for changing between four lighting installations:

- Base Case: A regular array of parabolic-louvered luminaires
- Best Practice: A linear system of direct / indirect luminaires, together with some wall-washing to brighten the walls, with the same average illuminance as the Base Case
- Switching Control: The same as the Best Practice but with a moveable desk lamp having three manually switched light outputs, allowing the individual to increase the illuminance in the work area.
- Dimming Control: Direct / indirect luminaires suspended over the center of each cubicle, together with the wall-washing system. The direct component of each suspended luminaire could be dimmed using an interface on the occupant's computer, allowing the individual to adjust the illuminance in the work area over a wide range.

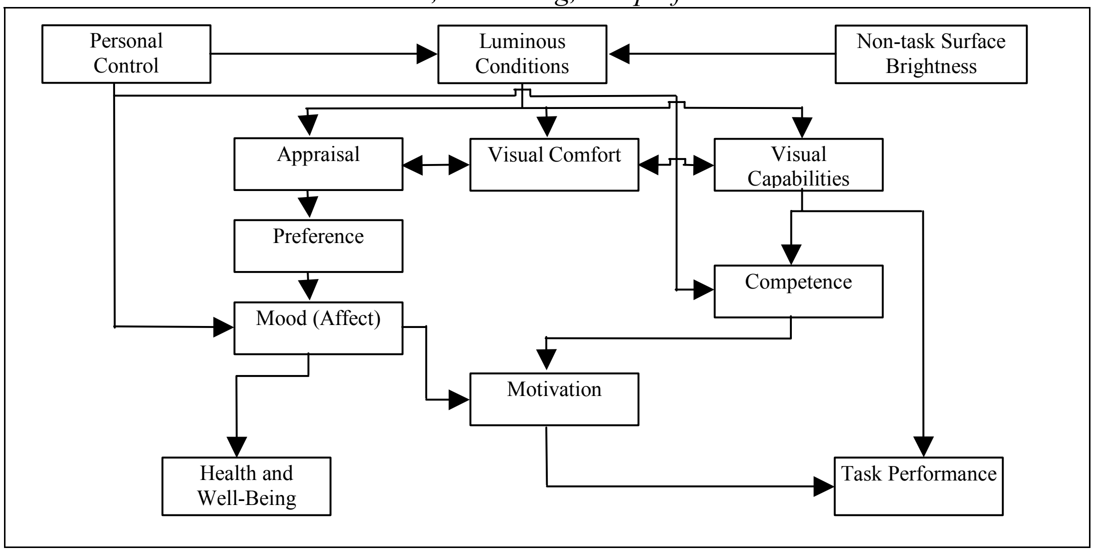
Experiment 2 contrasted two lighting installations:

- Base Case 2: A regular array of recessed prismatic luminaires
- Best Practice 2: A linear system of direct / indirect luminaires, together with some wall-washing to brighten the walls, at an illuminance level approximately 27% lower than the Best Practice in Experiment 1.

We hypothesized that occupants would prefer, and would perform better in, the Best Practice condition compared to the Base Case, and that having control would result in a further improvement in outcomes. In addition, workstations were decorated using one of two surface reflectances, a light gray and a dark blue. This provided greater variation in vertical surface luminance, which previous research has suggested might influence responses to the luminous environment in concert with lighting system effects.

The experiments were designed to test the effects of lighting systems on performance and well-being, as well as indications of the processes that might mediate these effects. What these processes might be is indicated in the Linked Mechanisms Map (Figure ES1), which evolved from the original RFP issued by the Light Right Consortium.

Figure ES1. Linked Mechanisms Map hypothesized to link luminous conditions with health, well-being, and performance.



Both experiments collected data from temporary office workers, who were hired to work under one of the lighting installations for a complete day. During that day the participants carried out tasks involving many forms of clerical and cognitive office work, evaluations of the physical environment, and assessments of their mood. In Experiment 1, we also collected data on switching/dimming choices in the designs in which control was offered.

Experiment 1 had two parallel experimental designs:

1. A between-groups design, in which each person was randomly assigned to one of the four lighting conditions, none of which they had previously seen. A total of 181 individuals contributed data to this portion of the experiment.
2. A repeated-measures design, in which 45 individuals returned for a second day of testing under a different lighting condition. This portion of the experiment involved the Base Case and Dimming Control conditions only.

Experiment 2 had a between-groups experimental design, in which each person was randomly assigned to one of the two lighting conditions, neither of which they had previously seen. A total of 107 people participated in this experiment. Experiment 2 was an extension and partial replication of Experiment 1.

The data from the two experiments were analyzed and interpreted separately. Within each experiment, comparisons between different lighting and partition reflectance conditions were analyzed using orthogonal planned comparisons, which provide for independent tests of specific effects chosen based on the underlying theory. Multivariate analysis of variance was used for sets of measurements relating to concepts for which there were multiple dependent measures, or univariate analysis of variance when there was only one measurement relating to a concept. To examine the validity of the Linked Mechanisms Map, each link in Figure 1 was tested individually, across all lighting conditions.

Results

- **Linked Mechanisms Found.** For the lighting installations examined, the Linked Mechanisms Map supports 5 conclusions, two of which involve paths that were not in the original model (see Figure ES2). The first three findings were obtained in both experiments (personal control was not a variable in Experiment 2):

1. Vision Path (Luminous Conditions ➔ Visual Capability ➔ Task Performance). Lighting and task conditions that improve visibility lead to better task performance. This extends laboratory knowledge to a more realistic setting.
2. Appraisal Path (Luminous Conditions ➔ Appraisal ➔ Preference ➔ Mood ➔ Health & Well-Being). People who are more satisfied with their lighting rate the space as more attractive, are happier, and are more comfortable and satisfied with their environment and their work. This is the first time that this complete path has been demonstrated.
3. Cross Links. The proposed linkages between the Appraisal Path and the Vision Path were not found. Visual Comfort did not predict Visual Capability; this might be because none of the lighting conditions were very uncomfortable. A similar explanation might hold for the (absent) link between Mood and Motivation.
4. Personal Control Path 1 (Personal Control ➔ Health & Well-Being). People with dimming control reported higher ratings of lighting quality, overall environmental satisfaction, and self-rated productivity. This does not appear to be mediated by effects on Mood.
5. Personal Control Path 2 (Personal Control ➔ Motivation ➔ Task Performance). People with dimming control showed more sustained motivation, and improved performance on a measure of attention.

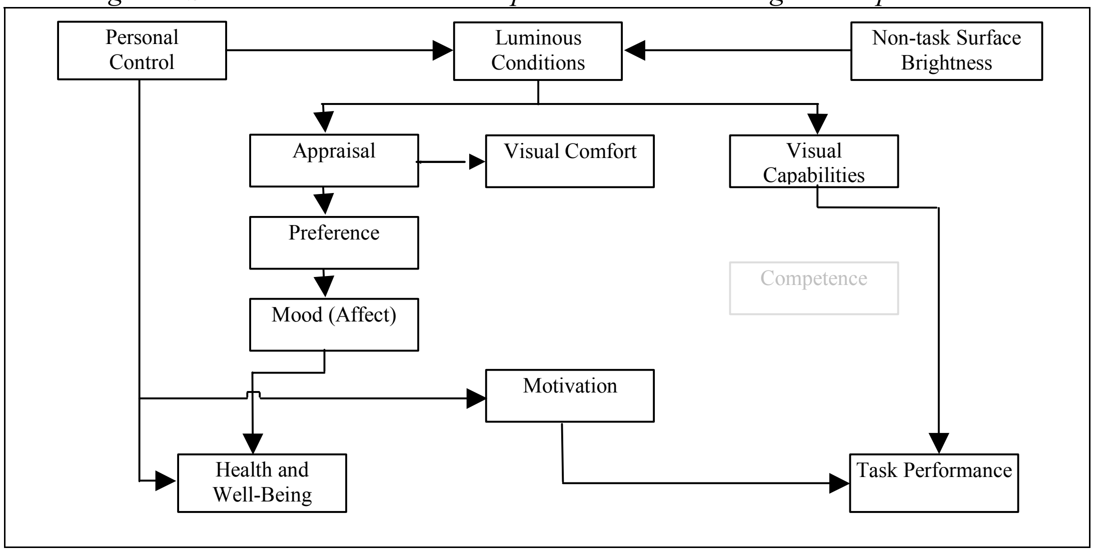
- **Lighting designs discriminated.** The lighting designs were rated for comfort as expected from normative data. The Base Case was considered comfortable by 71% of participants, and Base Case 2 by 69%, whereas the four better practice designs received more favorable ratings (81-91% comfortable). The highest ratings were obtained by the Dimming Control condition.

- **Task characteristics, practice, and fatigue important.** These effects occurred in both experiments. Performance on a vision test was best when the targets were presented at higher contrasts and larger sizes. Transcription typing performance improved when the source text, printed on paper, was in larger print (16 pt vs. 8 pt). Performance on tasks first encountered on the testing day improved from the start to the end of the day, reflecting practice. Performance on a psychophysical reaction time task declined over the day, indicating fatigue. This confirms the sensitivity of the tasks to the known influences of visibility, practice, and fatigue.

- **Current lighting practice attenuates lighting effects on performance.** Complex cognitive tasks showed no simple main effects of lighting design in either experiment. This is consistent with other short-term experiments, in which visibility was high and no condition was uncomfortable, and in which participants might be motivated to work hard for the one day regardless of conditions. There is ambiguous evidence concerning possible interactions between lighting design and partition reflectance, particularly for direct / indirect systems.

- **Control used sparingly, but effectively.** When they had control, most people used it once, near the start of the day, to choose a preferred condition. People with dimming control were particularly satisfied with their ability to achieve desired conditions. The median desktop illuminance for the people with the ability to adjust the lighting by dimming was generally lower than recommended practice, but consistent with other laboratory and field work. As mentioned above, people with dimming control showed more sustained motivation, and improved performance on a measure of attention.

Figure ES2. Linked Mechanisms Map: demonstrated linkages in Experiment 1.



Implications for Lighting Practice

- Current lighting practice (as represented by the Base Case and Base Case 2) is acceptable: levels of comfort were high, and task performance was the same as for the fixed, Best Practice lighting designs.

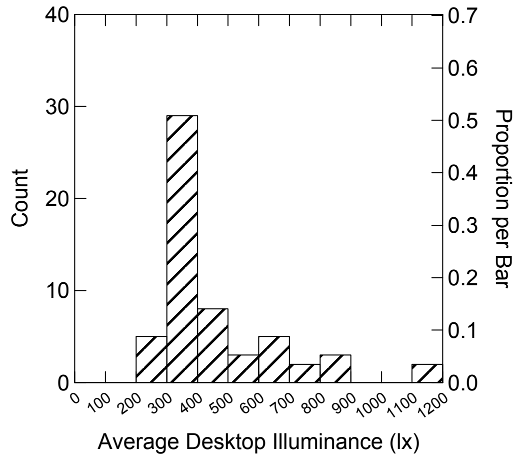
- Nevertheless, there is room for improvement. The direct / indirect installations, particularly those with individual dimming control, had even higher levels of comfort and satisfaction.

- People show a wide range of preference for illuminance (see Figure ES3). Individual overhead dimming control as well as desktop personal control are effective means of accommodating this range.

- In addition, on average, people with dimming control chose lower illuminances than current recommended practice. This implies that individual overhead dimming control has potential for energy savings.

- Although people *on average* chose lower illuminances than recommended practice, the *diversity* of preferences suggests that if a fixed lower ambient illuminance is chosen it must be supplemented with desktop personal control to allow higher local illuminances for those who prefer them. (It is important to note that the illuminances people chose were in the context of an environment with high room surface brightness due to the wall washing and an indirect component on the ceilings. Illuminance preferences might differ if the room surface brightness was reduced.)

Figure ES3. Mean desktop illuminance chosen by participants with Dimming Control. Includes data from both between-groups and within-groups designs.



Implications for Future Research

This study has implications for both future laboratory and field studies addressing the consequences of lighting design.

- Future laboratory studies should focus on specific pathways in the Linked Mechanisms Map. In particular, testing the Linked Mechanisms Map under more extreme lighting conditions that violate common design guidelines (in order to expand the range of conditions, thereby improving the sensitivity of the test). More robust measures of motivation should also be developed in order to test the hypothesized Mood ➔ Motivation ➔ Performance link.

- Field studies are important because they provide a more realistic context for research. Lighting is only part of the work context, and any laboratory experiment can only simulate part of that context. Effects involving visibility occur regardless of context, but mood and motivation effects are context-dependent. Studying these effects requires collecting data from real employees doing real jobs in real organizations over periods longer than a single day. This would also allow measures of consequence to the organization, rather than just the individual, to be explored, and would allow tests of the connections between performance and well-being.

Table of Contents

Lighting Quality and Office Work: A Field Simulation Study

1. Introduction

For many years studies of the optimum lighting conditions for offices were focused on maximizing visual performance and avoiding visual discomfort. The result was a steady growth in the understanding of what lighting conditions are necessary to achieve high levels of visual performance (Rea and Ouellette, 1991) and to avoid visual discomfort (Boyce, 2003). Further, lighting practice has taken this knowledge on board to such an extent that it is rare today to find offices lit in such a way that either visual performance is limited or visual discomfort occurs. However, this should not be taken to mean that all is well in the world of office lighting. It is all too easy to find offices that are perceived to be gloomy and uninteresting. This perception has been the impetus behind the burgeoning field of lighting quality (Veitch 1998; Veitch and Newsham, 1998a, 1998b). Lighting quality means different things to different people but the one thing all agree on is that a high quality lighting installation needs to consider not just the visibility of the tasks but the appearance of the space. This report describes a study done to determine to what extent different levels of lighting quality affect the task performance of office workers, and their health and well-being.

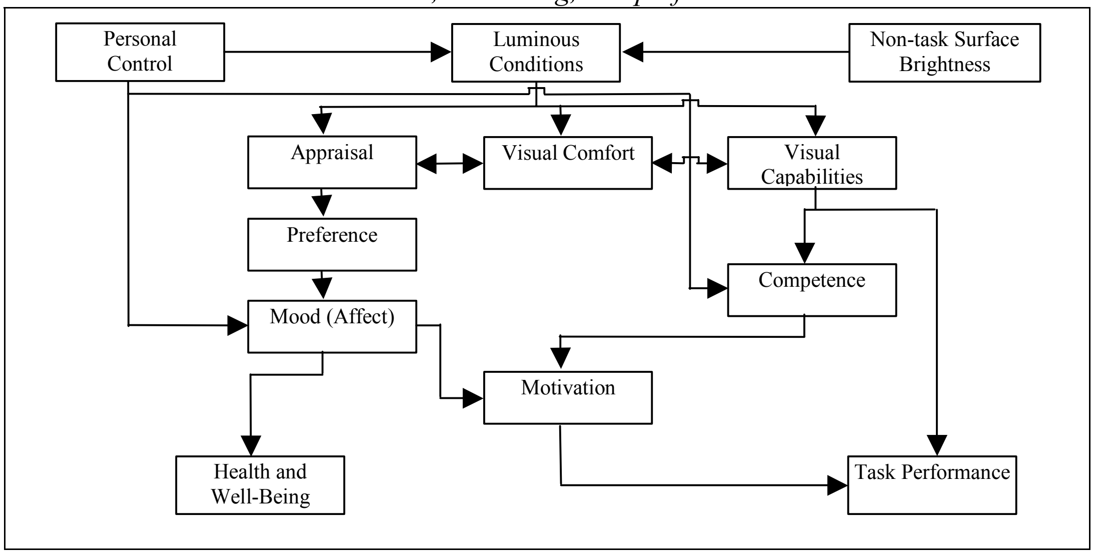
The approach adopted to address this question is that of a field simulation study. Two separate field simulation experiments have been conducted. In both experiments, "Best Practice" office lighting design conditions have been compared with a "Base Case" lighting condition representative of modern office lighting practice. The comparison is done in terms of office workers' task performance and feelings of health and well-being, over a complete working day. The "Best Practice" lighting conditions used enhanced non-task surface brightness and gave people different levels of control of the luminaires associated with their own workspace. The office workers who experienced the lighting conditions were temporary office workers, with no expert knowledge of lighting. The tasks examined were representative of those done in a modern office, but with an emphasis on knowledge-based tasks.

This approach has two major advantages. The first is that it has a high level of external validity. It uses a real office and real office workers experiencing a number of realistic lighting conditions while doing a wide range of realistic tasks. The second is that it is cost-effective, in that it provides a sensitive test of whether lighting conditions can have an effect on the task performance and/or health and well-being of office workers, even when all the task details are highly visible and visual discomfort is absent. The approach is sensitive because it tests the combined effects of several different mechanisms that may influence task performance and health and well-being and that are influenced by lighting conditions. If a significant effect of the lighting conditions is found, then the effectiveness of the individual mechanisms can be tested to determine their relative importance. If no significant effect of the lighting conditions is found, then it is unlikely that the individual mechanisms are significant - a conclusion that would have taken much longer and much more money to reach using an approach that studied one mechanism at a time.

2. Proposed Mechanisms

Wyon (1996) introduced the concept of a linked mechanisms map as a guide to well-conducted research projects that aim to provide useful practical guidance. They provide a rational basis for answering the question "Why do you expect your independent variable to affect your dependent variables?" Figure 1 shows the linked mechanisms map that sets out the mechanisms by which it is believed that surface brightness and personal control may improve task performance and enhance feelings of health and well-being.

Figure 1. Linked Mechanisms Map hypothesized to link luminous conditions with health, well-being, and performance.



The rationale behind this linked mechanisms map is as follows:

- Changing the luminous conditions in an office can immediately affect office workers in three ways; by changing visual capability (Rea & Ouellette, 1991); by changing visual comfort (Wibom & Carlsson, 1987); and, by changing the perception (appraisal) of the conditions (Flynn, Hendrick, Spencer, & Martyniuk, 1979).
- These three aspects interact over time. For example, luminous conditions that cause visual discomfort or distraction, over time, will affect visual capability and the appraisal of the conditions. Similarly, luminous conditions that limit visual capability will, over time, affect visual comfort.
- Visual capability has a direct effect on task performance (e.g., Eklund, Boyce, & Simpson, 2000, 2001). They also affect the perception of competence to do the task and hence the motivation to do the task.
- Visual comfort has a direct effect on mood and, through mood, an effect on feelings of health and well-being (although it could also be considered a form of well-being in itself).
- As for the appraisal of the lighting conditions, these are evaluated against expectations to determine whether the office worker likes or dislikes the luminous conditions (preference) and this, in turn, leads to an effect on mood (Baron, Rea, & Daniels, 1992; Newsham & Veitch, 2001).
- Mood, itself, directly affects feelings of health and well-being. Mood and the perception of competence together affect the motivation to do the task and hence task performance (Bandura, 1982; Baron et al. 1992).
- Finally, giving an office worker personal control directly affects their mood and feelings of competence to do the task, even if the control is little used (Barnes, 1981; Becker, 1986).

This study has been designed around this Linked Mechanisms Map (Figure 1). Measurements taken will allow all the elements in the map to be measured, using established methods or at least methods that have been used before.

3. Method - Experiment 1

3.1 Setting

The setting for the study was a Class C construction office building in downtown Albany, New York (Figure 2). Figure 3 shows a plan of the 239 m^2 (2,573 ft^2), air-conditioned, second floor office used. On entering the office from the elevator, the participant faced an array of nine office cubicles, arranged in three groups of three, four, and two (Figure 4). Turning to the left, the participant saw a reception desk and sofa (Figure 5), while a turn to the right revealed a long tapering hallway leading to a break area furnished with tables and chairs (Figure 6). There were a number of rooms off this open area. Starting from the elevator they were two restrooms, a kitchen, a large storage room, a janitors room, a large office, another storage room, a conference room (Figure 7) and a small room used for visual screening. Participants had access to the restrooms, the kitchen, the cubicle area, the hallway and the conference room. Of these areas, there were rectangular windows in the cubicle area, the conference room, and the end of the hallway. The positions of these windows are marked by letter in Figure 3. The dimensions of the windows, identified by letter, are given in Table 1.

The view through windows A to D, i.e., to the west, was of a small dark atrium. The view through windows E, F, G, H and I, i.e., to the north, were of a similar brick building some 10.7 m (35.1 ft) away across a pedestrian passage. The view through window J, i.e., to the east, was of some small buildings, the embankment of an interstate highway and the sky. Windows E and G were fitted with both vertical blinds and a perforated roller blind. Windows B, C, D, F, H, I, and J were fitted with vertical blinds. Both types of blinds were beige in color, the vertical blind having a reflectance of 0.56 and a transmittance of 0.21, the roller blind having a reflectance of 0.51 and a transmittance of 0.06.

By adjusting the position of the blinds on each window, the view of the sky from almost all of the cubicle area was eliminated, the amount of daylight entering the cubicle area was limited, yet a view out was preserved. For windows B, C, D, H and I, the vertical blinds were drawn across and slanted at 45 degrees. For window J, the vertical blinds were drawn back. For windows E and G the vertical and roller blinds were set so that the open part of the window started at sill height and was symmetrical about the vertical centerline of the window. For window F, the roller blind was pulled down to cut off the top part of the window. The actual dimensions and the effective dimensions of windows E, F, and G used throughout the experiment are given in Table 2.

Figure 2. The office building in downtown Albany, New York containing the study office.

Figure 3. A plan of the 239 m² (2,573 ft²), second floor office used.

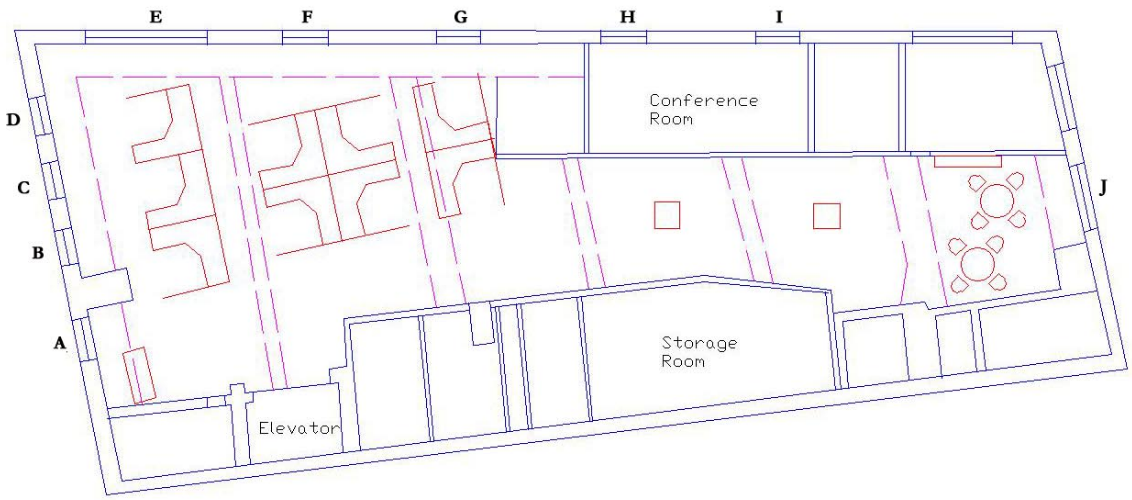


Figure 4. The array of nine office cubicles used in the study, as seen on participant arrival.

Figure 5. The reception area.

Figure 6. Hallway leading to the break area.

Figure 7. The conference room.

Table 1. Window positions and sizes

Identification letter	Location	Sill height (m)	Window width (m)	Window height (m)	Blinds
A	West wall, reception desk	0.82	1.35	2.18	None
B	West wall, cubicle area	0.82	1.35	2.18	Vertical
C	West wall, cubicle area	0.82	1.35	2.18	Vertical
D	West wall, cubicle area	0.82	1.35	2.18	Vertical
E	North wall, cubicle area	0.82	3.63	2.13	Vertical and roller
F	North wall, cubicle area	1.37	0.76	1.63	Roller
G	North wall, cubicle area	0.84	1.24	2.13	Vertical and roller
H	North wall, conference room	0.89	1.07	1.32	Vertical
I	North wall, conference room	0.89	1.07	1.32	Vertical
J	East wall, break area	0.89	2.28	2.08	Vertical

Table 2. Blind positions for windows E, F, and G.

Identification letter	Sill height (m)	Window width without blinds (m)	Window height without blinds (m)	Window width with blinds (m)	Window height with blinds (m)
E	0.82	3.63	2.13	2.89	1.04
F	1.37	0.76	1.63	0.76	0.58
G	0.84	1.24	2.13	0.91	1.55

3.2 Finishes and Furnishings

The ceilings throughout the office were suspended T bar carrying 600 mm by 1,200 mm (2 ft by 4 ft) fiber ceiling tiles with a reflectance of 0.88. In the open area, this ceiling was 2.90 m (9.5 ft) above the floor. However, the ceiling was interrupted by a series of enclosed structural beams whose underside was 2.75 m (9.0 ft) above the floor, and there was a 0.63 m (20.7 in) wide dropped soffit along the north edge of the building, whose underside was 2.44 m (8.0 ft) above the floor (see Figure 8). These features were painted flat white and had a reflectance of 0.88. In the conference room, the soffit and ceiling tiles were arranged so as to form a long elevated rectangular section above the conference table (see Figure 7). The height of the raised ceiling area was 2.75 m (9.0 ft) above floor level, while the height of the rest of the ceiling was 2.44 m (8.0 ft) above floor level.

The walls throughout the office were sheetrock with the exception of a small area near the reception desk which was painted brick. All these surfaces were painted flat white with a reflectance of 0.73.

The floor throughout the office was carpeted. The carpet was gray in color, with a reflectance of 0.20.

The cubicles were square in plan, the actual dimensions being 2.28 m by 2.28 m (7.5 ft by 7.5 ft). The height of the panels forming the cubicles were either1.67 m (66 in) or 1.37 m (54 in). Half the wall area of the cubicles had the 1.67 m height and half the 1.37 m height. Centered on the corner of the cubicle directly opposite the entrance to each cubicle was an L-shaped gray work surface. This work surface was positioned 0.75 m (29.5 in) above the floor. The work surface extended the full length of the two panels it was adjacent to (see Figure 9).

Above one arm of the L was a 0.40 m by 1.14 m (15.7 in by 44.9 in) filing bin, the bottom of which was mounted 0.48 m (18.9 in) above the work surface. A bare T5HO fluorescent lamp was recessed under the filing bin close to the panel. This lamp was mechanically filtered using electrical tape to limit its light output. The purpose of this lamp was to soften any shadows that were cast by the filing bin on the work surface. This lamp was on at all times. It produced an illuminance of 40 lx at the edge of the work surface opposite the center of the lamp, 100 lx on the work surface directly under the front of the filing bin, and 250 lx on the work surface directly under the lamp.

A 17 inch computer monitor was positioned at the angle of the L-shaped work surface, with the CPU on the floor below the work surface and the keyboard and mouse resting on a retractable shelf that could be pulled out and raised over a limited range as desired by the user. On one side of the monitor was a document holder, angled so as to place a sheet of paper in the document holder in a similar plane as the screen of the computer monitor. The location of the monitor on the desk was fixed, as was that of the document holder except that this latter could be positioned in one of two symmetric positions to the left or right of the monitor, as desired by the participant.

Completing the furnishing in each cubicle was a wheeled and padded chair of adjustable seat height, adjustable back position, and adjustable arm location.

The fabric finishes of the panels forming the cubicles were either gray with a reflectance of 0.30 or dark blue with a reflectance of 0.05 (see Figure 10). In five cubicles, all the panels had the gray finish. In two others, all the panels had the dark blue finish. In yet two others, three panels had the dark blue finish and the one containing the entrance to the cubicle had the gray finish. In all cubicles, the work surface and the filing bin were metal painted gray with a reflectance of 0.52. In all cubicles, the chair seats and backs were blue fabric with a reflectance of 0.11, while the rest of the chair was of black plastic with a reflectance of 0.04.

As for the conference room, this was furnished with a long mahogany table and a mahogany credenza, both of reflectance 0.12. Around the conference table were chairs with red seats and backs with a reflectance of 0.05, the rest of the chairs being black plastic with a reflectance of 0.03.

The only significant furnishing in the hallway area were two sets of four filing draws, arranged between the columns (see Figure 6). The filing draws were arranged back to back to form a 0.91m square in plan with a height of 0.75 m. The top of the units was covered with a mahogany panel of reflectance 0.11. The sides were painted light gray with a reflectance of 0.42.

A variety of paintings were hung on the walls of the office and various decorative items were placed on non-working surfaces to provide some visual interest (see Figures 5, 6, and 10).

Figure 8. The dropped soffit along the north edge of the building.

Figure 9. Typical work station configuration (with Lighting Design 4, Dimming Control).

Figure 10. General view of workstations showing different fabric reflectances, and decorative elements external to workstations, and Lighting Design 1 (Base Case).

Figure 11. General view of workstations, Lighting Design 2 (Best Practice).

Figure 12. Conference Room, Lighting Design 2 (Best Practice).

Figure 13. The desk lamp used in Lighting Design 3 (Best Practice with Switching Control).

Figure 14. General view of workstations, Lighting Design 4 (Dimming Control).

3.3 Lighting Manipulations

Four different lighting installations were used in this study. The four lighting installations were all designed to be representative of current office lighting practice. A check on the how well this objective was achieved was made by having a panel of experienced lighting designers visit the office and evaluate the design. The panel consisted of James Benya (Benya Lighting Design), Naomi Miller (Naomi Miller Lighting Design), and Michael White (Johnson Controls). At the time of their visit, the office was fully furnished but with different areas lit using the different lighting installations. After discussion and some minor modifications the panel agreed that the four lighting installations were all representative of current office lighting practice.

The four lighting installations were:

1. Base Case: This installation was designed to represent the most common approach to office lighting in use today. It was designed to meet the minimum requirements for office lighting as set out in the IESNA/ANSI Standard guide to office lighting (RP-1-1993) (IESNA, 1993). It consisted of a regular array of three lamp, 18 cell, 3 inch deep, semi-specular, parabolic-louvered luminaires, mounted on 8 ft by 10 ft centers in the suspended ceiling (Figure 10). Each luminaire contained three T8, 32W, fluorescent lamps of correlated color temperature (CCT) of 3,500K and a CIE General Color Rendering Index (CRI) of 82. These three fluorescent lamps were operated from an electronic dimming ballast. This installation provides an adequate illuminance on the working plane but tends to produce dark walls and ceilings. Its strongly downward directional light distribution also tends to produce shadows and veiling reflections. This lighting was installed throughout the space. Only the areas that the participants frequently used during the experiment, i.e., cubicle area, hallway and conference room, were changed to other lighting systems during the experiment. Other areas i.e., restrooms, kitchen, storage rooms, research office, and vision screening room, were lit by this installation throughout the experiment.

2. Best Practice: This installation was designed to provide a similar illuminance on the working plane as the Base Case installation but also to produce brighter walls and ceiling and to reduce shadows and veiling reflections. The part of the office containing the cubicles was lit by five runs of continuous direct / indirect luminaires suspended 16 inches below the ceiling (Figure 11). Each luminaire was either eight feet or four feet long; the eight foot luminaire containing two four foot, 54W, T5 high output fluorescent lamps, while the four foot luminaire contained only one of these lamps. All these lamps had a CCT of 3,500K and a CRI of 85. All these lamps were controlled by electronic dimming ballasts.

The north and west walls of the area containing the cubicles were illuminated by a wall-washing system using track-mounted linear luminaires, each luminaire containing one 50W 24 inch twin tube compact fluorescent lamp. These lamps have a CCT of 3,500K and a CRI of 82. The track on which these luminaries were mounted was set 3 feet in from the wall. The luminaires were positioned so as to light significant areas of the wall, i.e., areas with paintings but not windows (Figure 8). All these lamps were controlled by electronic dimming ballasts.

The hallway area was lit by three sets of four, 8 feet long indirect luminaires, arranged in a rectangle, and suspended 16 inches from the ceiling (Figure 6). Each luminaire contains two 54W, T5, high output fluorescent lamps, each with a CCT of 3,500K and a CRI of 85. All these lamps were controlled by electronic dimming ballasts.

The conference room was lit by two, 8 foot, direct / indirect luminaires, arranged above the long axis of the table. These luminaires were the same as those used in the cubicle area and contained the same lamps and ballasts (see Figures 7 and 12). The walls of the conference room were lit by five track mounted luminaires of the same type, with the same lamp and ballast as were used in the cubicle area.

3. Best Practice with Switching Control: In this installation, the lighting of the hallway and conference room was the same as in the Best Practice lighting. In the cubicle area, the ambient lighting (including wall-washing on the perimeter walls) was the same as for the Best Practice installation but, in addition, each cubicle was fitted with a free-standing desk lamp with a luminous shade placed on the unobstructed work surface (Figure 13). The desk lamp was a model. Each desk lamp contained one 40W 2D compact fluorescent lamp which could be mechanically switched to three different light outputs, equivalent to 11W, 19W and 35W. This lamp has a CCT of 3,500K and a CRI of 82. Moving and switching this desk lamp provides an element of control to the lighting of the cubicle.

4. Dimming Control: For this installation, the lighting of the hallway and conference room was the same as in the Best Practice installation. In the cubicle area, the wall-washing was also the same as in the Best Practice installation,. However, suspended above the center of each cubicle was one modified four foot, direct / indirect luminaire (Figure 14). Each of these luminaires contained three 54W T5HO fluorescent lamps with a CCT of 3,500K and a CRI of 85. The modification (made by the manufacturer for this experiment) was an opaque baffle that separated the three lamps so that one contributed only to the indirect component, while the other two could contribute to both the direct and indirect component but primarily the direct component. The light output of the lamp that contributed only to the indirect component was fixed, but the light output of the other two lamps could be dimmed from zero to full light output from the computer in the cubicle using a simple on-screen interface.

As for the area between the cubicles and the elevator, this was lit by thirteen compact fluorescent downlights (Figure 14), each luminaire containing a four-pin, 32W triple tube compact fluorescent lamp of CCT of 3,500 K and CRI of 82. In addition, two desk lamps, each fitted with one 40W, 2D compact fluorescent lamp which could be mechanically switched to three different light outputs, equivalent to 11W, 19W and 35W, were placed on the small filing cabinets near the cubicles (see Figure 14). For this application the middle, 19 W setting was used.

3.4 Lighting Monitoring

All the lighting installations, with the exception of the direct / indirect luminaires centered over each cubicle in the Dimming Control lighting installation, were operated through one lighting control system. This system allowed the luminaires forming the installation to be put into functional groups, e.g., for the Best Practice installation, the direct / indirect luminaires over the cubicle area formed one functional group and the wall washing luminaires formed another. Each functional group could be dimmed as desired. Further, the power demand for each functional group of luminaires was recorded every five minutes. This is particularly useful for monitoring the use of the individual desk lamps in the Best Practice with Switching Control installation.

The direct / indirect luminaires suspended over the center of each cubicle in the Dimming Control lighting installation were operated through a separate control network proprietary to the manufacturer. This network provided the occupant of the cubicle with a simple interface displayed on the monitor to control the light output of the two lamps in the luminaire over the cubicle that contributed to the direct and indirect components. It also recorded the time, magnitude and direction of any changes made.

3.5 Photometric Conditions

3.5.1 Illuminances

Table 3 shows the mean illuminances provided by the four lighting installations on the work surface, on the monitor screen, on the keyboard, and at the participant's eyes, in the cubicles with the gray and navy blue panels. Where there is some degree of control of light output, two average illuminances are given, reflecting the range of illuminances possible. For the Best Practice with Switching Control lighting, the range covers from the desk lamp being switched off to the desk lamp at full light output. For the Dimming Control installation, the range covers from the direct / indirect luminaire at minimum light output to maximum light output.

Table 3. Means of the illuminances (lx) provided on the work surface, on the monitor, on the keyboard, and at the participant's eye, by the four lighting installations in the cubicles with the gray and dark blue panels.

Lighting installation and panel type	Illuminance on work surface (lx)	Illuminance on monitor (lx)	Illuminance on keyboard (lx)	Illuminance at participant's eye (lx)
Base Case				
Gray	611	364	697	313
Blue	558	338	640	288
Best Practice				
Gray	601	426	649	341
Blue	528	378	544	316
Best Practice with Switching Control				
Gray	528 - 1199	371 - 447	599 - 764	295 - 422
Blue	473 - 1140	344 - 438	539 - 740	264 - 335
Dimming Control				
Gray	285 - 1032	166 - 852	269 - 1156	171 - 578
Blue	276 - 1048	188 - 873	283 - 1151	181 - 588

To indicate the contribution of daylight to these illuminances, Table 4 gives the mean illuminances on the work surfaces of the cubicles with gray and dark blue panels measured at noon on a sunny day with all the electric lighting turned off. From Table 4 it can be seen that the contribution of daylight to the measured illuminances given in Table 3 is modest.

Table 4. Means (and standard deviations) of the illuminances (lx) provided on the work surface by daylight alone.

Lighting installation and panel	Illuminance on work surface (lx)	Illuminance on monitor (lx)	Illuminance on keyboard (lx)	Illuminance at participant's eye (lx)
Daylight only Gray	19.8 (24.9)	18.8 (22.6)	22.2 (22.6)	18.2 (11.9)
Daylight only Blue	16.5 (9.4)	27.5 (24.9)	25.3 (18.1)	14.0 (1.4)

Illuminances were also measured at other salient locations in the space, i.e., the reception desk, the sofa near the reception desk, on the filing cabinets in the hallway, and on the tables at the end of the hall way that were used for meeting.
Table 5 shows the mean illuminances provided by the four lighting installations at these locations.

Table 5. Means (and standard deviations) of the illuminances (lx) provided on the reception desk, on the sofa near the reception desk, on the filing cabinets in the hallway, and on the tables at the end of the hallway, by the four lighting installations.

Lighting installation	Illuminance on reception desk (lx)	Illuminance on sofa (lx)	Illuminance on filing cabinets in hallway (lx)	Illuminance on tables at end of hallway (lx)
Base Case	390 (63.8)	492 (22.6)	445 (47.7)	392 (56.4)
Best Practice	414 (67.6)	451 (38.1)	429 (46.0)	451 (31.7)
Best Practice with Switching Control	414 (67.6)	472 (45.9)	429 (46.0)	451 (31.7)
Dimming Control	323 (51.3)	579 (94.9)	429 (46.0)	451 (31.7)

For the Best practice with Switching Control installation, the illuminance on the work surface, on the monitor screen, on the keyboard and at the participant's eyes varied as the desk lamp wattage is switched. Table 6 shows the mean illuminances measured at these locations for the three setting of the desk lamp wattage in the cubicles with the two panel types.

Table 6. Means (and standard deviations) of the illuminances (lx) provided on the work surface, on the monitor, on the keyboard, and at the participant's eyes, by the three wattage levels of the desk lamp used in the Best practice with switching control installation, for the gray and dark blue cubicles.

Desk lamp wattage (W) and panel	Illuminance on work surface (lx)	Illuminance on monitor (lx)	Illuminance on keyboard (lx)	Illuminance at participant's eye (lx)
13 W				
Gray	867 (58.1)	387 (66.5)	612 (56.0)	323 (39.9)
Blue	740 (72.7)	408 (81.8)	597 (76.8)	301 (45.9)
26 W				
Gray	1420 (179.7)	403 (38.1)	681 (23.4)	353 (17.4)
Blue	1215 (118.2)	383 (68.7)	657 (72.3)	302 (32.1)
37W				
Gray	2356 (284.0)	447 (52.4)	764 (18.3)	422 (26.0)
Blue	2248 (186.3)	438 (99.2)	740 (80.0)	335 (20.9)

For the Dimming Control installation, the illuminance on the work surface, on the monitor screen, on the keyboard and at the participant's eyes varied as the direct component of the lighting was dimmed. Table 7 shows the mean illuminances measured at these locations for three setting of the dimming control, for the cubicles with the gray and dark blue panels. One other aspect of the Dimming Control installation is the effect on adjacent cubicle when the direct component of the luminaire suspended directly above a cubicle is altered. To measure the contribution to an adjacent cubicle, the increase in illuminance on the work surface in one cubicle when the indirect component of the luminaire over the adjacent cubicle was varied from zero light output to 100% light output was measured. The mean increase in illuminance was 85 lx .

As for the conference room, the illuminance on the table in front of each chair and on the faces of people sitting in each chair were measured. The mean illuminances for the two lighting installations used in the conference room are given in Table 8.

Table 7. Means (and standard deviations) of the illuminances (lx) provided on the work surface, on the monitor, on the keyboard, and at the participant's eyes, at three dimming levels of the Dimming Control installation, for the gray and dark blue cubicles.

Dimmer Setting (%) and panel	Illuminance on work surface (lx)	Illuminance on monitor (lx)	Illuminance on keyboard (lx)	Illuminance at participant's eye (lx)
0%				
Gray	279 (35.3)	166 (16.3)	270 (34.4)	171 (21.7)
Blue	300 (27.1)	181 (32.1)	283 (12.8)	181 (32.1)
50%				
Gray	358 (48.9)	244 (22.2)	368 (39.6)	215 (43.3)
Blue	378 (23.7)	263 (40.1)	377 (22.0)	218 (12.1)
100%				
Gray	945 (88.7)	853 (64.5)	1151 (49.5)	588 (96.9)
Blue	964 (59.7)	874 (116.2)	1156 (32.0)	579 (66.8)

Table 8. Means (and standard deviations) of the illuminances (lx) provided on the conference room table, and at the eyes of participants sitting around the table, for the two lighting installations used in the conference room.

Lighting installation	Illuminance on table (lx)	Illuminance at participant's eye (lx)
Base Case	522 (37.2)	365 (52.3)
Best Practice	687 (43.5)	538 (58.7)

3.5.2 Luminances

3.5.2.1 Room surfaces

The luminances of the all the surfaces in the cubicle area, the hallway and the conference room, under each lighting installation, were measured using a Radiant Imaging 2610 digital imaging photometer, and the appearance of the space was recorded with a Sony Mavica FD92 digital camera.

3.5.2.2 Computer monitor

The computer monitors used extensively throughout the study were all set up to have a similar background color. The screen luminance of a blank white screen display, in the absence of any lighting in the room, was 120 cd/m^2.

3.6 Thermal and Acoustic Conditions

The air temperature and relative humidity conditions provided by the air-conditioning in the office were typical of offices and showed little variation over the time of the experiment The sound pressure levels occurring in the office were less than 50 dB(A). Such sound pressure levels are typical of a quiet office.

3.7 Experimental Design

3.7.1 Outline

There were two independent variables for the experiment, one being the four different lighting installations; the other being the two reflectances of the panels forming the background to the computer in the cubicle. (Although the panels also differed in color, we know of no reason why this color difference would be expected to influence any of the behavioral outcomes.)

The experimental design involved two experiments in one. One, called the between-groups experiment, collected data from every participant once only. The other, called the repeated measures experiment, exposed the same participants to two lighting conditions on separate occasions. The advantage of the between-groups approach is that the sample is larger and, therefore, is more likely to be representative of the population. It also reduces any bias due to hypothesis-guessing or to asymmetric transfer of training or mood that can occur under the repeated measures approach. The disadvantage of the between-groups approach is that more participants are required to achieve equal sensitivity to the repeated measures approach because the individual differences between participants cannot be removed from the variability. In the repeated measures approach each participant acts as their own control, so the differences between individuals can be removed from the variability, thereby increasing the sensitivity for the same number of participants.

The experimental design is shown schematically in Table 9. By using both approaches, this experimental design provided a partial replication of the experiment. Replication is important in persuading people about the reliability of the conclusions of a study. Demonstrating that lighting conditions that enhance non-task surface brightness and but allowing personal control leads to better task performance and improved feelings of health and well-being once would be interesting, but demonstrating it twice is convincing.

3.7.2 Target Sample Size

In a between-groups experiment, each participant is randomly assigned to one of the experimental conditions (in this case, one of eight conditions: four lighting conditions crossed by two cubicle panel reflectances). This procedure ensures that individual differences are equally distributed between the groups, and allows the inference that if any differences exist between the groups on any dependent measure, then the treatment is the cause of the effect (Kerlinger & Lee, 2000). A power analysis was undertaken to determine the sample size necessary to detect medium-sized effects (7-10% explained variance; Cohen, 1988), which was the smallest effect size considered to have practical significance. Assuming the standard probability criterion of alpha=.05, a sample size of 200 people would provide adequate statistical power (power= .84 for the lighting design main effect, power = .93 for the partition luminance main effect, and power = .84 for the interaction [Borenstein, Rothstein, & Cohen, 1997]). To have enough statistical power to detect a small effect (1-2%) would have required 972 participants. Thus, we set our target as 50 participants in each of the four lighting conditions.

A random selection of people in the Base Case and Dimming Control conditions who at the end of the first day had indicated their willingness to participate again, returned for a second day of testing in which they experienced the other of these two lighting conditions from the one they had previously experienced. This procedure provided a counterbalanced presentation of the conditions (half had the base-case lighting first, and half had the best-practice with Dimming Control first). They experienced the same reflectance condition (either high or low) on both days.

The repeated measures experiment sample size target was set at 50 people, 25 each from the Base Case lighting condition and the Dimming Control condition. This repeated measures comparison required a minimum sample size of 34 individuals (each experiencing both lighting design conditions) to detect medium size effects, so a sample size of 50 was expected to provide more than adequate power. (Statistical power for the detection of a small effect size would require 199 people; Borenstein et al., 1997).

Table 9. Schematic experimental design with planned sample sizes (Experiment 1)

	Base Case	Best Practice	Best Practice with Switching Control	Dimming Control
Gray Panel	Between groups N = 25	Between groups N = 25	Between groups N = 25	Between groups N = 25
Blue Panel	Between groups N = 25	Between groups N = 25	Between groups N = 25	Between groups N = 25
	N = 200 participants randomly assigned to the 8 above conditions			
Same panel reflectance for both occasions:	Repeated measures: N = 25 from Dimming Control participated a second time			Repeated measures: N = 25 from Base Case participated a second time
	N = 50, each participant participating twice in counterbalanced order.			

3.7.3 Participant Characteristics

The participants were recruited from an office temporary services agency; and paid at the agency rate for each day they worked. They were all over 18 years of age and were required to have experience with Windows-based word processing and spreadsheet software and a minimum typing speed of 30 words per minute, and have normal or corrected to normal vision and hearing.

Several demographic characteristics were recorded by the participants themselves, guided by the software on their computers. In addition, during one of the rest periods built into the day, each participant was taken to the vision screening room. In this room, the participants' near-field visual acuity, contrast sensitivity, color vision and visual field size were measured. Visual acuity was measured using the Lighthouse continuous text card seen at a distance of approximately 40 cm with an illuminance of 320 lx. Contrast sensitivity was measured using the Vistech contrast test system, viewed from a distance of 2.28 m, and at a background luminance of 60 cd/m^2. Color vision was measured using the Ishihari Test for Color Blindness. Visual field size was measured using an apparatus consisting of a horizontal array of white LEDs, with one LED positioned every 10 degrees from the fixation point. The participant was asked to fixate on the central LED, and the experimenter flashed different LEDs, one at a time, until the locations where the participant was no longer able to see the LED flashing were identified.

3.7.3.1 Between-groups experimental design

A total of 181 people provided valid data in the between-groups experimental design. The sample size for each of the eight experimental conditions is shown in Table 10. (Equipment failure and experimental errors led to the loss of data and the unequal group sizes.)

Table 10. Between-groups participants by experimental condition

	Base Case	Best Practice	Best Practice with Switching Control	Dimming Control	Total
Light (Gray)	27	25	17	33	102
Dark (Blue)	24	16	16	23	79
Total	51	41	33	56	181

The demographic characteristics of the between-groups participants are summarized in Table 11.

Table 11. Between-groups sample characteristics

Sex	69 Male			112 Female		
Age	86 – 18-29	38 – 30-39	28 – 40-49	22 – 50-59	7 – 60-69	
Corrective Lenses	54 – None	23 – Reading	55 – Distance	15 – Bi- or Tri-focal	6 – Multi-focus	7 – Contact lenses
Education	41 - HS	38 - CC	35 - UC	42 - UD	25 - GD	

Note. HS = high school diploma. CC = community college diploma. UC = undergraduate courses. UD = undergraduate degree. GD = graduate degree.

We inquired about the frequency of hearing impairment or use of a hearing aid on the testing day. Eleven participants reported having at some time being diagnosed with a hearing impairment, and three wore a hearing aid at the test. In addition, no one reported any functional problems with hearing that prevented them from communicating with others during the session.

In the separate vision testing, participants' color vision was tested: Six people failed the Ishihara color test.

The contrast sensitivity test consisted of five rows of rings, from high to low spatial frequency; in each row were nine targets of diminishing luminance contrast. Scores on this test were the sum of correctly-identified rings, from 0 to 45. The mean contrast sensitivity score was 33.22 (*SD* = 4.41). Near-field visual acuity was generally very high; the distribution is shown in Table 12.

Table 12. Between-groups participants' near-field visual acuity

	Visual Acuity 20/							
	20	25	30	35	40	50	60	80
Total frequency	59	57	33	2	18	8	3	1

We also asked participants to report the duration of their time in the paid work force and as a temporary office worker. Overall, participants had been in the workforce for an average of 15.2 years (*SD* = 11.4) and as a temporary worker for an average of 1.5 years (*SD* = 3.1).

3.7.3.2 Repeated-measures participants

A subset of the between-groups participants returned for a second visit in a different lighting condition. We examined the distributions of the demographic variables for the repeating participants to see whether or not they were similarly distributed to the larger sample. These statistics are shown in Table 13. The distributions differed from the full sample for every variable (compared to the overall sample for expected values, using chi-squared). The people who completed a second session included a higher proportion of women, a lower proportion of people aged 40 – 49, more contact lens wearers, and fewer people with only a high school diploma, than the overall sample. Two people in this group reported a hearing impairment and one used a hearing aid. The people who repeated did not differ from the whole sample in their total years in the workforce (*M* = 14, *SD* = 12) or as a temporary office worker (*M* = 1.2, *SD* = 2.2).

Table 13. Repeated measures sample characteristics

Sex	13 Male			32 Female			$\chi^2(1) = 102.6$, $p<.01$
Age	20 – 18-29	9 – 30-39	11 – 40-49	3 – 50-59	2 – 60-69		$\chi^2(4) = 103.1$, $p<.01$
Corrective Lenses	13 - None	6 – Reading	13 – Distance	3 – Bi- or Tri-focal	2 - Multi-focus	8 – Contact lenses	$\chi^2(5) = 102.4$, $p<.01$
Education	14 - HS	7 - CC	9 - UC	10 - UD	5 - GD		$\chi^2(4) = 102.8$, $p<.01$

Note. HS = high school diploma. CC = community college diploma. UC = undergraduate courses. UD = undergraduate degree. GD = graduate degree.

Visual acuity for the people who repeated showed a different pattern than for the full sample ($\chi^2(7) = 103.5$, $p<.01$) (Table 14), but the median and mean did not differ from the full sample. The difference in the distributions were a higher proportion of people with 20/20, 20/30, or 20/40 vision, and a lower proportion of people with 20/25 vision. Two people who repeated failed the color vision test, which was a significantly higher proportion than in the full sample ($\chi^2(1) = 103.7$, $p<.01$). Contrast sensitivity test scores were the same for the repeating participants as for the full between-groups sample ($M = 33.51$, $SD = 3.89$).

Table 14. Repeated measures visual acuity

	Visual Acuity 20/								Median	Mean
	20	25	30	35	40	50	60	80		
Frequency	13	20	6	1	2	2	1	0	25	27

3.8 Independent Variables

This experiment had two independent variables, lighting design and panel reflectance, that are applicable to all the dependent variables. In addition some dependent variables were measured more than once, thereby providing an independent variable of time for these dependent variables. Further, one of the tasks, the typing task, had the task material presented in three print sizes. The timed vision test also varied in target contrast, with six levels.

3.9 Dependent Variables

3.9.1 Daily Schedule

The outcome measures encompassed several domains in order to cover all of the concepts in the linked mechanisms map: visual and cognitive task performance; mood; satisfaction; vision capability; ratings of task and environmental competence, motivation, room appraisal, and lighting preferences. In addition, demographic information was recorded. The activities undertaken by the participants during the day are listed in the planned chronological order in Table 15.

As is clear from this list of activities, the first session, from 8.30 a.m. to 10.00 a.m., was primarily used for collecting demographic and current state information from the participants and for training the participants in the tasks. Subsequent sessions involved performance on a mixture of visual and cognitive tasks, with participants working as individuals or in a group. Several of the tasks and surveys were repeated so that changes over time could be examined. The typing task, summaries task, and the group discussion task required the participant to leave the cubicle and go up the hallway to retrieve information from the filing cabinets there. The group discussion task required the participant to leave the cubicle and go into the conference room with all the other participants to hold a discussion.

The nature of each activity and the data collected from it are discussed below in conceptual groups. This same conceptual grouping guided the data analysis reported in the Results section.

Table 15. Activities undertaken during the day, in chronological order

Approximate time	Duration (min)	Activity
8.30 a.m.	15	Arrival, instructions, consent, assignment to cubicle
8.45 a.m.	10	Visual capability 1
8.55 a.m.	5	Training on survey response systems
9.00 a.m.	10	Demographics survey Visual discomfort survey 1 Physical discomfort survey 1 Room appraisal survey 1
9.10 a.m.	5	Training on typing task
9.15 a.m.	15	Training on summaries task
9.30 a.m.	5	Task competence
9.35 a.m.	10	Training on the conveyor task
9.45 a.m.	15	Lighting preferences survey 1 Mood survey 1 Environmental competence survey 1
10.00 a.m.	15-25	Break and visual screening
10.15 a.m.	15	Typing task 1
10.30 a.m.	40	Summaries task 1
11.10 a.m.	10	Conveyor task 1
11.20 a.m.	30	Social situations task
12.00	30	Lunch
12.30 p.m.	15	Typing task 2
12.45 p.m.	40	Summaries task 2
1.25 p.m.	10	Conveyor task 2
1.35 p.m.	40	Group discussion task
2.15 p.m.	15	Break
2.30 p.m.	15	Typing task 3
2.45 p.m.	40	Summaries task 3
3.25 p.m.	10	Conveyor task 3
3.35 p.m.	15	Mood survey 2 Room appraisal survey 2 Lighting preferences survey 2
3.50 p.m.	10	Visual discomfort survey 2 Physical discomfort survey 2 Environmental satisfaction survey Session satisfaction survey Environmental competence survey 2 Lighting control survey (if necessary)
4.00 p.m.	10	Visual capability 2
4.10 p.m.	10	Workday experiences survey
4.20 p.m.	5	Dismiss

3.9.2 Perceptions and Feelings

3.9.2.1 Appraisal

Two scales, both presented on the computer, were used to determine the opinions of the participants about the lighting they experience: the Office Lighting Survey (Eklund and Boyce, 1996) and the lighting quality scale developed by Veitch and Newsham (2000a). Although the two sets of questions are similar, they are not identical and they use different rating methods. Normative data are available for the Office Lighting Survey, and data from a similar experiment on control and lighting quality are available from Veitch and Newsham (2000a). The lighting preferences survey was administered at the beginning and end of the day

3.9.2.2 Preferences

Also using a questionnaire presented on the monitor, evaluative judgments of the appearance of the office were collected using a set of 27 semantic differentials based on Flynn, et al. (1979) and on Loe, Mansfield, & Rowlands (1994). Each adjective pair was presented on the computer at opposite ends of a sliding scale scored from 0 to 100 (left to right). This set of scales was used by Veitch and Newsham (1998a) and shown to be sensitive to differences in lighting design. The room appraisal survey was administered at the start of the day and at the end.

3.9.2.3 Mood

This survey was also presented on the monitor. The participant's mood was assessed using the Russell and Mehrabian 3-Factor Mood Scale (Russell & Mehrabian, 1974). This is a well-documented psychological measure that has been used successfully in previous research to detect subtle emotional effects of experimental conditions (e.g., Newsham & Veitch, 2001; Veitch, Gifford, & Hine, 1991). It provides scores on a 0-8 point scale for Pleasure, Arousal, and Dominance factors, each of which is the average of the scores on 6 semantic differential pairs relating to that factor. The mood survey was administered at the beginning and end of the day.

3.9.2.4 Competence

One of the cognitive tasks (described below) involved creating summaries from longer articles. After completion of the training of the summaries task, the participant's feeling of competence was measured by simply asking how many of the summaries they expected to evaluate during the three remaining forty-minute sessions. The more competent and capable they felt, the higher should be the expected number of summaries evaluated. This was the measure of task competence.

No scale exists to measure environmental competence in an experimental context, but a new measure was been adapted from Jones and Veitch (2000) for this study. Jones and Veitch developed a 5-item scale to assess office workers' feelings of competence to create desirable environmental conditions in their usual workplaces; it is analogous to common scales measuring job competence (e.g., Wagner & Morse, 1975). Participants indicate their agreement or disagreement with such statements as "I believe I know how to create good working conditions for myself." A four-item version of the scale demonstrated acceptable internal consistency (alpha =0.73, N=93). This survey was presented on the monitor. Individuals with personal control of the lighting are expected to report higher feelings of environmental competence than those without. The environmental competence survey was administered at the beginning and end of the day.

3.9.2.5 Health and well-being

This conceptual group included measurement of satisfaction (both with the environment and with one's own performance), visual and physical discomfort, and ratings of the difficulty of the tasks.

Satisfaction with the physical environment was measured in two different ways, both administered on the monitor. The first used the four-item scale developed by Sundstrom, Town, Rice, Osborn, and Brill (1985), which has good internal consistency and reliability, and is sensitive to changes in lighting conditions (Newsham & Veitch, 2001; Veitch & Newsham, 1998a). The items ask the participant to give their level of agreement, on a five point scale, with the statements "This workspace has helped me get the work done efficiently", "All things considered, I'm very satisfied with this workspace". "The physical layout of this workspace is well suited to the tasks I have done today" and "Compared to the other offices where I have worked, I like this workplace.

The second measure was the response to the question "How much do you think the office environment you experienced today influenced your work performance compared to the office environments your are used to?" The answer was given on the nine point scale raging from -40% or more to +40% or more; this survey was completed once in Session 4.

A similar approach was used to quantify session satisfaction. The participant was asked to indicate their level of agreement with four statements presented on the monitor. The level of agreement was given on a five point scale. The statements were "I worked well today", "All things considered, I'm very satisfied with my performance at work today". "With regard to work, I was at my best today" and "Compared to other short assignments I have worked, I'm pleased with my efforts today". This survey was also completed once in Session 4.

Using a questionnaire presented on the monitor, visual discomfort was measured using a modification of the scale developed by Wibom and Carlsson (1987), and with which they demonstrated that luminance ratios of greater than 10:1 between paper and a VDT screen tend to reduce visual comfort over several hours. The scale used asks the participants to indicate the intensity (not at all, a little, moderately, very much, extremely) at which they are experiencing eight symptoms of eye discomfort (smarting, itching, gritty feeling, aches, sensitivity to light, redness, teariness and dryness). Visual discomfort was scored as the total of the intensity ratings for the eight symptoms. The visual discomfort survey was administered at the beginning and end of the day.

Using a questionnaire presented on the monitor, physical comfort was measured using a scale which asked the participants to indicate the intensity (not at all, a little, moderately, very much, extremely) at which they were experiencing nine symptoms of physical discomfort (sore throat, hoarseness, nose congested, sore back, sore wrists and arms, excessive mental fatigue, nervousness and irritability, headaches, unusually tired). Physical discomfort was scored as the total of the intensity ratings for the nine symptoms. Veitch & Newsham (1998a) have found that this measure is sensitive to changes in lighting conditions. The physical discomfort survey was administered at the beginning and end of the day.

Participants were also asked to rate the difficulty of the various tasks , using a five point scale. There were seven ratings of task difficulty: the typing task, categorizing summaries, evaluating articles and summaries, extracting summaries, questionnaires, the timed vision test, and all tasks, overall. (The performance tasks to which these refer are described below.)

3.9.2.6 Lighting control survey

For lighting installations in which the participants had some degree of control over their lighting conditions (Best Practice with Switching Control and Dimming Control), a five-item survey was presented at the end of the day. The participant was asked to indicate their level of agreement with four statements presented on the monitor. The level of agreement was given on a five-point scale. The statements were "The lighting control system allowed me to create the lighting conditions I wanted", "The lighting control system was easy to use". The instructions on how to operate the lighting controls were adequate", "The interface for the lighting controls was easy to use", and "When I used the lighting controls, the lighting changed rapidly enough in response to my commands". This survey was presented once in Session 4.

3.9.3 Performance Measures

3.9.3.1 Timed vision test

In this task a series of square gratings of side 7.6 cm (3.0 in), oriented either vertically or horizontally, or a square of the same size but uniform in luminance, were presented to the participant, on the computer monitor, in random order. For each presentation the participant had to respond whether he/she perceived a grating to be present or not, by pressing keys on the keyboard.

Two measurements were taken for each presentation, whether the answer was right or wrong and the reaction time, this being the time from when the square was first presented to when the participant responded. The uniform square and the square gratings all had the same average luminance (28.5 cd/m^2) but the square gratings varied in luminance contrast from 0 to 16 gray levels. The square gratings were presented at three different spatial frequencies, 0.6 cycles / degrees, 6 cycles / degree and 10 cycles / degree.

To ensure that the spatial frequency of the grating was not changed by the participant changing the distance from the monitor, the participant viewed the monitor from a distance of 51 cm (21.3 in), this distance being fixed by the participant placing their chin in a chin cup mounted one end of a post placed at a marked position on the floor. This task was designed to quantify the contrast sensitivity function of the participant's visual system, a metric that reflects the sensitivity of the visual system to stimuli varying in size and contrast. This task occurred twice, once at the start of the day and once at the end.

3.9.3.2 Motivation

We used the NRC Conveyor Belt task (Newsham et al., 1995a) to assess motivation. This computer-based task presents two different red or blue moving symbols that enter the screen from the left and exit on the right, in random order. This movement was over a black, horizontal strip of width 8.9 cm (3.5 in), simulating a conveyor. Pairs of symbols were chosen from a list of hearts, crosses, squares, diamonds, circles and clubs. During passage across the screen the symbols passed through a defined region on the simulated conveyor of length 10.8 cm (4.25 in). Participants were instructed to use the space bar to delete one of the two symbols presented whenever it was in this region (Veitch and Newsham, 1998a).

The speed with which the symbols moved across the screen increased in a series of 8 steps, each speed being presented for 15 seconds, giving a total trial time of 2 minutes. The outcome measure was the speed at which the symbols were moving when the participant ceased to attempt to delete them; participants were instructed to keep trying as long as possible, but to stop pressing the space bar when the task becomes impossible to keep up with. Performance on this task is taken as a measure of motivation, motivation being conceptualized as the individual's willingness to persist at a difficult or impossible task (Feather, 1962). The conveyor task was done in all four sessions; data from sessions 2-4 were used for analysis (session 1 was a training session).

3.9.3.3 Vigilance

Accuracy of performance on the conveyor belt task at lower speeds was one measure of vigilance or attention. A second measure was the speed with which the participant responded to a random prompt. This task could occur at any time during the day apart from when the participants were on a break or when the conveyor task was being done. An outline envelope icon was present on the lower right corner of the monitor at all times. Participants were informed that when the envelope changed color to a solid yellow and a beep sounded, they were required to move the mouse to click on the icon as quickly as possible. Performance on the task was measured as the time taken to respond to the changes. This was used as a measure of alertness.

3.9.3.4 Typing task

This task used the NRC Typing Task software (Newsham et al., 1995b). This task required the participant to first collect a binder of material from a marked filing cabinet in the hallway. The material to be typed consisted of a number of 200-300 word passages. The participant's task was to place the piece of paper containing one passage in the document holder and type the passage into the computer exactly as it appeared on the paper. At the end of each word typed, the computer checked the word for accuracy. If an error had occurred a beep was sounded and the error was marked by a strikethrough mark. Performance was measured as the time taken to complete the typing of the passage exactly as it appeared on the paper, the number of errors corrected, and the time intervals between finishing one passage and starting the next. There were six passages to be typed, two of each of three different print sizes (8 point, 12 point and 16 point), all at the same high luminance contrast, done in a random order that differed between workstations. Every passage was presented on a different sheet of paper. Having the passages in different print sizes introduces a difference in visibility. This provides an index of sensitivity for the experiment because the effect of print size on visual performance is well understood and can be predicted from known models of visual performance (Rea and Ouellette, 1991; Eklund et al., 2000). A maximum time of 15 minutes was allowed for this task, with a minimum of four minutes per passage. This task occurred in all four sessions; data from sessions 2-4 were used for analysis.

3.9.3.5 Cognitive performance

There were both objective and subjective cognitive performance measures, both derived from work performed on a series of news articles drawn from commercial broadcasts in 1998 and 1999.
This task required the participant to answer questions relating to a series of articles approximately 300 words long, on popular topics, and written at approximately 8th-10th grade reading levels. The articles were of matched difficulty for vocabulary and comprehension. They covered topics in politics, sports/entertainment, business and science/health.

One sequence of tasks was called article categorization. In this the participant was shown a 50-word summary of an article on the monitor. The participant was asked to read the summary and then select a category for the article from four alternatives (business, entertainment or sport, science or health, politics). Performance on article categorization is measured by the time taken to select a category and accuracy compared to the choices made by expert raters. The total time for trial on this task (including the evaluations described below) was another objective cognitive performance measurement.

Upon finishing the article categorization, a question was presented on the monitor which asked the participant to rate, on a continuous scale, their level of interest in reading the article from which the summary they had just read was taken. Regardless of their level of interest, the whole article was then presented on the monitor alongside the summary. The participant was then asked the extent to which they agreed with the statements "The summary contains all the important information from the article". "The summary is grammatically correct", and "The summary is well written". These ratings were given on a five point scale. These provided four subjective measurements of cognitive function.

A separate task was called summary extraction. In this the participant first read an article printed in 12 point, high contrast print on paper. Then the participant viewed the monitor on which were displayed all the sentences forming the article, one per line, in order. The participant was asked to mark the four sentences that are most important in conveying the meaning of the article. This summary extraction task was scored by the time taken to select the sentences.

Categorization and extraction tasks occurred in all four sessions, but only data from sessions 2-4 were analyzed. Forty minutes (20 for each part) were allotted to these tasks in each session.

The preceding tasks were primarily computer-based. We also had a source of comparable data from material presented on paper, as part of the preparations for a group discussion task (described further below, under Social Behavior). For this task the participants were asked to consider themselves as editors of a fictional magazine entitle *markeTrend*. In the first part of the task, each participant reviewed two sets of four stories, all of which were similar in length, style and content to those used in the Summaries task. The stories were printed on paper in high contrast 12-point print and reviewed by the participant in his or her cubicle. The participant's task was to rate each article on a series of 7 point scales for interest to the reader, relevance to *markeTrend*, and ease of understanding, and to estimate the probability (0 to 100%) that the article would be chosen by the group of participants as the lead story for *markeTrend*.

3.9.3.6 Work structure

For the typing, conveyor belt, and summary extraction tasks, we also recorded the time the time intervals taken by the participant between completing one part of the task and starting the next. These were analyzed to give an indication of whether or not lighting conditions altered the work strategies of participants, as had been observed in a prior experiment (Boyce & Eklund, 1996).

3.9.4 Social Behavior

3.9.4.1 Liking for group members

After the participants had made their individual ratings of articles for the fictional magazine entitled *markeTrend,* they went to the conference room for a group meeting. For each article, the participants, as a group, were asked to reach consensus on how interesting the article would be to the readers of *markeTrend,* how relevant it was to *markeTrend*, and how easy it was to understand. Finally, the group had to select one article from each set of four as lead stories in two successive publications of *markeTrend* and suggest an appropriate title for the two selected stories. The purpose of this task was to provide a constant source of interaction for all groups, in order to facilitate the creation of opinions about other group members.

At this point the participants return to their individual cubicles and are asked to complete a simple paper questionnaire on the behavior of the group. The questionnaire asked the participant to rate his/her degree of agreement with the following statements: "The discussion gave an opportunity to get to know the others in the group; I like the other members of this group; The group worked effectively to complete the task; The group was argumentative; If I could, I would choose this group to work with in the future." The degree of agreement was given on a 7 point scale, ranging from very strongly disagree to very strongly agree. This task was done once in Session 3

3.9.4.2 Conflict resolution

Baron et al. (1992) found that lighting conditions influenced the likelihood that participants would choose co-operative rather than competitive forms of conflict resolution, when asked how they would resolve conflicts described in written scenarios. For this experiment, five scenarios involving workplace conflicts were presented on paper, each with five alternative responses. For each scenario, the participant was first asked to rate the likelihood that they would adopt each alternative response, the likelihood being given on a 7-point scale. After assessing each alternative response in this way, the participant was asked, without looking back at these answers, to rank order the five alternative responses, with 1 = most likely response and 5 = least likely response. this task was done once in Session 2.

3.9.4.3 Willingness to volunteer

As part of the final survey at the end of the day, the participants were asked how much time they would be willing to spend completing questionnaires on environmental issues regarding buildings, transportation, and energy conservation, at home, on a scale ranging from zero to ten hours, and how willing they would be to spend another day like they had just completed, the response being given on a continuous scale from "not at all willing" to "very willing".

3.10 Daily Procedure

Participants arrived at the site between 8:30 and 9:00 each day. Group sizes varied from three to nine (nine being the number of workstations available). Groups were either all-male or all-female to avoid any variation in gender balance in the group discussions. In addition, sessions were either composed entirely of people who were in the between-groups experiment (i.e., on their first visit), or of people who had returned for a second day as part of the repeated-measures experimental design. This ensured that all the between-groups participants had the same starting level of knowledge about the experiment, and could not acquire extra knowledge from a repeating participant in the same group.

As they arrived, the experimenter oriented the participants to the site; then they read and signed an informed consent form. When all participants for the day had arrived, the experimenter explained that the purpose of the experiment was to study office productivity. The experimenter then described the structure of the day, the nature of the tasks that participants would perform, and the types of additional vision testing that they would undergo. The participants were informed that the custom computer software would train them on the tasks, guide them through the sequence of tasks, and record their performance. Participants could then ask questions about what was expected of them.

When everyone understood the requirements, they went to their assigned cubicles, entered their assigned identification number, and began working. Assignment to high- and low-reflectance partitions was random for the between-groups groups, subject to the need to have equal numbers experiencing both reflectance panels. Participants who came for a second day stayed with the same reflectance for both lighting conditions. For the Best practice with Switching Control lighting, when the participant entered the assigned cubicle the desk lamp was on and had been set at the middle power setting. For the Dimming Control lighting, when the participant entered the assigned cubicle the direct component of the direct /indirect luminaire suspended over the center of the cubicle was switched on and the light output had been set to the midpoint of the control interface.

The experimenter monitored the participants closely at the beginning of the day to ensure that all were using the software correctly. He or she also watched participants using the chin rest during the visual capability tests at the beginning and end of the day. Throughout the day, he or she monitored participants' progress through the tasks and checked written materials for completeness.

To encourage people to move around in the office suite so that they experienced the lighting conditions throughout the space, they had to retrieve task materials from filing cabinets in the hallway at intervals, get coffee or their lunches from the kitchen area, and, during breaks, sit in the break area at the end of the hallway.

In the morning, participants completed two sessions of task work separated by a break and the administration of four vision tests (see Table 15). At the break, they could go to the kitchen or the break area; only if they were smokers were they allowed to leave the office.

Participants took a 30-minute lunch break, as close to noon as possible. If some participants finished their tasks early, they were given a paper version of the summaries task and were told to work on the summary extraction and categorization parts of the task. This activity was a time filler and was not scored. During lunch, participants were encouraged to stay in the office break area, although those who had not brought their lunch were allowed to buy it from the restaurant in the building opposite. Some did choose to stay outside socializing, walking, or smoking, especially when the weather was pleasant.

After lunch, participants went back to their cubicles for the third session of task work. As they completed the readings for the Group Discussion task, they could take a break until all participants were ready for the task. Some went immediately to the conference room, some waited at their cubicles, and some went to the break area. When all participants were ready to begin and seated in the conference room, the experimenter joined them to explain the task, then left the room and closed the door. Approximately 30 minutes later, he or she let the group know how much time remained for the task, if they had not already finished. He or she collected the participants' completed forms as they left the conference room. They returned to their cubicles, filled out the Group Discussion Feedback form, gave it to the experimenter, and were given their last afternoon break. The fourth session of task work followed, and then the end-of-day questionnaires administered by the computer software.

Most participants finished their work by 4:15-4:30 p.m. The experimenter signed their time cards and thanked them for their work. If the participants had any questions about the experiment, the experimenter deflected these questions by telling them that they would receive an explanation of the experiment and a summary of the results when the project had been completed.

3.11 Ethical Issues

The experiment protocol was submitted for approval to the Institutional Review Board at Rensselaer Polytechnic Institute, to the Research Ethics Board at the National Research Council of Canada, and to the Ethics Review Board of the Pacific Northwest National Laboratory prior to the commencement of data collection. All eventually approved the protocol, with some minor modifications. An informed consent procedure was used in which participants received information in advance about the general nature of the study, and more detailed information at the time of their participation concerning the risks and benefits of participation, and rights accorded to research participants. At the start of the experimental session they were asked to sign a consent form.

The information provided to the participants prior to the experimental session deliberately omitted any mention of the existence of varying lighting conditions. This was done so as to minimize the likelihood that participants would have expectations about the purpose of the study. To persist with this deception, participants were told to report to "Office Solutions", the name intending to convey that the study was about all aspects of the office and not just the lighting. Further, data collection for the between participants experiment was completed before data collection for the within participants experiment started.

4. Results and Discussion - Experiment 1

4.1 Analytic Strategy

Conservative practices from the behavioral sciences guided the data analysis. This section describes the general approach taken. For detailed discussions of the issues, consult general works such as those by Ghiselli, Campbell, and Zedeck (1981), Keppel (1982), Kerlinger and Lee (2000), and Tabachnick and Fidell (2001).

4.1.1 Research Design

4.1.1.1 Statistical model.

As described above, this experiment combined two research designs, one a between-groups design and one a repeated measures design. Data from each research design were analyzed using analysis of variance models. Separate analyses were conducted for each conceptual group of measurements (groups of dependent variables). When there was more than one dependent variable relating to the concept, the analysis was a multivariate analysis of variance (MANOVA). This technique gives an overall test of significance for all dependent variables and individual, univariate significance tests. The univariate effects on individual measurements were interpreted only if the multivariate test reached statistical significance ($\alpha < .05$). This practice limits the possibility of Type I statistical errors. If the group had only one dependent variable, univariate analysis of variance was used (ANOVA).

4.1.1.2 Between-groups design.

The general between-groups experimental design was a 4 x 2 (Lighting Design X Reflectance) factorial experiment, in which participants were randomly assigned to a lighting design and partition reflectance condition. For Lighting Design, three orthogonal planned comparisons were tested: the comparison of Base Case versus Best Practice; Best Practice versus Best Practice + Switching Control; and, Best Practice versus Dimming Control. Each comparison addressed our starting hypotheses, and, as a group, the comparisons had the added benefit of being independent of one another (thereby providing a further protection against obtaining significant results merely because of conducting many interrelated statistical tests). For Reflectance there was one possible comparison, between Light and Dark partitions. We also tested all possible interactions of the three Lighting Design planned comparisons and the Reflectance comparison.

4.1.1.3 Repeated measures design.

The repeated measures experimental design was a 2-level (Lighting Design) comparison between Base Case and Dimming Control. This allowed us to test the widest difference between lighting conditions with the greatest statistical power. Participants who repeated kept the same reflectance on both occasions, making it a between-groups variable; we ignored it in the repeated measures design because it lacked statistical power.

4.1.1.4 Non-lighting variables.

For both the between-groups and repeated measures designs there were additional within-subject (i.e., repeated measures) independent variables for some of the dependent variables. Some variables were measured twice, and others were measured three times during the day. For the transcription typing task only, the source text came in three print sizes (8, 12, or 16 point) at all three times. The visual performance task could be analyzed by six contrast levels or three spatial resolutions (these were not crossed). For these variables we examined polynomial trends (linear and quadratic, where appropriate), and interactions with all of the other variables in the analysis.

4.1.1.5 Interaction effects and interpretation.

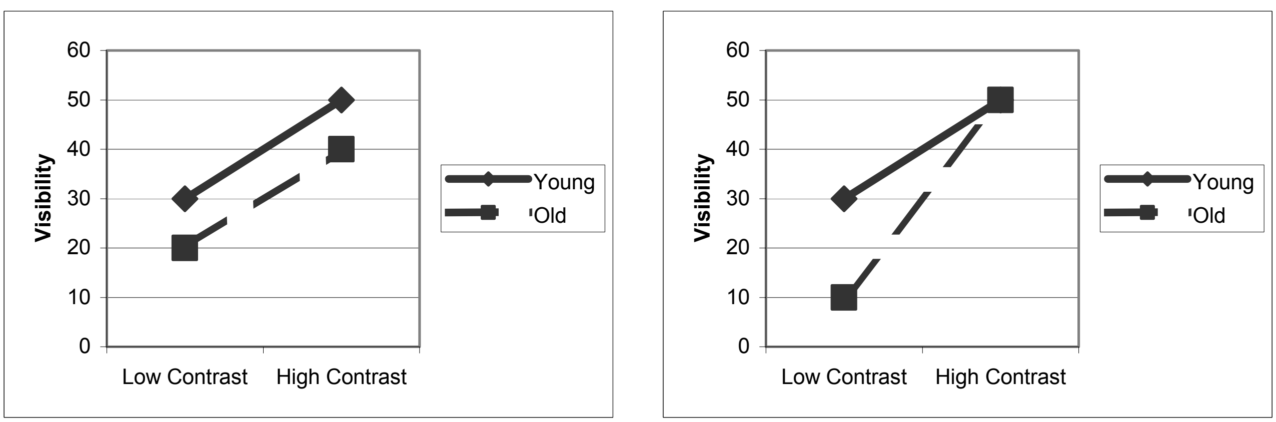
When we found statistically significant interaction effects, we further probed them using *post hoc* tests to ascertain the exact nature of the interaction. In this case, we chose to conduct planned comparisons between specific pairs of means in the interaction as if they stood alone. We did not apply any additional corrections for Type I (e.g., Bonferroni probabilities) because we judged them to be overly conservative for our purposes. The focused tests already had fewer degrees of freedom than the original tests and were therefore less powerful.

Interpretation followed the standard rules. A significant interaction means that the effect of one variable is different depending on the level of the second variable. The nature of the interaction qualifies the interpretation of the main effect. Therefore, statistically significant interactions were interpreted first, followed by statistically significant main effects only if the main effect was strong.

Consider a hypothetical example, shown in Figure 15, of the effects of age and task contrast on visibility (these data are made up for the purposes of this example). The left-hand panel shows results that might have two statistically-significant main effects. Increasing contrast appears to improve visibility by the same amount for both old and young people, and young people appear to score higher on visibility than old people, for each task contrast. The fact that the two lines are parallel tells us that there is no interaction of age and contrast. Although the average visibility is always higher for young people than for old (an average, or mean (*M*) of 40 versus 30), the difference between scores for high and low contrast is 20 units for both age groups.

The right-hand example, however, shows lines that intersect. In this example, it appears than the effect of increasing contrast is larger for older people than younger people (equally, the effect of age is larger for low contrasts than for high contrasts). If our statistical tests tell us that the interaction is statistically significant, we cannot reach conclusions about the effects of one variable without knowing the state of the other. We would conduct post hoc tests on all possible pairs of means to aid interpretation. In this example, it appears that although young people score better than old at low contrasts, the two are roughly the same when the task has high contrast. In addition, it would be generally true that visibility for high-contrast tasks is better than for low-contrast tasks. Although in this example there is the same difference between the means for young and old adults (*M* = 40 versus 30), it is probably not true in this case that young people always score better than old people. That is, there is a main effect but knowing the interaction tells us more. The important fact about the interpretation is that the effects of one variable depend on the level of the other. Old people can do as well as young ones if the task contrast is high.

Figure 15. Hypothetical example of main effects and interactions.



4.1.1.6 Statistical versus practical significance.

We reported effect sizes for all statistically significant effects. Our effect sizes are reported as percentages of variance explained (R^2 or $\eta^2_{partial}$). We interpreted them using Cohen's (1988) guidelines for small, medium, and large effects. A small effect explains 1% of the variance; a medium effect 9%, and a large effect 25%.

The importance of a given effect size depends, however, on the specific outcomes being tested. For example, a famous study tested the effect of aspirin on heart attacks. Those who took aspirin daily had fewer heart attacks than those who took a placebo pill. The study was prematurely stopped because it was concluded that it would be unethical not to give aspirin to the placebo group given the clear benefits of aspirin in the treatment group. Using the measure that we are using (% of variance explained), the effect size was R^2 = .011 (Rosenthal, 1991). This small effect is important because it means that three out of every hundred people who might otherwise have died will survive if they take an aspirin every day (the R^2 = .011 corresponds to r = .034, or three in one hundred people if expressed as a binomial). This is a common effect size in biomedical research with large sample sizes.

The outcomes in the Light Right Albany Experiment are not as fundamentally important as life and death, so one might say that a small effect is not important enough for us to interpret. However, the range of independent variables also influences effect sizes. Consider the effect of age on the heights of teenage girls (Cohen, 1988) (Table 16). If there is a small change in the predictor (age, in this case), the effect size will be small. This is analogous to the present experiment, in which horizontal illuminances were kept in a tight range and all lighting systems more than adequately lit the working area.

Table 16. Effect sizes for effect of age on height of teenage girls.

Age Range (years)	Effect Size (R^2)
15-16	.014
14-18	.058
13-18	.137

Another comparison is to a meta-analysis of the effects of illuminance on office task performance (paper-and-pencil, reading-intensive tasks, but not just involving vision) (Gifford, Hine, & Veitch, 1997). A meta-analysis is an investigation that statistically combines many separate studies to see what the overall effects are. Overall the effect of increasing illuminance on office task performance had an R^2 = .032 over the range of very low (average 70 lx) to very high (average 1962 lx) illuminances.

Group means (*M*s) and standard deviations (*SD*s) are also indicators of the practical significance of an effect, and we report these also. Most of the dependent variables do not have direct parallels to workplace measurements but are interpreted in a relative manner between experimental conditions. The larger the difference in relation to the scale of measurement, the greater the potential importance of the effect.

4.1.2 Dependent Variables

For dependent variables that used established scales, we used the established rules to form the final scores. If there were no established scales associated with a set of related questions (e.g., for the Room Appearance judgments), we used principal components analysis with varimax rotation to find an interpretable simple structure, and created subscales using unweighted averages of the answers to questions loading on each component. For performance variables that included repeated trials within a session, the dependent variables were averages within the session.

The frequency distributions of dependent variables were examined and tested for normality. Many of the dependent variables were not normally distributed, although the direction and magnitude of the skew varied widely (as one would expect for such a diverse set of measures). Transformations were used where possible (and as appropriate) to convert the distributions to normality. If no such transformation was possible or appropriate, analyses proceeded with the original data. The number of participants included varies slightly from one analysis to another because of missing data.

4.1.3 Lighting Choices

For people in the Best Practice + Switching Control and Dimming Control conditions, we also recorded their lighting choices (power levels or dimmer settings) throughout the day. These were converted to illuminance values using measured photometric data. We examined the luminous conditions that each group chose, and also conducted multiple regression analyses to determine whether or not the photometric conditions directly influenced performance or mood. These regression analyses were conducted separately for the two forms of lighting control. For the Dimming Control group, we combined data from the between-groups and repeated-measures experimental designs and controlled statistically for the order in which it was experienced. Data from the Switching and Dimming control conditions could not be combined; participants in Switching control had no opportunity to lower illuminance beyond the Best Practice design setting, whereas Dimming Control people were able to do so.

4.1.4 Linked Mechanisms Map

We had planned to use structural equation modeling (SEM) to conduct simultaneous tests of all of the links in the linked mechanisms map. However, the data were not distributed appropriately for us to use this method; the data failed to meet the statistical assumptions. Instead, we tested each link as a single, *a priori* hypothesis in which an independent variable predicted a dependent variable. We collapsed all the data across Lighting Design and Reflectance conditions for these tests. For each independent variable we formed two non-random groups by splitting the data at the median (50^{th} percentile) of that variable. We then conducted a MANOVA or ANOVA (as appropriate) to test whether or not levels of the independent variable affected the dependent variable(s) at the next level of the link. For example, we formed two groups for ratings of Lighting Quality, and tested whether being above or below the median in lighting quality rating influenced ratings on the Room Appearance scales. We conducted these analyses separately for the between-groups experimental design (people on their first visit) and for the additional data from people in the repeated measures design (people on their second visit), thereby providing a partial internal replication. This is not a perfect test of the map because the tests are not independent; however, it provides support to theoretically-derived predictions and provides guidance for further research.

4.2 Manipulation Checks

4.2.1 Demographics

4.2.1.1 Between-groups design.

The between-groups research design rests on the assignment of participants to groups being truly random, so that individual characteristics are approximately equally distributed across groups. Differences in the dependent measures can then be attributed to the between-groups treatments rather than to any inherent differences between the group compositions.

We checked this assumption against several demographic characteristics. There were no differences between the groups (4 Lighting Designs or 2 Reflectances) for sex (Lighting Design, $\chi^2(3) = 2.30, p>.05$; Reflectance, $\chi^2(1) = 0.07, p>.05$) , age (Lighting Design, $\chi^2(12) = 14.21, p>.05$; Reflectance, $\chi^2(4) = 6.03, p>.05$), education (Lighting Design, $\chi^2(12) = 20.53, p>.05$; Reflectance, $\chi^2(4) = 4.52, p>.05$), or type of corrective lenses (Lighting Design, $\chi^2(15) = 22.03, p>.05$; Reflectance, $\chi^2(5) = 4.51, p>.05$). In addition, the distributions did not differ across Lighting Design (for hearing impairment, $\chi^2(3) = 5.03, p>.05$; for hearing aid, $\chi^2(3) = 1.11, p>.05$) or Reflectance groups (for hearing impairment, $\chi^2(1) = 0.60, p>.05$; for hearing aid, $\chi^2(1) = 0.13, p>.05$). Of the six people who failed the Ishihara color test, one was in the Base Case condition (2%), one in the Best Practice condition (2%), one in the Best Practice + Switching Control condition (3%) and three in the Dimming Control condition (5%). The frequency of impaired color vision did not differ between the lighting conditions ($\chi^2(3) = 1.1, p>.05$), nor across Reflectances ($\chi^2(1) = 0.28, p>.05$). For contrast sensitivity, we conducted a 4 x 2 Lighting Design X Reflectance ANOVA on the overall scores and found no between-groups differences.

We tested near-field visual acuity, and examined its distribution across Lighting Design (Table 17) and Reflectance conditions. In this case, we did notice a difference. Visual acuity is slightly better for the Base Case participants than for others ($\chi^2(21) = 33.8, p<.05$). However, the median visual acuity for all groups is the same, 20/25, and the means differed by less than 5 when the test itself cannot distinguish more finely than 5. We further examined this issue by testing whether or not adding Visual Acuity as a covariate would alter results for the NRC Vision test or for a measure of Room Appearance, judging that these were the two dependent variables that would be most likely to be affected by differences in visual acuity. There was no difference in results with the added covariate; therefore we judged that the difference between groups on near-field acuity was unlikely to have influenced the overall results.

Table 17. Near-field visual acuity by lighting design

	Visual Acuity 20/									
	20	25	30	35	40	50	60	80	Median	Mean
Base Case	21	20	9	0	1	0	0	0	25	24.1
Best Practice	10	12	10	0	5	3	1	0	25	29.5
Best Practice + Switching Control	15	8	2	0	6	0	1	1	25	28.5
Dimming Control	13	17	12	2	6	5	1	0	25	29.7
Total	59	57	33	2	18	8	3	1		

We also asked participants to report the duration of their time in the paid work force and as a temporary office worker. These continuous variables we tested in a 4 x 2 (Lighting Design by Reflectance) MANOVA design with planned comparisons as described above. This revealed one statistically significant multivariate effect, for the planned comparison of Best Practice vs Best Practice + Switching Control (Wilks' $\Lambda = 0.94$, $F(2,168) = 5.14$, $p<.05$), which was associated with a significant univariate test for years as a temporary worker ($F(1,169) = 10.24$, $p<.01$). The Best Practice participants had spent a shorter time as temporary workers than the Switching Control participants ($M = 0.6$ ($SD = 0.8$) versus $M = 2.9$ ($SD = 5.8$) years). Eight of the 33 people in the Switching Control group, but none of the Best Practice people, had worked as a temporary office worker for more than three years. However, the overall years in the workforce did not differ between the groups. This difference seemed unlikely to be a source of bias in the results.

4.2.1.2 Repeated-measures participants.

The repeated-measures design exposes individuals to all the treatment conditions (in this case, two lighting designs, Base Case and Dimming Control), so that in effect each participant serves as his or her own control. Extraneous influences cannot bias the results; this reduces random variation in the data and provides a more powerful test of the hypothesis. However, bias may result because of the order of presentation of the treatment conditions. Forty-five participants provided complete data for two lighting conditions. Of these, 25 participated first in the Dimming Control condition, and 19 participated first in the Base Case condition. Scheduling difficulties prevented a perfect split in orders (which would have been 22 or 23 in each direction). Using 23 for the expected value for Base Case and 22 as the expected value for Dimming Control, we tested the observed distribution and found that it did not differ significantly from the expected ($\chi^2(1) = 1.1$, $p>.05$). We do not believe this order difference was a likely source of bias in the results.

4.2.2 Participant Expectancies

At the end of each testing day, participants were asked to report their beliefs about the purpose of the experiment and their feelings about what might have influenced their performance or responses. These questions we ask routinely to probe the extent to which participants had expectancies about the expected outcome of the investigation, or beliefs about how the researchers had wanted them to act. In reading these responses we looked particularly for indications that participants had identified lighting as the focus of the experiment.

A content scoring of responses to the question “What did you think was the purpose of the study?” revealed differences between the lighting conditions. For the between-groups participants, those in the Base Case, at 20%, were less likely to identify lighting as the independent variable, than those in the Best Practice (42%), Best Practice + Switching Control (36%), or Dimming Control (41%) conditions. A likely reason for this is the fact that the Base Case lighting was probably more familiar to them than any of the other conditions. Participants who experienced Base Case on a second visit (having seen Dimming Control on their first visit) were very likely to say that lighting was the focus of the investigation (46%). Of those who experienced Dimming Control on the second visit, 37% reported that lighting was the focus of the experiment.

Although most responses were not specific about the nature of the research question, a small number of participants reported having been influenced by the lighting, usually saying that they had developed headache or eyestrain, or had experienced problems with glare. Those who expressed definite opinions about office lighting (2 or 3) said that they disliked fluorescent lighting. Not one person reported a preference for fluorescent lighting over other types.

Smaller percentages of people responded that the purpose of the experiment was to study vision, eye strain, or optical problems: Sixteen percent of the Base Case made this attribution, 2% of Best Practice, 18% of Switching Control, and none in Dimming Control. Two people (4%) in Dimming Control said specifically that the purpose of the study was to determine what lighting choices people make with individual controls.

We judged that these expectancies were unlikely to have significantly biased the results for the between-groups comparison, but cannot rule out these opinions as having intensified responses for the repeated measures participants on their second visits. This underscores the importance of counterbalancing the order of exposure to the two repeated conditions.

4.3 Perceptions and Feelings

In this section we report the results of analyses of both the between-groups and repeated measures designs. The results are organized by their conceptual groups. Except for the LRC Office Lighting Survey, all analyses proceeded using the MANOVA and ANOVA strategy outlined above.

4.3.1 Lighting Appraisal

4.3.1.1 LRC's Office Lighting Survey.

Although participants completed this survey in both the morning and afternoon, we focus here on the afternoon responses. Most participants did not change their responses from the morning to the afternoon; for those who did, we view the afternoon responses as more reliable indicators of their opinions of the lighting. The agree/disagree format of the survey lends itself to cross-tabulation and nonparametric analyses. For the between-groups design, we examined the tables of responses by Lighting Design and by Reflectance. Lighting Design was the only breakdown for the repeated measures design. We tested the distributions using the chi-squared test of independence to seek differences in the pattern of responses across lighting designs and across partition reflectances. We also compared the results individually for each lighting design against normative responses based on over 800 individuals in 13 offices with new or energy-efficient upgraded lighting, all in the upper North-East (Eklund & Boyce, 1996). For these tests we used the normative values as the expected values, and tested the difference between the observed and the expected values using chi-squared tests.

Tests of contingency must be interpreted cautiously. The chi-square statistic is unreliable under certain conditions. Its power can be low when the expected frequency drops below five in many cells, and particularly when more than 20% of cells have low expected frequencies (Tabachnick & Fidell, 2001). In the between-groups design a further limitation is the discrepancy in group sizes. One case accounts for a smaller percentage difference for the Base Case and Dimming Control conditions than for the Best Practice + Switching Control condition (which had fewer participants), which means than the tests for the Best Practice + Switching Control condition will have less statistical power than for the others. In interpreting cross-tabulations such as these, one must examine the observed frequencies as well as the statistical tests in order to come to an understanding of the results.

Results are in Table 18. One important observation is that all of the lighting conditions were relatively comfortable: even the Base Case lighting, which was intended to be representative of the majority of present-day office lighting, was rated as comfortable by 71% of respondents in the between-groups design and 82% of the repeated-measures participants. The Base Case responses were not different from the normative sample. The Best Practice and Dimming Control conditions both had higher-than-expected percentages of people rating the lighting as comfortable.

Equally importantly, the responses reveal that the participants distinguished between the various lighting conditions. The distinctions are stronger for the repeated-measures participants. The most substantive difference for the between-groups data is between the Base Case and the other conditions. For both overall judgments (questions 1 and 10), the Base Case group responded in a similar way to the normative sample. The percentage agreement that the lighting is comfortable rose for all conditions with direct / indirect systems, and was highest for the Dimming Control group. Similarly, the Base Case condition was judged most similar to other workplaces (question 10) and the other conditions tended to be rated as better than lighting in other workplaces. Base Case lighting was also more likely to be rated as uncomfortably bright or for the fixtures to be too bright.

Table 18. Afternoon OLS appraisals.

1. Overall, the lighting is comfortable.			
Norm: 69% Agree		Between-Groups $X^2(3) = 7.52$	Repeated Measures $X^2(1) = 6.23$**
Base Case	Count of total	36 of 51	36 of 44
	% Agree	71	82
	$X^2(1)$	0.09	3.77
Best Practice	Count of total	34 of 40	
	% Agree	85	
	$X^2(1)$	4.29*	
Best Practice + Switching Control	Count of total	26 of 32	
	% Agree	81	
	$X^2(1)$	2.33	
Dimming Control	Count of total	49 of 54	44 of 45
	% Agree	91	98
	$X^2(1)$	12.36***	17.52***
2. The lighting is uncomfortably bright for the tasks that I perform.			
Norm: 16% Agree		Between-Groups $X^2(3) = 8.18$*	Repeated Measures $X^2(1) = 5.87$*
Base Case	Count of total	17 of 51	13 of 45
	% Agree	33	29
	$X^2(1)$	12.00***	6.09*
Best Practice	Count of total	8 of 39	
	% Agree	21	
	$X^2(1)$	0.12	
Best Practice + Switching Control	Count of total	7 of 33	
	% Agree	21	
	$X^2(1)$	0.94	
Dimming Control	Count of total	6 of 56	4 of 45
	% Agree	11	9
	$X^2(1)$	1.19	1.52

3. The lighting is uncomfortably dim for the tasks that I perform			
Norm: 14% Agree		Between-Groups $X^2(3) = 2.86$	Repeated Measures $X^2(1) = 2.05$
Base Case	Count of total	2 of 50	0 of 45
	% Agree	4	0
	$X^2(1)$	4.15*	6.92**
Best Practice	Count of total	3 of 40	
	% Agree	8	
	$X^2(1)$	1.76	
Best Practice + Switching Control	Count of total	2 of 33	
	% Agree	6	
	$X^2(1)$	2.12	
Dimming Control	Count of total	7 of 56	2 of 45
	% Agree	13	4
	$X^2(1)$	0.32	3.08
4. The lighting is poorly distributed here.			
Norm: 25% Agree		Between-Groups $X^2(3) = 0.26$	Repeated Measures $X^2(1) = 0.00$
Base Case	Count of total	8 of 50	5 of 45
	% Agree	16	11
	$X^2(1)$	1.75	4.33*
Best Practice	Count of total	7 of 40	
	% Agree	18	
	$X^2(1)$	1.20	
Best Practice + Switching Control	Count of total	6 of 33	
	% Agree	18	
	$X^2(1)$	0.66	
Dimming Control	Count of total	8 of 55	5 of 45
	% Agree	15	11
	$X^2(1)$	3.45	4.36*
5. The lighting causes deep shadows			
Norm: 15% Agree		Between-Groups $X^2(3) = 1.40$	Repeated Measures $X^2(1) = 0.19$
Base Case	Count of total	6 of 50	3 of 44
	% Agree	12	7
	$X^2(1)$	0.60	2.72
Best Practice	Count of total	4 of 40	
	% Agree	10	
	$X^2(1)$	0.78	
Best Practice + Switching Control	Count of total	5 of 33	
	% Agree	15	
	$X^2(1)$	0.00	
Dimming Control	Count of total	4 of 54	2 of 43
	% Agree	7	5
	$X^2(1)$	2.35	3.10

6. Reflections from the light fixtures hinder my work.			
Norm: 19% Agree		Between-Groups $X^2(3) = 5.70$	Repeated Measures $X^2(1) = 4.11^*$
Base Case	Count of total	15 of 51	14 of 45
	% Agree	29	31
	$X^2(1)$	3.11	3.47
Best Practice	Count of total	7 of 41	
	% Agree	17	
	$X^2(1)$	0.16	
Best Practice + Switching Control	Count of total	13 of 33	
	% Agree	39	
	$X^2(1)$	9.98***	
Dimming Control	Count of total	12 of 56	6 of 45
	% Agree	21	13
	$X^2(1)$	0.11	1.25
7. The light fixtures are too bright.			
Norm: 14% Agree		Between-Groups $X^2(3) = 6.56$	Repeated Measures $X^2(1) = 6.69^{**}$
Base Case	Count of total	19 of 50	14 of 45
	% Agree	38	31
	$X^2(1)$	23.92***	12.31***
Best Practice	Count of total	8 of 41	
	% Agree	20	
	$X^2(1)$	0.78	
Best Practice + Switching Control	Count of total	7 of 32	
	% Agree	22	
	$X^2(1)$	2.25	
Dimming Control	Count of total	10 of 54	4 of 44
	% Agree	19	9
	$X^2(1)$	0.50	0.77
8. My skin is an unnatural tone under the lighting.			
Norm: 9% Agree		Between-Groups $X^2(3) = 9.17^*$	Repeated Measures $X^2(1) = 0.00$
Base Case	Count of total	11 of 50	8 of 45
	% Agree	22	18
	$X^2(1)$	13.32***	4.39*
Best Practice	Count of total	5 of 40	
	% Agree	13	
	$X^2(1)$	0.28	
Best Practice + Switching Control	Count of total	2 of 32	
	% Agree	6	
	$X^2(1)$	0.37	
Dimming Control	Count of total	17 of 56	8 of 44
	% Agree	30	18
	$X^2(1)$	31.62***	4.40*

9. The lights flicker throughout the day.			
Norm: 4% Agree		Between-Groups $X^2(3) = 2.89$	Repeated Measures $X^2(1) = 0$
Base Case	Count of total	2 of 50	0 of 45
	% Agree	4	0
	$X^2(1)$	0.00	2.09
Best Practice	Count of total	0 of 40	
	% Agree	0	
	$X^2(1)$	2.11	
Best Practice + Switching Control	Count of total	2 of 33	
	% Agree	6	
	$X^2(1)$	1.03	
Dimming Control	Count of total	1 of 55	0 of 45
	% Agree	2	0
	$X^2(1)$	0.52	2.09

10. How does the lighting compare to similar workplaces in other buildings?.							
Norm: 19% worse – 60% same – 22% better		Between-Groups $X^2(6) = 11.63$			Repeated Measures $X^2(2) = 5.91^*$		
		Worse	Same	Better	Worse	Same	Better
Base Case	Count	4	35	12	3	28	14
	%	8	69	24	7	62	31
	$X^2(2)$		4.52			5.94	
Best Practice	Count of total	1	18	21			
	%	3	45	53			
	$X^2(2)$		22.64***				
Best Practice + Switching Control	Count of total	3	17	13			
	%	9	52	39			
	$X^2(2)$		7.09*				
Dimming Control	Count of total	4	24	28	0	22	23
	%	7	43	50	0	49	51
	$X^2(2)$		28.87***			30.70***	

Note. * $p<.05$. ** $p<.01$. *** $p<.001$.

One notable finding is the response pattern for the Best Practice + Switching Control group on question 6, "Reflections from the light fixtures hinder my work." A higher-than-expected percentage of this group agreed with the statement (39%), indicating that reflected images were somewhat of a problem for them. The percentage of agreement for the other groups were all consistent with the normative data (19%). It is not surprising that the Best Practice + Switching Control participants might have found reflected images problematic, as the luminous desk lamp had the potential to cause such reflections, whereas all of the other luminaires were designed to prevent them.

There was one unexpected finding, the observation that people in the Best Practice and Best Practice + Switching Control group rated skin tone to be as natural as the normative sample. People in both Base Case and Dimming Control had higher-than-expected percentages of people rating their skin tone as unnatural. This is difficult to interpret because the Best Practice and Dimming Control conditions both had T5HO lamps (Base Case had T8 lamps). All the lamps had a CRI over 80 and had a CCT of 3500K. We know of no property of the lamps or lighting distribution that can account for this finding, which persisted in the repeated-measures design.

Similar cross-tabulations were examined for Reflectance. There were no statistically significant differences between lighting appraisals for the two partition reflectances.

4.3.1.2 NRC's lighting quality scale.

Participants also completed this scale at the beginning and end of the day, but a software problem resulted in data loss for many participants' morning responses on this questionnaire. Therefore we analyzed only the afternoon scores. The dependent measures were Lighting Quality (LQ) and Bothersome Glare (Glare). LQ was a composite score, the average of responses on four questions concerning the level of satisfaction with the lighting. It had a theoretical range from 0 – 4, where higher values indicate higher ratings of lighting quality. It showed good internal consistency reliability, with Cronbach's alpha = 0.87 for the between-groups experimental design. Glare was the average of responses to two questions. It also had a theoretical range from 0 – 4, with higher values indicating more bothersome glare. Its internal consistence was acceptable, with Cronbach's alpha = 0.79.

For the between-groups design, these two variables were examined using the 4 x 2 MANOVA model described above. There were no statistically significant differences between ratings of either Lighting Quality or Bothersome Glare by Lighting Design, Reflectance, or their interaction. Examination of the overall descriptive statistics revealed that all the lighting designs were favorably rated. The mean Lighting Quality rating was 2.93 (*SD* = 0.89), and its median was 3, which means than half of the participants rated the lighting as 3 or 4 out of 4. Bothersome Glare, conversely, had low scores (few problems), with an overall mean of 1.13 (*SD* = 1.06), and median of 1. Although both variables had scores throughout the range of possible scores, for most participants the lighting was quite good. This reiterates the observation made above on the basis of the LRC Office Lighting Survey results.

For the repeated-measures design, the planned comparison was between Base Case and Dimming Control. For the MANOVA model, 14 cases were excluded because of missing data on Bothersome Glare. Subsequent comments revealed that some participants skipped the glare questions because they saw none. Nonetheless, even with the reduced sample size the results were strong, and clearly favored the Dimming Control condition (Table 19). This result is also consistent with the observations from the OLS. Although for both conditions the Lighting Quality was rated on the good end of the scale, the people who experienced both conditions clearly believed that the Dimming Control condition offered better lighting quality than the Base Case. As might be expected for a repeated-measures comparison, the effect size for the univariate effect is large.

Table 19. Repeated-measures lighting design effect on lighting quality

					Base Case		Dimming Control	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	0.78	4.18*	2, 29	.13				
LQ		8.65**	1, 30	.22	2.84	0.97	3.43	0.60
Glare		1.07	1, 30	.03				

Note. *p<.05. **p<.01. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.3.2 Preference

We derived four scales from the 27 semantic differential (SD) items in the Room Appearance questionnaire using the between-groups data from the afternoon. Principal components analysis with Varimax rotation was used to identify a simple structure. The most interpretable solution had four components (see Table 20). SD items loading higher than 0.500 were interpreted for each component, which were given the labels Attractiveness, Comfort, Visibility, and Spaciousness. Four items failed to load on any component, and one item (like-dislike) loaded on two components.

We used this structure to develop four scale scores, which were calculated by averaging the scores for the high-loading items, after first reverse-coding those items with negative factor loadings so that all higher scores were more favorable judgments. Possible scores range from 0 to 100. These scales were measured at two time points. Internal consistency reliability ranged from very good (Attractiveness) to just acceptable (Spaciousness), and was higher in the afternoon than in the morning.

Table 20. Rotated factor loadings from principal components analysis of afternoon room appearance judgments.

	Comfort	Attractiveness	Visibility	Spaciousness
Formal – Casual	-0.72			
Complex – Simple	-0.60			
Cluttered – Uncluttered	-0.53			
Comfortable – Uncomfortable	0.51			
Tense – Relaxing	-0.51			
Attractive – Unattractive		0.77		
Sombre – Cheerful		-0.76		
Colourful – Colourless		0.75		
Interesting – Monotonous		0.75		
Subdued – Stimulating		-0.70		
Beautiful – Ugly		0.70		
Like – Dislike		0.64	(0.50)	
Pleasant – Unpleasant		0.54		
Nonuniform – uniform		0.52		
Constant – Flickering			0.77	
Clear – Hazy			0.70	
Faces obscure – Faces clear			-0.69	
Distinct – Vague			0.64	
Gloomy – Radiant			-0.59	
Bright – Dim			0.50	
Spacious – Cramped				0.67
Overhead – Peripheral				0.51
Small – Large				-0.70
Dramatic – Diffuse				
Cool – Warm				
Public – Private				
Not glaring – Glaring				
% Variance Explained	10.07	19.80	15.53	8.74
Cronbach's alpha (PM)	.70	.90	.82	.58
Cronbach's alpha (AM)	.62	.86	.73	.45

Note. This analysis used only data from the between-groups experimental design. Only loadings above the 0.50 criterion are shown. Like-Dislike, which loaded at or above the criterion on two components, was used only in the Attractiveness scale, where it had the higher loading.

The between-groups design was examined using a 4 x 2 x 2 (Lighting Design X Reflectance X Time) MANOVA with the four scales as dependent variables and using the previously-described planned comparisons. There were no statistically significant effects of Lighting Design, Reflectance, or their interactions. The only statistically significant effect was the main effect of Time (Table 21). Both Comfort and Visibility scores dropped from the morning to the afternoon. The change in Comfort scores has a medium effect size, although the change in mean values is only 4 units. The Visibility change is small. The direction of both effects is consistent with the development of fatigue over the working day.

Table 21. Between-groups Time main effect on preference

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	AM *M*	AM *SD*	PM *M*	PM *SD*
Multivariate	0.83	8.86**	4, 170	.05				
Attractiveness		0.07	1, 173	.00				
Comfort		29.30**	1, 173	.14	70.19	12.99	66.05	14.37
Visibility		5.00*	1, 173	.03	64.76	13.83	62.95	15.16
Spaciousness		0.92	1, 173	.01				

Note. *p<.05. **p<.01. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

For the repeated-measures design the statistical model was a 2 x 2 (Lighting Design x Time) MANOVA. There were no statistically significant effects of either Lighting Design or Time, or their interaction, on the four preference scales. We examined the overall scores on these variables for the repeated-measures participants on their first visit in comparison to the other between-groups participants. It appears that the people who returned for a second session tended towards the upper end of the overall between-groups distribution. They had slightly higher scores, and experienced less of a decline from morning to afternoon, on comfort and visibility, than those people who did not return for a second visit (regardless of lighting condition) (Table 22). The data for these people has a negative skew, with more high scores than low. If repeating participants tend to be extreme scorers this will limit the power of statistical tests, even for a repeated measures design.

Table 22. Descriptive statistics for repeaters vs non-repeaters.

	Comfort - AM	Comfort – PM	Visibility – AM	Visibility - PM
Non-repeaters				
N of cases	136	136	136	136
Median	68.50	63.60	62.42	59.75
M	69.99	64.87	64.42	61.83
SD	13.18	14.69	13.97	14.89
Repeaters				
N of cases	45	45	45	45
Median	71.60	68.40	60.67	63.50
M	70.76	69.60	65.81	66.33
SD	12.53	12.88	13.50	15.63

4.3.3 Mood

There were three measures of mood, all derived from the Russell and Mehrabian Three-Factor Mood Scale (Russell & Mehrabian, 1974). Of the eighteen semantic differential items, six each are averaged to form scores for Pleasure, Arousal, and Dominance. These were collected in the morning and the afternoon. Each scale had a theoretical range from 0 – 8, and higher scores indicated stronger feelings (i.e., more pleasure, more arousal, or more dominance). Internal consistency reliability, as indicated by Cronbach's alpha, was typical of previously-obtained values (for Pleasure, α = .89 AM, α = .92 PM; for Arousal, α = .60 AM, α = .74 PM; for Dominance, α = .62 AM, α = .62 PM).

These were analyzed in the same manner as the Preference scales described above. The between-groups design showed no statistically significant differences associated with Lighting Design, Reflectance, or their interaction. Time had a statistically significant main effect, described in Table 23. The significant multivariate effect was associated with significant univariate effects for Pleasure and Arousal. Participants felt less pleasure (a medium-large effect) and less aroused in the afternoon than in the morning (a small effect). This is consistent with the development of fatigue and boredom over the working day. Several participants remarked in their open-ended comments that they found the day repetitive and tiresome, which further supports this interpretation.

Table 23. Between-groups time main effect on mood.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	AM *M*	AM *SD*	PM *M*	PM *SD*
Multivariate	0.80	13.27***	3, 160	.07				
Pleasure		35.06***	1, 162	.18	4.98	1.44	4.32	1.66
Arousal		5.03*	1, 162	.03	3.65	1.03	3.42	1.27
Dominance		0.00	1, 162					

Note. *p<.05. ***p<.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

The repeated measures analysis showed no effects involving lighting design, but did show a main effect of time (Table 24). This multivariate effect was associated with a large, statistically significant effect on Pleasure. Pleasure dropped from morning to afternoon, as it had in the between-groups design. Note that the means are higher for this analysis that they are for the between-groups design, which is consistent with the findings for the mean values of the preference (room appearance) scores reported in the previous section.

Table 24. Repeated-measures time main effect on mood.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	AM *M*	AM *SD*	PM *M*	PM *SD*
Multivariate	0.61	7.75	3, 36	.12				
Pleasure		19.48***	1, 38	.34	5.17	1.39	4.57	1.74
Arousal		0.64	1, 38	.02				
Dominance		0.00	1, 38	.00				

Note. *p<.05. ***p<.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.3.4 Competence

4.3.4.1 Task competence.

Participants reported in the morning a goal for the number of summaries they would process in a session; this was the only measurement of task competence. It was analyzed for the between-groups design in a 4 x 2 (Lighting Design X Reflectance) factorial ANOVA, and for the repeated measures design in a one-way Lighting Design ANOVA. There were no statistically significant differences in any of these analyses.

4.3.4.2 Environmental competence.

Environmental competence was measured using a previously-developed scale (Jones & Veitch, 2000). Scores on this scale have a range from 0 to 8, with higher scores indicating feelings of greater environmental competence (the ability to successfully alter environmental conditions). There are morning and afternoon scores for this dependent variable. Internal consistency reliability for this four-item scale was good (Cronbach's α = .71 in the AM and α = .85 in the PM). The values for both the morning and afternoon scores on this variable were slightly negatively skewed; the median scores for both times were 6 of 8, indicating that most participants felt somewhat environmentally competent. There were very few scores below 4, the scale midpoint.

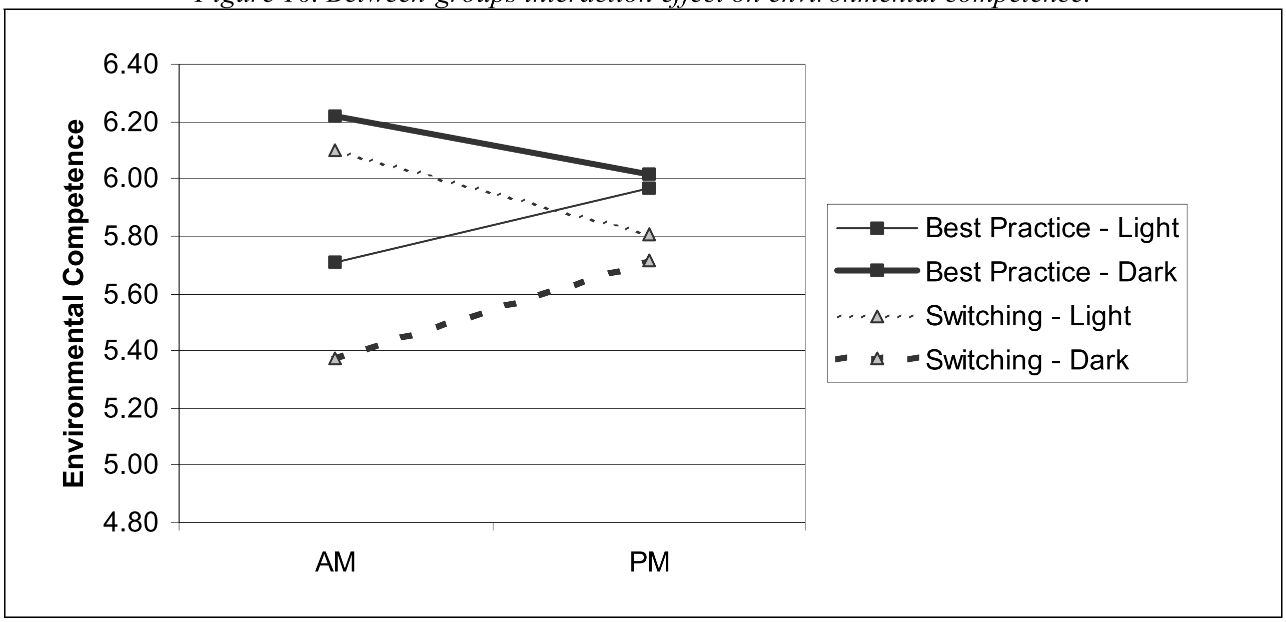
The between-groups analysis was a 4 x 2 x 2 Lighting Design X Reflectance X Time factorial ANOVA, with planned comparisons for Lighting Design. There was one statistically-significant effect in this analysis, a three-way interaction of Lighting Design (Best Practice vs Switching Control) X Reflectance X Time. This effect is detailed in Table 25 and shown graphically in Figure 16.

Table 25. Between-groups interaction effect on environmental competence.

	F	*df*	$\eta^2_{partial}$	
	5.28*	1, 173	.03	
	AM		PM	
	M	*SD*	*M*	*SD*
Best Practice – Light	5.71	1.05	5.96	1.25
Best Practice – Dark	6.22	0.81	6.02	1.43
BP + Switching Control – Light	6.10	0.93	5.81	0.84
BP + Switching Control - Dark	5.38	1.06	5.72	1.06

Note. *p<.05.

Figure 16. Between-groups interaction effect on environmental competence.



This interaction was expected and, like all three-way interactions, difficult to interpret. We conducted *post-hoc* tests on this interaction. These revealed that the only statistically significant differences between the eight points in this interaction are at the extremes: The planned comparison between Best Practice – Dark in the AM and Switching – Dark in the AM was statistically significant ($F(1, 30) = 6.36$, $p<.05$), as was the planned comparison between Switching-Light and Switching-Dark in the AM ($F(1, 31) = 4.40$, $p<.05$). The effect appears to be a statistical artifact known as regression to the mean, in which extreme Time 1 scores converge on the mean at Time 2 because of random influences. Therefore, we made no further attempt to interpret it.

The 2 X 2 (Lighting Design X Time) ANOVA for the repeated measures design showed no statistically significant effects.

4.3.5 Health & Well-Being

4.3.5.1 Satisfaction.

There were three indicators of satisfaction measured at the end of the session. Performance Satisfaction concerned the individual's sense of having worked effectively. Environmental satisfaction concerned their feelings that the workplace had been suitable for their work. Both were averages of four questions, scored on scales from 0 –4, with higher scores indicating greater satisfaction. For Environmental Satisfaction, internal consistency reliability was very good, with Cronbach's alpha = .89.

Overall, most participants were satisfied with the environment: Average Environmental Satisfaction was 2.76 (*SD* = 0.79), and the median was 3. For Performance Satisfaction, alpha was .90, and the overall mean was 2.75 (*SD* = 0.84).

Self-rated productivity was a single-item measure on a scale from –4 through +4. Negative scores indicate that the participant felt that the environment had reduced his or her productivity relative to other workplaces; positive scores indicate that the participant felt the environment had improved productivity relative to other workplaces. The overall mean was 0.67 (*SD* = 1.94), indicating than on average the participants believed that the Office Solutions environment improved their ability to work relative to most places where they usually work.

There were no statistically significant multivariate effects on these measures in the 4 x 2 MANOVA for the between-groups design. For the repeated-measures design (a one-way MANOVA comparing Base Case and Dimming Control), the multivariate test was statistically significant. It was associated with significant univariate effects on environmental satisfaction and self-rated productivity (Table 26). These effects are medium-sized ($\eta^2_{partial} > .10$), and in both cases show greater environmental satisfaction and self-rated productivity for the Dimming Control condition over the Base Case.

Table 26. Repeated-measures lighting design effect on satisfaction.

					Base Case		Dimming Control	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	0.66	6.75***	3, 39	.11				
Env. Sat.		7.27**	1, 41	.15	2.83	0.76	3.17	0.59
Perf. Sat.		0.58	1, 41	.01				
Self-Prod.		8.68**	1, 41	.17	0.69	1.96	1.64	1.78

Note. ***p*<.01. ****p*<.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.3.5.2 Comfort.

There were two measures of comfort, one specifically visual comfort and the other more general physical comfort. Each was the sum of severity ratings (from 0 = none, to 4 = extremely severe) on each of a list of related symptoms. Thus, high scores are more uncomfortable. For visual comfort the maximum possible score was 32, and for physical comfort it was 36. Participants completed these scales in both the morning and afternoon. Internal consistency reliability was good to very good for visual comfort (Cronbach's α = .73 AM, α = .83 PM) and acceptable to good for physical comfort (Cronbach's α = .58 AM, α = .78 PM).

The between-groups analysis used a 4 x 2 x 2 (Lighting Design X Reflectance X Time) MANOVA model with the above-described planned comparisons. There were no effects involving lighting design or reflectance; the only statistically significant effect involved time (Table 27). Both visual and physical comfort got worse over the course of the day, which is consistent with other research (Amick et al., 2002). These are large effects in terms of explained variance, although the scores always remained low in relation to the possible scale maxima, indicating that in absolute terms people were never very uncomfortable.

Table 27. Between-groups time main effect on comfort.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	AM *M*	AM *SD*	PM *M*	PM *SD*
Multivariate	0.70	35.97***	2, 169	.26				
Visual Comfort		62.81***	1, 170	.27	2.34	2.97	4.87	5.02
Physical Comfort		54.09***	1, 170	.24	2.46	2.67	4.65	4.55

Note. ***$p<.001$. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

The repeated-measures design was a 2 x 2 (Lighting Design X Time) MANOVA. It also showed no effect of lighting design, but did show a main effect of time on both physical and visual comfort (Table 28). Here too, comfort got worse (higher symptom severity scores) from morning to afternoon.

Table 28. Repeated measures time main effect on comfort.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	AM *M*	AM *SD*	PM *M*	PM *SD*
Multivariate	.79	5.54**	2, 42	.18				
Visual Comfort		8.06**	1, 43	.16	2.07	2.97	3.20	4.46
Physical Comfort		10.87***	1, 43	.20	2.45	2.68	3.86	4.79

Note. **$p<.01$. ***$p<.001$. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.3.5.3 Task difficulty ratings.

At the end of the day, participants rated the difficulty of six individual tasks and all the tasks, overall, on a scale from 0 – 4 with higher scores indicating greater difficulty. These seven ratings were not combined into one scale, but kept as separate indicators and analyzed in MANOVA models for the between-group and repeated measures designs. The between-groups MANOVA was a 4 x 2 Lighting Design X Reflectance MANOVA, and the repeated measures design had only the comparison between the two lighting conditions. There were no statistically significant effects on task difficulty rating in any of these analyses.

4.3.6 Summary: Perceptions and Feelings

The results for this set of dependent variables may be summarized thus:

- **Appraisals of the lighting differentiated between the lighting designs.** All the lighting conditions were considered comfortable by the majority, and all were rated fairly highly on the NRC Lighting Quality scale, but nonetheless people were able to discriminate between them when asked to evaluate specific dimensions using the LRC Office Lighting Survey. On this measure, the four lighting conditions were rated as we would have expected. The ratings for the Base Case were similar to those for the normative sample, whereas the other three conditions received more favorable ratings.

- **The experiment was sensitive to known effects of fatigue.** The lower afternoon scores for Pleasure and Arousal are typical of previous research (e.g., Veitch & Newsham, 1998a; Veitch & Newsham, 2000a). The decrease in visual and physical comfort from morning to afternoon is also consistent with the literature (Wibom & Carlsson, 1987). The change in ratings of room appearance over the day follows a similar pattern.

- **Dimming control can be very satisfactory.** The repeated-measures comparison between the Base Case and Dimming Control showed a strong pattern favoring Dimming Control. These participants rated the lighting quality as superior and also showed greater environmental satisfaction and higher self-rated productivity when they had spent the day with Dimming Control, than when the lighting was in the Base Case configuration. They also were more likely to rate the lighting as better than in other workplaces, when asked directly using the LRC Office Lighting Survey.

Except for the findings related to the repeated measures design, there were no dramatic effects of the luminous environment on perceptions and feelings. However, in this they are consistent with other laboratory-based lighting research that considered similar lighting designs and luminous conditions. For example, Veitch and Newsham (1998a) did not find statistically significant benefits of direct / indirect systems (in comparison to direct systems) on mood or room appearance judgments in a one-day exposure to an open-plan space. Neither did Eklund, et al. (2000) find differences in room appearance judgments during a one-day exposure to an achromatic private office with either a deep-cell parabolic luminaire or a suspended direct / indirect luminaire. Preferences for direct / indirect systems have been found in longer, field studies (e.g., Hedge, Sims, and Becker, 1995).

We further attempted to vary the luminous environment by varying the partition reflectance, thereby making some of the near surfaces dark. The hypothesis was that darker workstations would be rated as less satisfactory than the lighter ones, based on findings such as those of Loe, et al. (1994) on the benefits of lighter vertical surfaces. There were no effects of partition reflectance on any of the measures of preference (room appearance) or lighting appraisal. It is possible that the darker interior surface of the modular workstation was insufficient to engage this effect, because all participants could see over the workstation to more distant walls to a limited extent, and saw them when walking around the space (e.g., to retrieve material from the filing cabinets). The outer walls of the office were all quite light and in three of the lighting designs there was supplemental wall washing. Moreover, there is some evidence that these more distant surfaces are where the higher luminances are desired. In a simulation study conducted after the design of the Albany experiment, Newsham, Marchand, and Veitch (in press) found that preferred luminances for workstations showed lighter surfaces at the periphery of the space – distant walls and ceilings – than for the near surfaces of the workstation.

4.4 Performance

In this section we report the results of analyses of both the between-groups and repeated measures designs. The results are organized by their conceptual groups, which relate to constructs in the linked mechanisms map. All analyses proceeded using the MANOVA and ANOVA strategy outlined above.

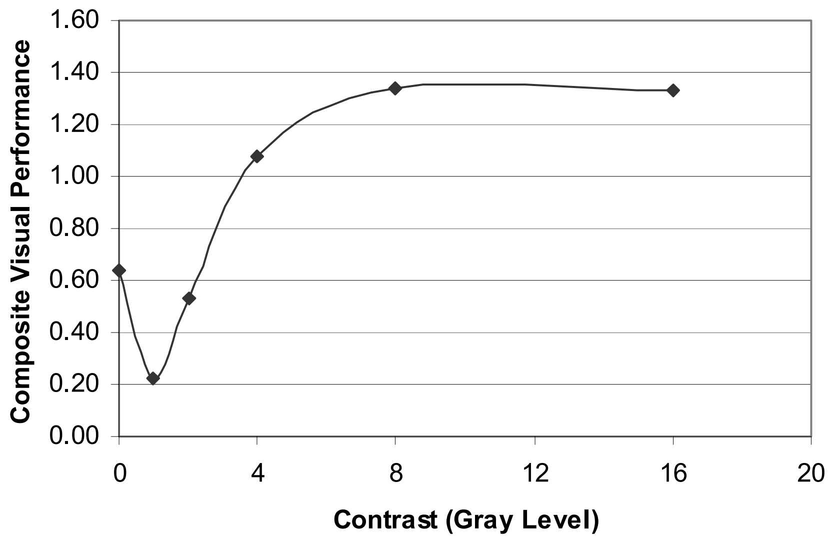
4.4.1 Timed Vision Test

Twice during the day, participants completed a computer-based contrast sensitivity test that was scored for both speed and accuracy. Response times and accuracy scores were combined to form a composite visual performance score: total correct/total time (summed across three target sizes) for all trials in each of six contrast levels. Thus, time and contrast were two additional within-subjects variables for the vision test analyses.

4.4.1.1 Overall contrast sensitivity.

The overall pattern of visual performance by contrast had the expected shape (Figure 17). Performance was lowest when the target was just barely visible (a contrast of one gray level), and improved exponentially with diminishing returns for increased contrast, leveling off at a contrast of eight gray levels, and slightly for lower contrast (0 gray levels) when the uncertainty was lower. We performed a MANOVA on the overall data, collapsed across lighting conditions and times (Table 29), and determined that the effect of contrast on composite visual performance was statistically significant, and furthermore that all of the polynomial components were themselves statistically significant. These effects are large, as one would expect based on knowledge of visual processes (e.g., Boyce, 2003) and as one would expect for a repeated-measures design (all participants did all contrasts).

Figure 17. Composite visual performance by target contrast level.



Note. This figure uses only the data from the between-groups experimental design.

Table 29. Between-groups contrast effect on composite visual performance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Multivariate	.07	450.98***	5, 168	.70
Linear trend		1895.09***	1, 172	.92
Quadratic trend		487.84***	1, 172	.74
Cubic trend		105.95***	1, 172	.38
Quartic trend		683.23***	1, 172	.80
Quintic trend		311.00***	1, 172	.64

Note. ***p<.001. Lighting condition, reflectance, and time were controlled for in this analysis. Contrast is weighted by gray level value. Only between-groups data were used here.

We simplified the remainder of the analyses on the visual performance data by focusing on the linear trend only, as it was the largest and the most interpretable. We further simplified the analyses by treating the gray-level values as being equally spaced. The repeated measures analysis was a 2 x 2 x 6 (Lighting Design X Time X Contrast) ANOVA. There were no effects in this analysis involving lighting design, but there were three non-lighting effects, which are discussed in a section below. The between-groups analysis was a 4 x 2 x 2 x 6 (Lighting Design X Reflectance X Time X Contrast) ANOVA. The lighting planned comparisons were as described above. For contrast, we examined only the linear effect of contrast on composite visual performance because it was of greatest interest. There were several statistically significant interactions and some non-lighting main effects. The interactions will be presented first, organized around the planned comparisons for lighting design.

4.4.1.2 Base Case versus Best Practice.

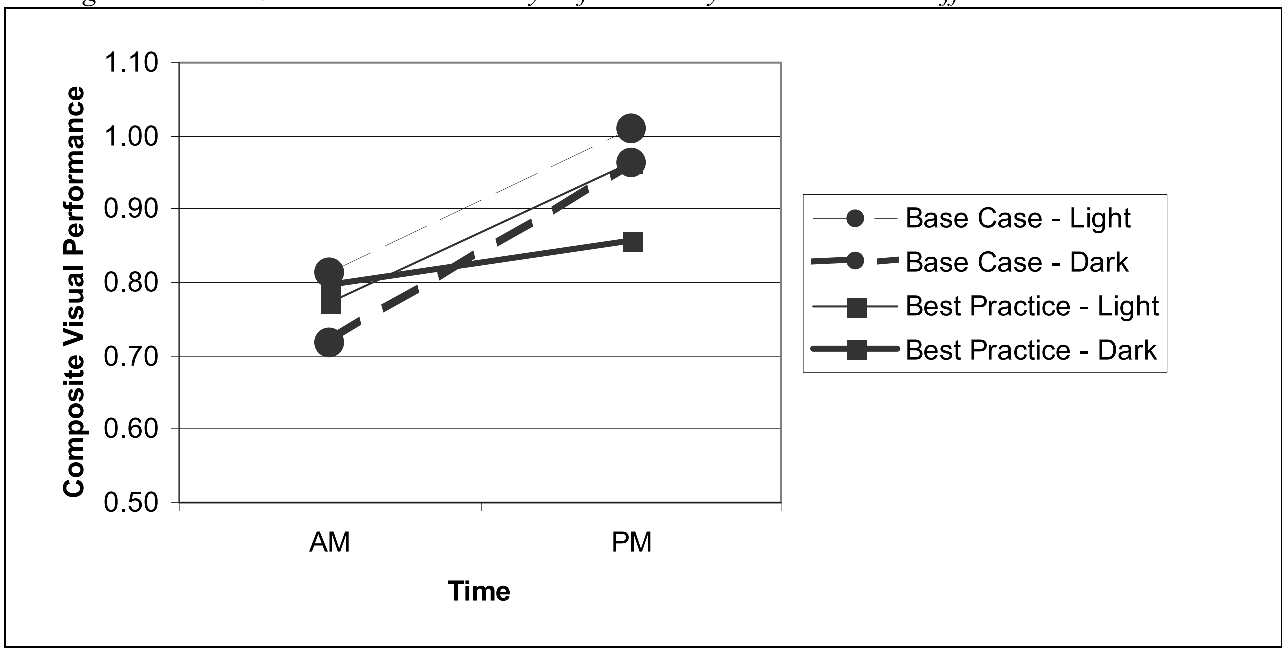
There was a small, statistically significant interaction effect of Base Case vs Best Practice by Reflectance by Time (Table 30 and Figure 18). We probed this interaction effect with *post hoc* tests and found that for the Base Case – Light, Base Case – Dark, and Best Practice – Light, the linear trends from AM to PM are statistically significant: performance improves with time. For Best Practice – Dark, there is no difference between AM and PM. Thus, it appears that if the ambient lighting system is direct / indirect, there is a penalty for dark workstations, in that there was no improvement in composite visual performance from morning to afternoon. We also compared all possible pairs of means, and found that in the morning there were no between-groups differences; whereas, in the afternoon there was a statistically significant difference between the Base Case – Light and Best Practice – Dark means (*F* (1, 41 = 4.09, *p*=.05).

Table 30. Base Case vs Best Practice by Reflectance by Time interaction effect on timed vision test.

	F	*df*	$\eta^2_{partial}$		
	4.06*	1, 172	.02		
	AM		PM		AM-PM Trend
	M	*SD*	*M*	*SD*	
Base Case – Light	0.81	0.48	1.01	0.57	*F* (1, 26) = 19.66***, $\eta^2_{partial}$ = .43
Base Case – Dark	0.72	0.47	0.96	0.59	*F* (1, 23) = 52.24***, $\eta^2_{partial}$ = .67
Best Practice – Light	0.77	0.55	0.96	0.62	*F* (1, 23) = 27.55***, $\eta^2_{partial}$ = .54
Best Practice – Dark	0.80	0.57	0.86	0.60	*F* (1, 15) = 2.20

Note. *p<.05. *** p<.001.

Figure 18. Base Case vs Best Practice by Reflectance by Time interaction effect on timed vision test.



This three-way interaction effect modifies a two-way interaction of Base Case vs Best Practice X Time (Table 31). Composite visual performance increased in the afternoon for both groups, but the increase is greater for the Base Case. However, we know from the significant three-way interaction with reflectance that the real difference is for the Best Practice – Dark group, rather than for the Best Practice lighting design as a whole.

Table 31. Base Case vs Best Practice by Time interaction effect on timed vision test.

	F	*df*	$\eta^2_{partial}$		
	4.62*	1, 172	.03		

	AM		PM		AM-PM Trend
	M	*SD*	*M*	*SD*	
Base Case	0.77	0.47	0.99	0.58	*F* (1, 50) = 57.70***, $\eta^2_{partial}$ = .54
Best Practice	0.78	0.56	0.92	0.61	*F* (1, 39) = 23.27 ***, $\eta^2_{partial}$ = .37

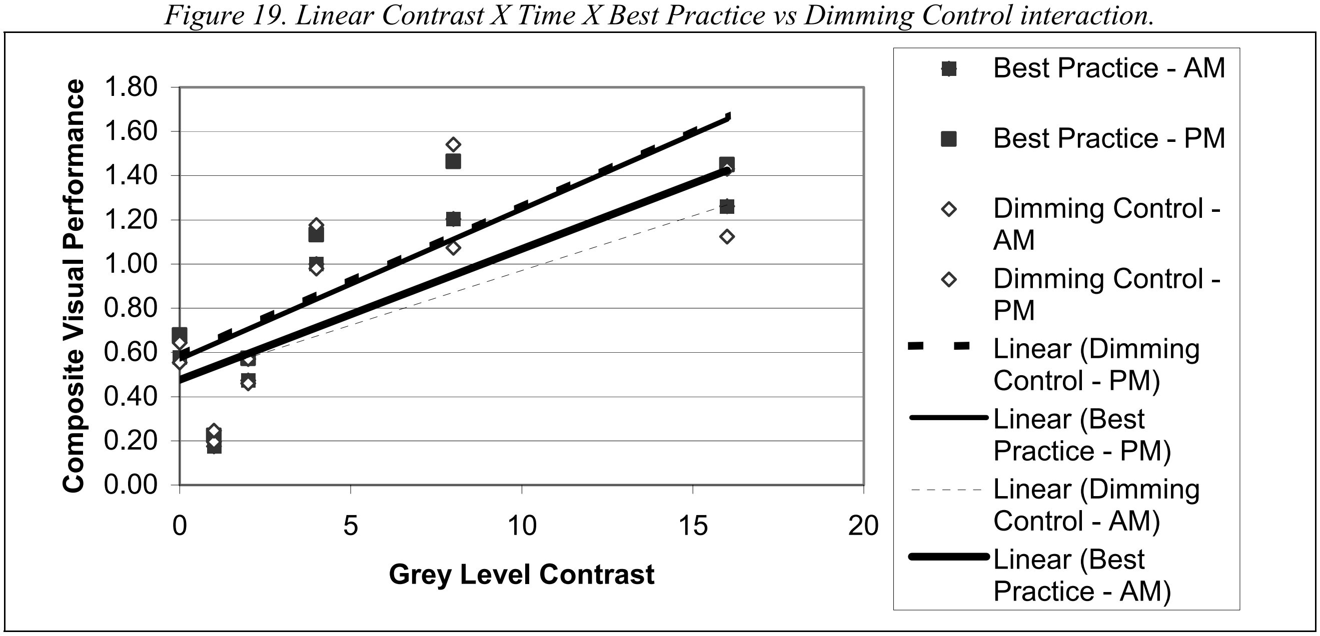
Note. *$p<.05$. *** $p<.001$.

4.4.1.3 Best Practice versus Dimming Control.

There was also a statistically significant three-way interaction of Best Practice vs Dimming Control by Contrast – Linear Trend by Time (Table 32 and Figure 19). Performance increases with increasing contrast, but the contrast effect is greater in the afternoon than the morning (i.e., contrast interacts with time, so that the slope of each line is greater for the afternoon than the morning). The three-way interaction shows that the Contrast X Time interaction effect is larger for the Dimming Control group than for the Best Practice group.

Table 33 shows the more important of the *post hoc* tests. All the linear trends are statistically significant. Figure 19 reveals that the slope is steeper for the afternoon trends, and the increase in slope from morning to afternoon is greatest for the Dimming Control condition. Contrasting the scores at two selected contrast levels, where differences might be expected to be smallest (0) and greatest (8), we see that the scores for Best Practice do not differ from Dimming Control in the morning, nor in the afternoon (Table 33). As discussed above, this interaction effect tells us that the effect of one variable (time, in this case) differs depending on the level of the other variables.

Figure 19. Linear Contrast X Time X Best Practice vs Dimming Control interaction.



Note. Trend lines show the linear effect of contrast = 0 through contrast = 16 for Best Practice and Dimming Control in the morning and afternoon.

Table 32. Best Practice vs Dimming Control by Contrast by Time interaction effect on timed vision test.

	F	*df*	$\eta^2_{partial}$	
	5.57*	1, 172	.03	
	AM		PM	
	M	*SD*	*M*	*SD*
Best Practice – 0	0.58	0.32	0.68	0.32
Best Practice – 1	0.18	0.18	0.22	0.19
Best Practice – 2	0.47	0.29	0.57	0.32
Best Practice – 4	1.00	0.51	1.13	0.51
Best Practice – 8	1.20	0.45	1.46	0.56
Best Practice – 16	1.26	0.50	1.45	0.39
Dimming Control – 0	0.55	0.29	0.64	0.33
Dimming Control – 1	0.20	0.13	0.25	0.16
Dimming Control – 2	0.46	0.26	0.57	0.26
Dimming Control – 4	0.98	0.47	1.18	0.43
Dimming Control – 8	1.07	0.44	1.54	0.46
Dimming Control – 16	1.12	0.46	1.43	0.36

Note. *p<.05. *** p<.001.

Table 33. Post hoc tests for Best Practice vs Dimming Control by Contrast by Time interaction effect on timed vision test.

	F	*df*	$\eta^2_{partial}$
Best Practice – AM	165.79***	1, 38	.81
Best Practice – PM	314.26***	1, 38	.89
Dimming Control – AM	206.37***	1, 54	.79
Dimming Control - PM	472.44***	1, 54	.90
Best Practice vs Dimming Control			
AM, contrast = 0	0.12	1, 94	
PM, contrast = 0	0.28	1, 94	
AM, contrast = 8	1.99	1, 94	
PM, contrast = 8	0.53	1, 94	

Note. *** p<.001. All of the contrasts in the lower half of the table have p>.05.

4.4.1.4 Non-lighting effects.

There were three statistically significant effects that did not involve lighting variables – that is, these effects apply regardless of lighting condition. Both the between-groups analysis and the repeated measures analysis showed these effects.

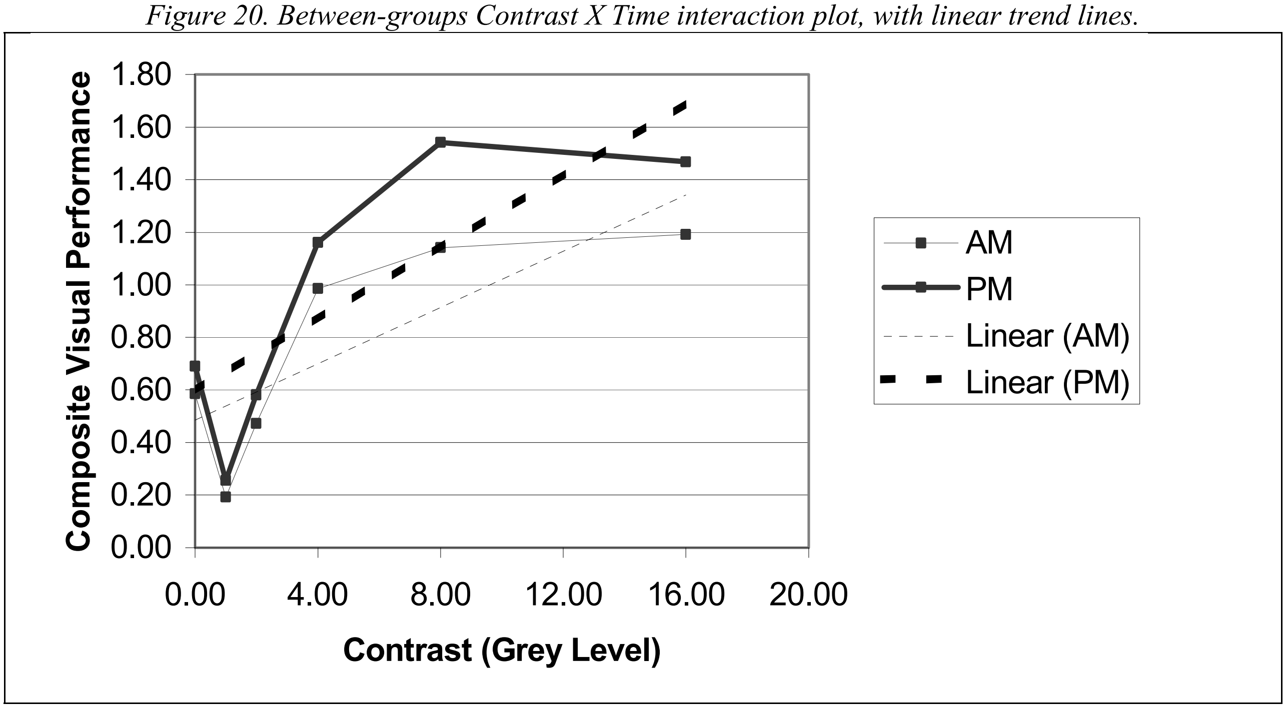
The first of these was a large two-way interaction of Contrast x Time, we interpret because of its effect size (in keeping with the guidelines for the hierarchy of interpreting interactions). As seen in Figure 20 and Table 34, the effect has the same basic shape as the interaction above: The effect of contrast on performance is stronger (steeper linear trend) in the afternoon than in the morning. The effect is large for both the between-groups and repeated-measures designs.

This was accompanied by large statistically significant main effects of both Time and Contrast for both the between-groups and repeated-measures designs. Composite visual performance increased as the target contrast increases (a linear effect). Performance also improved from the morning to the afternoon. This latter effect was unexpected in that we had anticipated that fatigue would lead to a decline in performance; instead, we see an improvement associated with having more practice on this novel computer task. Descriptive statistics for these effects are in Table 34.

Table 34. Contrast -Linear and Time effects on timed vision test.

Between-groups design	*F*	*df*	$\eta^2_{partial}$			
Contrast X Time Interaction	63.75***	1, 172	.27			
Contrast Main Effect	1613.27***	1, 172	.91			
Time Main Effect	129.08***	1, 172	.43			
	AM		PM		Contrast Effect	
Contrast (Grey Level)	*M*	*SD*	*M*	*SD*	*M*	*SD*
0	0.58	0.29	0.69	0.32	0.64	0.31
1	0.19	0.17	0.26	0.19	0.22	0.19
2	0.47	0.29	0.58	0.30	0.53	0.30
4	0.99	0.45	1.16	0.44	1.07	0.46
8	1.14	0.41	1.54	0.44	1.34	0.47
16	1.19	0.44	1.47	0.36	1.33	0.42
Time Main Effect	0.76	0.51	0.95	0.59		
Repeated-measures design	*F*	*df*	$\eta^2_{partial}$			
Contrast X Time Interaction	13.84***	1, 43	.24			
Contrast Main Effect	406.61***	1, 43	.90			
Time Main Effect	31.37***	1, 43	.42			
	AM		PM		Contrast Effect	
Contrast (Grey Level)	*M*	*SD*	*M*	*SD*	*M*	*SD*
0	0.61	0.30	0.73	0.33	0.67	0.32
1	0.24	0.18	0.29	0.22	0.26	0.20
2	0.55	0.34	0.62	0.34	0.59	0.34
4	1.08	0.51	1.25	0.48	1.16	0.50
8	1.26	0.44	1.61	0.45	1.43	0.48
16	1.31	0.46	1.49	0.40	1.40	0.44
Time Main Effect	0.84	0.56	1.00	0.61		

Figure 20. Between-groups Contrast X Time interaction plot, with linear trend lines.



Note. Solid line links the average performance at each contrast level. Trend lines (dotted) show the linear effect of contrast = 0 through contrast = 16 in the morning and afternoon.

4.4.2 Motivation

Persistence at a very difficult, almost impossible, task was the indicator of motivation. This variable was measured using the conveyor belt task, which had moving targets at increasing speeds from 1 through 9. We derived the target speed at which responding on this task became random by determining the speed at which two conditions were met: the accuracy of performance dropped to random levels (based on the frequency of targets, 0.50 of all symbols) and the number of space bar presses (the means of responding) was twice the number of symbols. This gave the speed of random responding for each trial; there were four trials in each session, and the final score was the average speed of random responding for the session. When we examined the distribution of this variable we saw that it was not normally distributed, so we transformed the variable by cubing it. This enlarged the scale, so that the maximum possible score was 729. The data for this variable were very skewed. Over all times and lighting conditions in the between-groups experimental design, the mean score was 620.39 (*SD* = 202.17), and the median was 729, indicating that half of the scores were the maximum value.

This variable gave scores at three times: session 2 (before lunch), session 3 (after lunch) and session 4 (after the afternoon coffee break). Therefore for the repeated-measures analysis the model was a 2 x 3 (Lighting Design X Time) ANOVA. There was one non-lighting effect, which is discussed below. For the between-groups analysis the model was a 4 x 2 x 3 (Lighting Design X Reflectance X Time) ANOVA, with the planned comparisons discussed earlier. It had three significant interaction effects and one non-lighting effect.

4.4.2.1 Best Practice vs Switching Control.

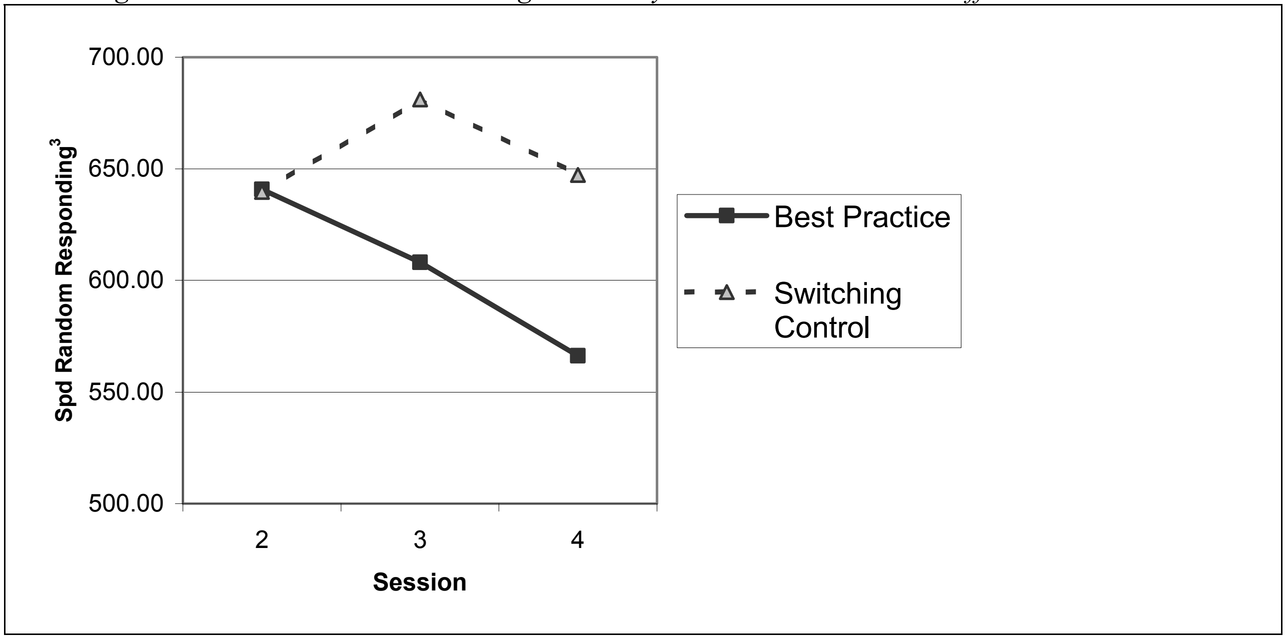
For the planned comparison between Best Practice and Best Practice + Switching Control, the interaction of this planned comparison with the linear Time effect was statistically significant (Table 35), although small. Figure 21 shows that the speed of random responding remained stable from session 2 to session 4 for the Best Practice + Switching Control group, but declined over time for the Best Practice group. Post hoc tests confirmed this interpretation: the linear trend was statistically significant for Best Practice (a drop from Session 2 to Session 4), but not for the Best Practice + Switching Control group. The two groups did not differ overall (there was no main effect for the comparison of the two lighting design conditions), and in the post hoc tests the average for the two groups was not different at any of the three sessions. Thus, the interpretation of this interaction is that the effect of one variable (time) is different for different levels of the other variable (lighting design).

Table 35. Best Practice vs Switching Control by Time-linear interaction effect on motivation.

	F	*df*	$\eta^2_{partial}$				
	4.25*	1, 173	.02				
	Session 2		Session 3		Session 4		Time-linear Trend
	M	*SD*	*M*	*SD*	*M*	*SD*	
Best Practice	640.80	177.28	607.99	195.60	566.33	243.13	$F(1, 39) = 6.10$*, $\eta^2_{partial} = .14$
Switching Control	639.59	168.56	681.01	144.02	647.15	176.76	$F(1, 31) = 0.13$

Note. *p<.05.

Figure 21. Best Practice vs Switching Control by Time-linear interaction effect on motivation.



4.4.2.2 Best Practice vs Dimming Control.

For the planned comparison between Best Practice and Dimming Control, the interaction of this planned comparison with the linear Time effect was statistically significant (Table 36, Figure 22), although small. The effect is consistent with the switching control planned comparison above: People with Dimming Control maintained their motivation scores across time, whereas the Best Practice group declined over time. There was no main effect of lighting design, and post hoc tests confirmed that there were no differences between the two groups at any of the three sessions. It appears that the availability of dimming control acts as an inoculation against the drop in motivation that occurs over the working day.

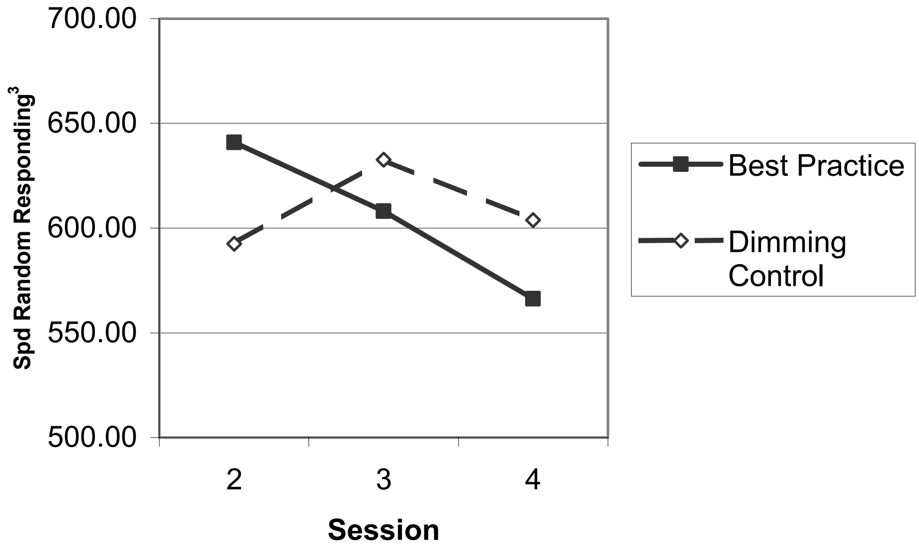
Table 36. Best Practice vs Dimming Control by Time-linear interaction effect on motivation.

	F	*df*	$\eta^2_{partial}$
	6.09**	1, 173	.03

	Session 2		Session 3		Session 4		Time-linear Trend
	M	*SD*	*M*	*SD*	*M*	*SD*	
Best Practice	640.80	177.28	607.99	195.60	566.33	243.13	*F*(1, 39) = 6.10*, $\eta^2_{partial}$ = .14
Dimming Control	592.39	232.99	632.55	193.54	603.69	210.49	*F*(1, 54) = 0.29

Note. ***p*<.01.

Figure 22. Best Practice vs Dimming Control by Time-linear interaction effect on motivation.



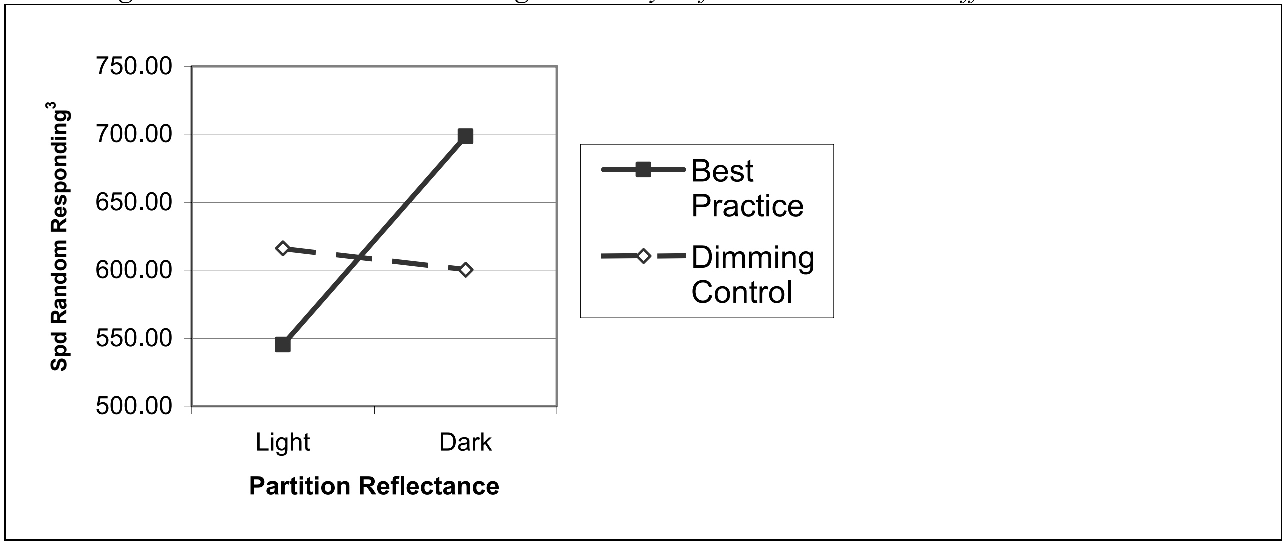
There was also an interaction of Best Practice vs Dimming Control X Reflectance, summarized in Table 37 and illustrated in Figure 23. Post hoc tests revealed that the nature of the interaction was a significant difference between motivation levels for Light versus Dark partitions under the Best Practice design, with motivation higher for the people in darker workstations, but no difference in motivation levels for the two partition reflectances under the Dimming Control condition. There was no main effect of lighting design and the two lighting design groups did not differ within any reflectance condition.

Table 37. Best Practice vs Dimming Control by Reflectance interaction effect on motivation.

	F	*df*	$\eta^2_{partial}$		
	4.82*	1, 173	.03		
	Light		Dark		Reflectance *post hoc* tests
	M	*SD*	*M*	*SD*	
Best Practice	545.31	228.47	698.36	123.60	$F(1, 39) = 7.47$**, $\eta^2_{partial} = .16$
Dimming Control	616.03	207.37	600.23	220.51	$F(1, 54) = 0.09$
Lighting Design *post hoc* tests	$F(1, 56) = 1.93$		$F(1, 37) = 3.02$		

Note. *$p<.05$. **$p<.01$.

Figure 23. Best Practice vs Dimming Control by Reflectance interaction effect on motivation.



4.4.2.3 Non-lighting effects.

In addition to the between-groups effects discussed above, there were main effects of time for both the repeated measures and the between-groups experimental designs. For the between-groups design this took the form of a small but significant Time-Quadratic trend, summarized in Table 38. The quadratic trend takes the form of an inverted-U, with performance increasing slightly from session 2 to 3 and then dropping between session 3 and 4. For the repeated-measures design the time trend is a linear decline from session 2 through session 4. This is a medium-sized effect. In both cases we see a drop in motivation towards the end of the day, consistent with the development of fatigue and boredom. The linear drop for the repeated measures design, as compared to a midday rise for the between-groups design, is probably related to greater familiarity with the experiment for those who participated more than once.

Table 38. Time effects on motivation.

Between-groups design	*F*	*df*	$\eta^2_{partial}$			
Time-Quadratic Main Effect	5.22*	1, 173	.03			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
	625.02	199.93	632.80	188.84	603.35	216.75
Repeated-measures design	*F*	*df*	$\eta^2_{partial}$			
Time-Linear Main Effect	6.03*	1, 44	.12			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
	645.63	168.38	636.56	186.59	603.63	208.96

Note. *p<.05.

4.4.3 Vigilance

There were two measures of vigilance. One was derived from the conveyor belt task: it was the average hit rate [(hits-false positives)/targets] over the three middle speeds in each trial (during which participants were still responding systematically rather than randomly), averaged over trials to give a session average score. The other was the speed of response to a prompt (response/sec), similar to the beep and symbol flash of an e-mail arriving. These two dependent measures were used in MANOVA models for the between groups and repeated measures experimental designs. For the repeated measures design the model was a 2 x 3 (Lighting Design X Time) MANOVA; only non-lighting effects were observed, which are discussed below. For the between-groups design the analysis was a 4 x 2 x 3 (Lighting Design X

Reflectance X Time) MANOVA with the previously-described planned comparisons. There was one statistically significant interaction effect involving lighting and two non-lighting main effects.

4.4.3.1 Best Practice versus Dimming Control.

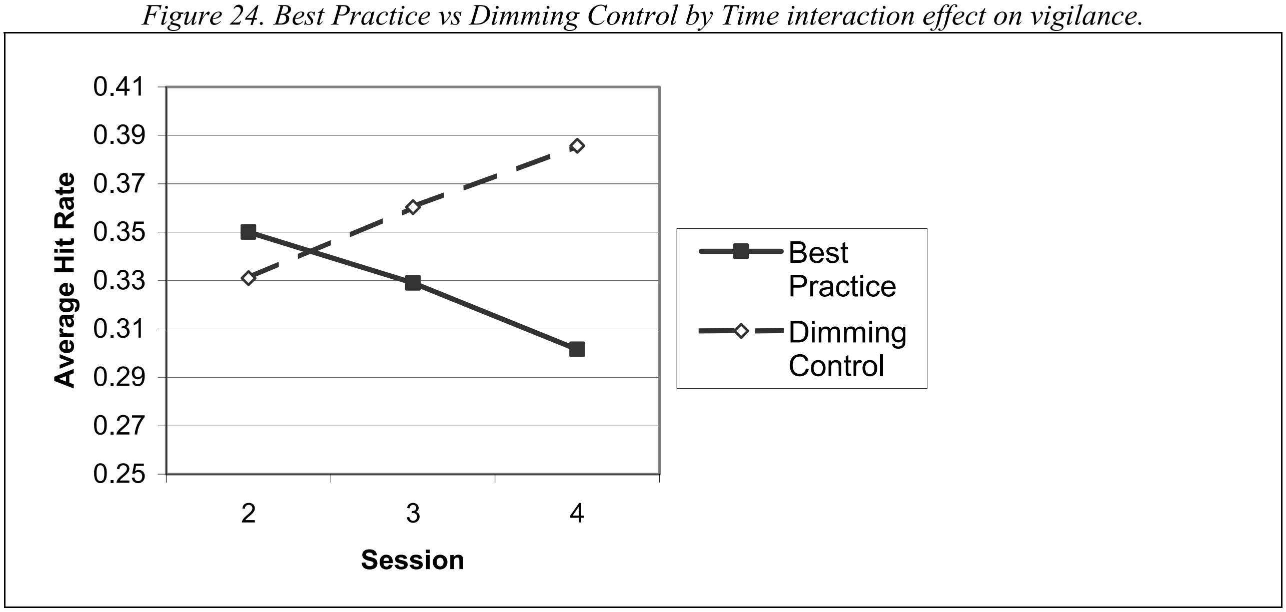
The one statistically significant interaction involving this planned comparison was between the lighting design and time-linear effects. The significant multivariate interaction effect was associated with a statistically significant medium-sized effect for conveyor belt hit rate (Table 39 and Figure 24). The *post hoc* tests showed that the hit rate for the Dimming Control group increased over time. Although the means for the Best Practice group dropped from sessions 2 through 4, this linear decline was not statistically significant. There was no main effect of lighting design, and the *post hoc* tests showed that the two groups did not differ at any of the three times.

Table 39. Best Practice vs Dimming Control by Time-linear interaction effect on vigilance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$				
Multivariate	0.95	4.22*	2, 149	.08				
Hit rate		8.10**	1, 150	.16				
		Session 2		Session 3		Session 4		
		M	*SD*	*M*	*SD*	*M*	*SD*	Linear Trends
Best Practice		0.35	0.26	0.33	0.28	0.30	0.30	*F*(1, 39) = 1.80
Dimming Control		0.33	0.24	0.36	0.23	0.39	0.24	*F*(1, 54) = 4.09*, $\eta^2_{partial}$ = .07
		F	*df*	$\eta^2_{partial}$				
Envelope Speed		0.17	1, 150	.00				
Best Practice								
Dimming Control								

Note. **p*<.05. ***p*<.01. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

Figure 24. Best Practice vs Dimming Control by Time interaction effect on vigilance.



4.4.3.2 Non-lighting effects.

There were statistically significant Time main effects for both the between-groups and the repeated measures experimental designs. These are summarized in Table 40. There was a small Time-Linear effect on envelope speed in the between-groups experimental design, but the dominant trend was quadratic, with a large effect on envelope speed. In the repeated measures design there was a large, statistically significant Time-Quadratic effect for both hit rate and envelope speed. The quadratic effects for both experiments showed the same shape: a drop in vigilance right after lunch, reflected in slower responding to the envelope prompts and a lower hit rate on the conveyor belt task at session 3 relative to sessions 2 or 4.

Table 40. Time effects on vigilance.

Between-groups design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	
Time – Linear Multivariate	0.94	4.70**	2, 149	.03	
Hit Rate		0.40	1, 150	.00	
Envelope Speed		9.31***	1, 150	.06	
Time-Quadratic Multivariate	0.57	55.78***	2, 149	.22	
Hit Rate		0.20	1, 150	.02	
Envelope Speed		112.20***	1, 150	.43	
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Envelope Speed	0.41	0.13	0.28	0.11	0.37	0.13

Repeated-measures design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Time-Quadratic Multivariate	0.28	38.05***	2, 30	.44		
Hit Rate		11.88***	1, 31	.28		
Envelope Speed		44.81***	1, 31	.59		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Hit Rate	0.41	0.25	0.38	0.27	0.42	0.27
Envelope Speed	0.44	0.14	0.29	0.11	0.39	0.11

Note. ***p*<.01. ****p*<.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.4.4 Typing

There was one dependent variable for the transcription typing task: correct characters per second. In addition to the lighting design and time variables (three sessions), there were three print sizes (8, 12, and 16 point) for the source text. Thus, for the between-groups experimental design the statistical model was a 4 x 2 x 3 x 3 Lighting Design X Reflectance X Print Size X Time ANOVA. The usual planned comparisons and interactions were examined for the lighting design and reflectance effects, and linear and quadratic trends for the Time and Print Size effects, with all possible interactions. For the repeated measures design the statistical model was a 2 x 3 x 3 Lighting Design X Print Size X Time ANOVA. The repeated measures design showed only non-lighting effects; there were several interaction effects in the between groups design.

4.4.4.1 Base Case vs. Best Practice.

There was a small but statistically significant three-way interaction of Base Case vs Best Practice X Print Size-Quadratic X Time – Quadratic. This effect is shown visually in Figure 25 and summarized in Table 41. It is a difficult effect to interpret (as are most three-way interactions). One way to express it is to say that the quadratic shape for Print Size varies quadratically by Time, in different ways for the Base Case and Best Practice conditions.

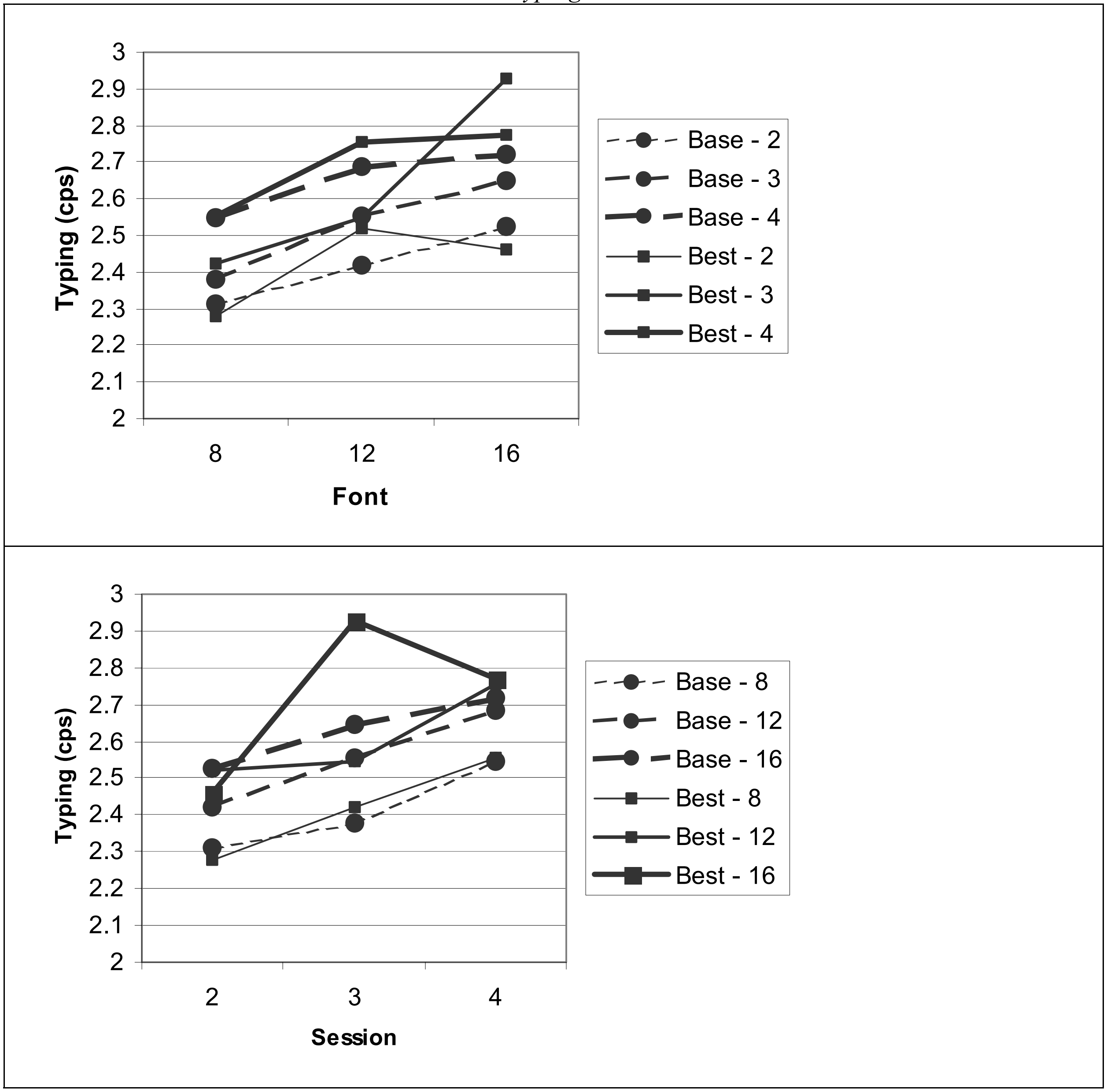
We tested the quadratic trends for each lighting design both by Print Size at each time, and by time for each Print Size (the two ways of looking at the interaction, both illustrated in Figure 25). We also tested the largest between groups differences between points (e.g., the maximum difference between points at session 2, considering both groups and all Print Size sizes). These *post hoc* tests revealed that the only statistically significant differences were the quadratic effect of Print Size for the Best Practice at session 2 (Table 42), and the difference between the lowest and highest Session 3 scores (Print Size=8 for Base Case and Print Size = 16 for Best Practice) [$F(1, 90)=4.01$, $p<.05$, $\eta^2_{partial} = .04$). One way to look at this pattern is that the Best Practice group's typing performance for 16 point source text was unusually low at Session 2 (hence the quadratic shape), and then rebounded to be unusually high at Session 3 (where it is the highest of all points and also has the highest standard deviation). There were no other differences between the groups and it is unclear why lighting design should cause this pattern. It seems most probable that a few odd cases underlie this statistically significant effect.

Table 41. Base Case vs Best Practice by Print Size-Q by Time-Q interaction effect on typing performance.

	F	*df*	$\eta^2_{partial}$			
	3.84*	1, 173	.02			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Base Case – 8 point	2.31	1.04	2.38	1.03	2.55	1.09
Base Case – 12 point	2.42	1.13	2.55	1.15	2.69	1.21
Base Case – 16 point	2.52	1.11	2.65	1.16	2.72	1.24
Best Practice – 8 point	2.28	0.93	2.42	1.06	2.55	1.10
Best Practice – 12 point	2.52	0.99	2.55	1.18	2.75	1.12
Best Practice – 16 point	2.46	1.07	2.93	1.60	2.77	1.18

Note. *$p<.05$.

Figure 25. Two views of Base Case vs Best Practice X Print Size-Q x Time-Q interaction effect on typing.



Note. Top panel shows the interaction effect plotted against Print Size. Bottom panel shows the same data reorganized to display against session.

Table 42. Post hoc tests for Base Case vs Best Practice X Print Size-Q x Time-Q interaction effect on typing.

	F	*df*	$\eta^2_{partial}$
Base Case Print Size – Q at Session 2	0.00	1, 49	
Base Case Print Size – Q at Session 3	0.49	1, 49	
Base Case Print Size – Q at Session 4	0.62	1, 49	
Best Practice Print Size – Q at Session 2	7.12**	1, 39	.15
Best Practice Print Size – Q at Session 3	1.25	1, 39	
Best Practice Print Size – Q at Session 4	1.44	1, 39	
Base Case Time – Q at 8 point	0.75	1, 49	
Base Case Time – Q at 12 point	0.00	1, 49	
Base Case Time – Q at 16 point	0.47	1, 49	
Best Practice Time – Q at 8 point	0.02	1, 39	
Best Practice Time – Q at 12 point	1.07	1, 39	
Best Practice Time – Q at 16 point	3.36	1, 39	

Note. ** $p<.01$.

4.4.4.2 Best Practice vs Best Practice + Switching Control.

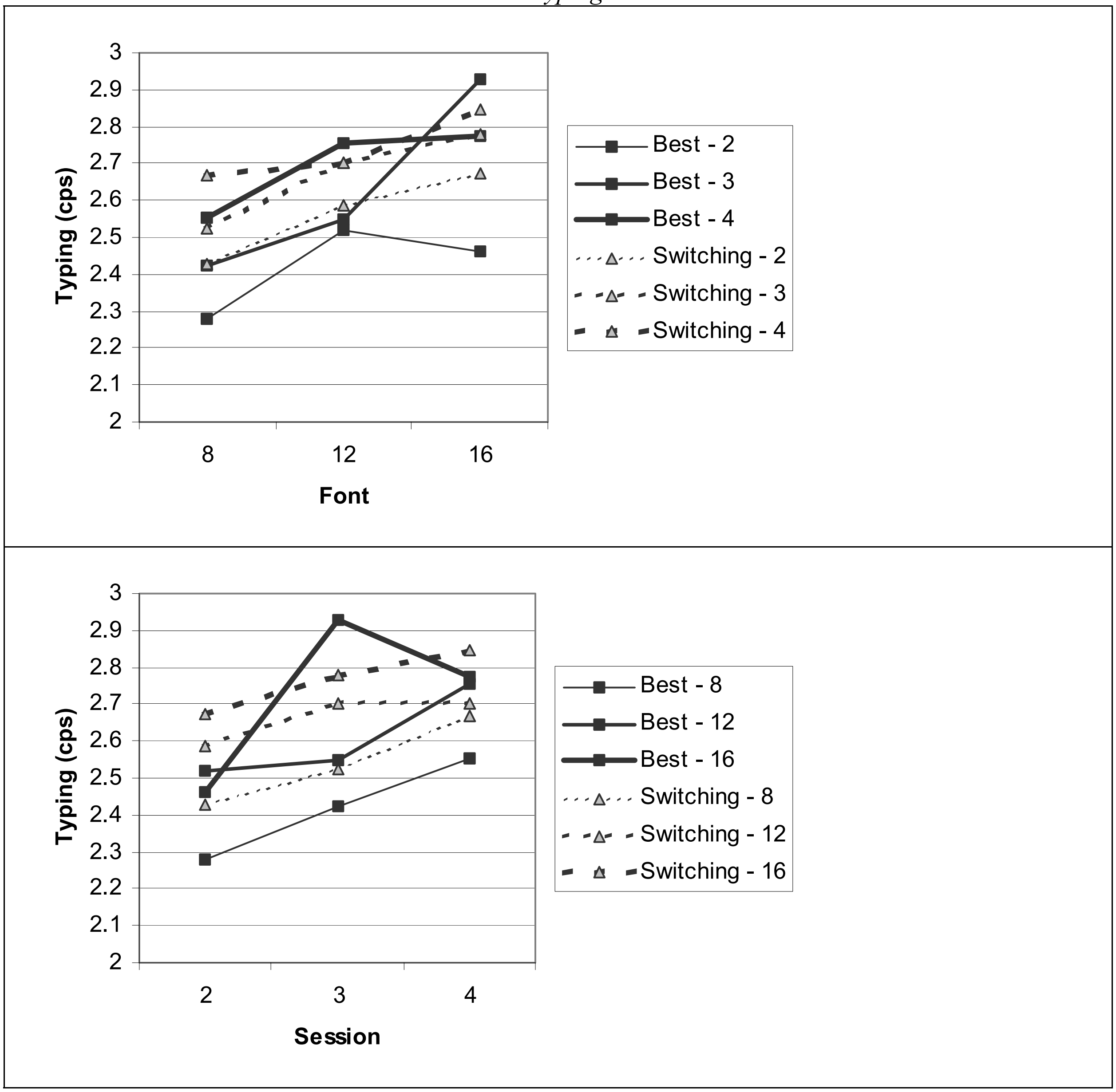
This planned comparison also showed a small, statistically significant three-way interaction of Best Practice vs Best Practice + Switching Control by Print Size-Quadratic by Time-Quadratic (Table 43). For this interaction we also plotted two views (Figure 26), and conducted *post hoc* tests (Table 44). As for the interaction with Base Case vs Best Practice, the *post hoc* tests revealed that the only significant quadratic trend was the Print Size trend for Best Practice at Session 2, where the performance for 16-point text was unusually low, giving a quadratic trend. We also tested the statistical significance of the maximum between-groups differences at each session (across lighting conditions and Print Sizes) and at each Print Size (across lighting conditions and sessions). None of these pairs achieved statistically significant differences, not even the pairs that included the high mean value for Best Practice with16 point Print Size at session 3 (which had greater variability than all of the other means).

Table 43. Best Practice vs Switching Control by Print Size-Q by Time-Q interaction effect on typing performance.

	F	*df*	$\eta^2_{partial}$			
	4.14*	1, 173	.02			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Best Practice – 8 point	2.28	0.93	2.42	1.06	2.55	1.10
Best Practice – 12 point	2.52	0.99	2.55	1.18	2.75	1.12
Best Practice – 16 point	2.46	1.07	2.93	1.60	2.77	1.18
Switching – 8 point	2.43	1.06	2.53	0.98	2.67	1.18
Switching – 12 point	2.58	1.12	2.70	1.07	2.70	1.11
Switching – 16 point	2.67	1.00	2.78	1.04	2.85	1.01

Note. *$p<.05$.

Figure 26. Two views of Best Practice vs Switching Control X Print Size-Q x Time-Q interaction effect on typing.



Note. Top panel shows the interaction effect plotted against Print Size. Bottom panel shows the same data reorganized to display against session.

Table 44. Post hoc tests for Best Practice vs Switching Control X Print Size-Q x Time-Q interaction effect on typing.

	F	*df*	$\eta^2_{partial}$
Best Practice Print Size – Q at Session 2	7.12**	1, 39	.15
Best Practice Print Size – Q at Session 3	1.25	1, 39	
Best Practice Print Size – Q at Session 4	1.44	1, 39	
Switching Print Size – Q at Session 2	0.25	1, 31	
Switching Print Size – Q at Session 3	0.41	1, 31	
Switching Print Size – Q at Session 4	0.26	1, 31	
Best Practice Time – Q at 8 point	0.02	1, 39	
Best Practice Time – Q at 12 point	1.07	1, 39	
Best Practice Time – Q at 16 point	3.36	1, 39	
Switching Time – Q at 8 point	0.06	1, 31	
Switching Time – Q at 12 point	0.38	1, 31	
Switching Time – Q at 16 point	0.09	1, 31	

Note. ** $p<.01$.

4.4.4.3 Best Practice vs Dimming Control.

This planned comparison also showed a small, statistically significant three-way interaction of Best Practice vs Dimming Control by Print Size-Quadratic by Time-Quadratic. As for the previous two interactions, we have plotted it in two orientations (Figure 27) and also present a summary table (Table 45) and a table for the *post hoc* tests (Table 46).

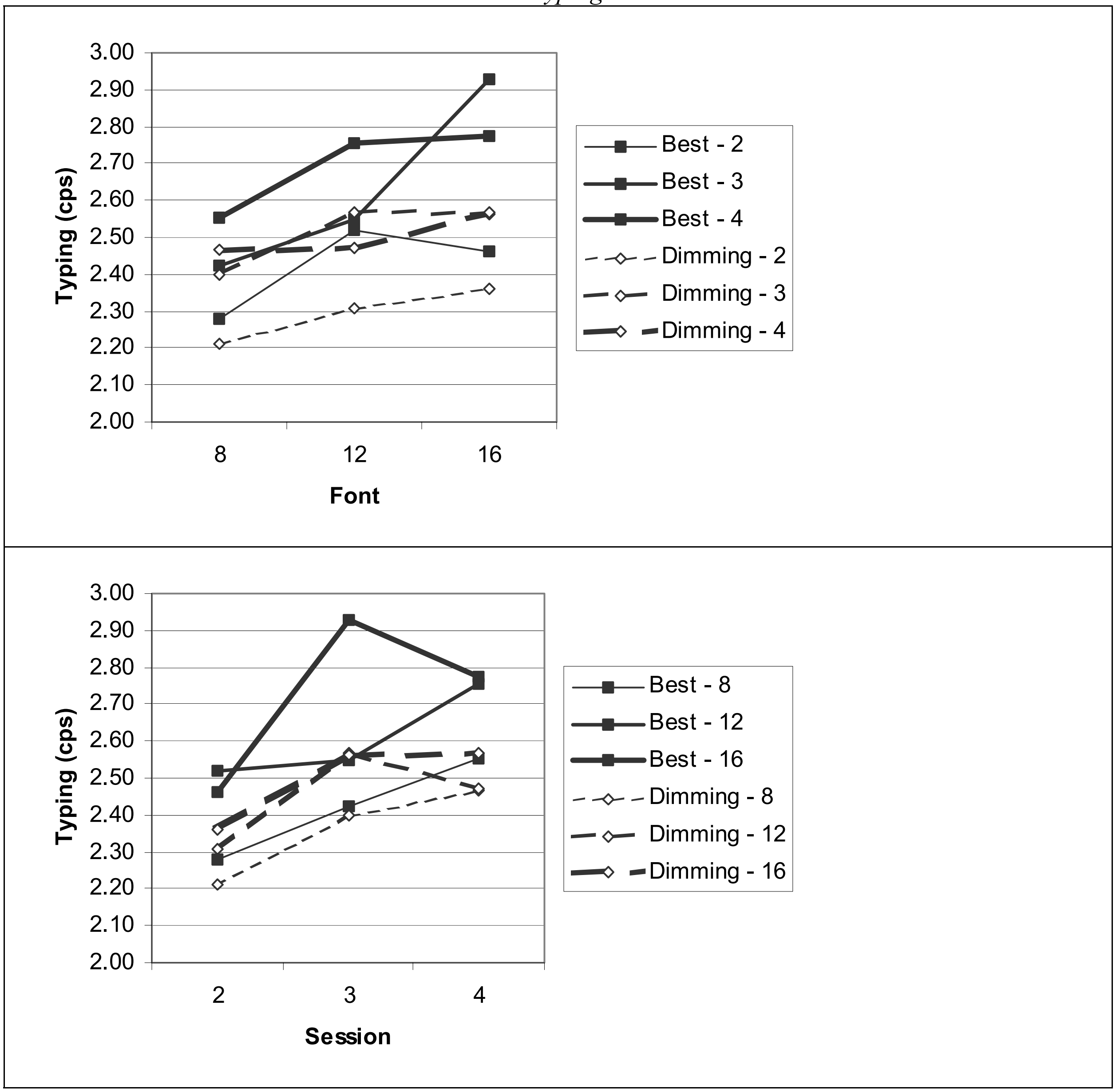
This interaction is somewhat different from the others in that the post hoc tests showed two significant quadratic trends for the Dimming Control group (Time-Q for 12 point and 16 point Print Sizes), in addition to the one significant trend for the Best Practice group (Print Size-Q at session 2). These two trends reveal that performance in the Dimming Control group for 12 and 16 point Print Sizes increased from session 2 to session 3, but then did not increase further in session 4. We also tested the largest between-groups differences across Print Sizes for each session, and across sessions for each Print Size. One of these planned comparisons was statistically significant, the comparison at 16 point Print Size between Best Practice in Session 3 and Dimming Control in Session 2 ($F(1, 95) = 4.09$, $p<.05$, $\eta^2_{partial} = .04$). As one may see in Figure 27, the Best Practice-Session 3-16 point mean is anomalous.

Table 45. Best Practice vs Dimming Control by Print Size-Q by Time-Q interaction effect on typing performance.

	F	*df*	$\eta^2_{partial}$			
	6.73**	1, 173	.04			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Best Practice – 8 point	2.28	0.93	2.42	1.06	2.55	1.10
Best Practice – 12 point	2.52	0.99	2.55	1.18	2.75	1.12
Best Practice – 16 point	2.46	1.07	2.93	1.60	2.77	1.18
Dimming – 8 point	2.21	1.11	2.40	1.23	2.47	1.27
Dimming – 12 point	2.31	1.16	2.57	1.22	2.47	1.23
Dimming – 16 point	2.36	1.18	2.56	1.24	2.57	1.24

Note. **$p<.01$.

Figure 27. Two views of Best Practice vs Dimming Control X Print Size-Q x Time-Q interaction effect on typing.



Note. Top panel shows the interaction effect plotted against Print Size. Bottom panel shows the same data reorganized to display against session.

Table 46. Post hoc tests for Best Practice vs Dimming Control X Print Size-Q x Time-Q interaction effect on typing.

	F	*df*	$\eta^2_{partial}$
Best Practice Print Size – Q at Session 2	7.12**	1, 39	.15
Best Practice Print Size – Q at Session 3	1.25	1, 39	
Best Practice Print Size – Q at Session 4	1.44	1, 39	
Dimming Print Size – Q at Session 2	0.47	1, 54	
Dimming Print Size – Q at Session 3	2.40	1, 54	
Dimming Print Size – Q at Session 4	0.11	1, 54	
Best Practice Time – Q at 8 point	0.02	1, 39	
Best Practice Time – Q at 12 point	1.07	1, 39	
Best Practice Time – Q at 16 point	3.36	1, 39	
Dimming Time – Q at 8 point	0.53	1, 54	
Dimming Time – Q at 12 point	7.93**	1, 54	.13
Dimming Time – Q at 16 point	4.17*	1, 54	.07

Note. ** *p*<.01.

There was also a statistically significant two-way interaction between Best Practice vs Dimming Control by Print Size-linear. It is modified by the three-way interaction. As may be seen in Figure 28, the shape of the interaction is that although increasing Print Size improves typing performance for both groups, the increase is greater for the Best Practice group than the Dimming Control group. *Post hoc* tests revealed that both lines are themselves statistically significant, but the performance levels for the two groups never differ by a statistically significant degree (Table 47). That is, although it might appear from Figure 28 that the mean performance is higher for the Best Practice group, this is not a statistically significant difference.

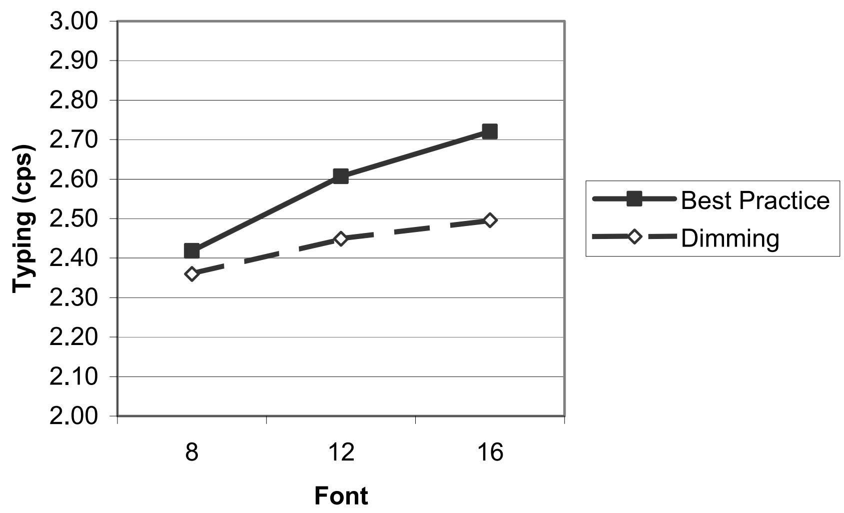
Taken together, the two interactions involving this between-groups planned comparison are puzzling. Among those with Dimming Control, typing performance at higher, more visible Print Size sizes increased from session 2 to session 3 but not further at session 4 (the three-way interaction), contributing to the overall smaller increase in performance across Print Size sizes (the two-way interaction). Given that the two groups are not different overall (despite on average a 100-lx difference in illuminance on the document holder, 501 lx for Best Practice vs 404 lx for Dimming Control), one interpretation of this effect is that making the source text more visible (by increasing print size) has a stronger effect on performance when there is no control available, and the effect seems to have increasing influence as the day progresses.

Table 47. Best Practice vs Dimming Control by Print Size-L interaction effect on typing.

	F	*df*	$\eta^2_{partial}$				
	7.24**	1, 173	.04				
	8 pt		12 pt		16 pt		Print Size-L *post hoc* tests
	M	*SD*	*M*	*SD*			
Best Practice	2.42	1.03	2.61	1.10	2.72	1.30	*F*(1, 39) = 19.96***, $\eta^2_{partial}$ = .34
Dimming Control	2.36	1.20	2.45	1.20	2.50	1.22	*F*(1, 54) = 10.73***, $\eta^2_{partial}$ = .17
Lighting Design *post hoc* tests	*F*(1, 95) = 0.07		*F*(1, 95) = 0.48		*F*(1, 95) = 0.35		

Note. **p*<.05. ***p*<.01. ****p*<.001

Figure 28. Best Practice vs Dimming Control by Print Size-L interaction effect on typing.



4.4.4.4 Reflectance effects.

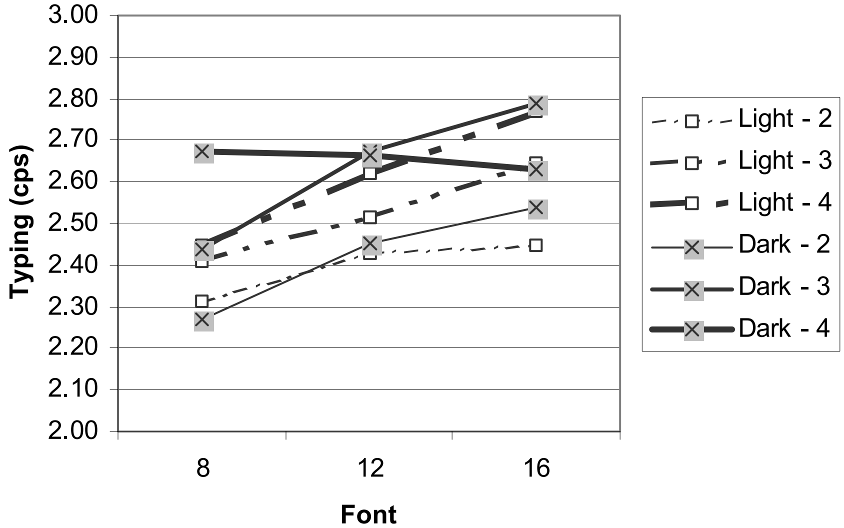
There were two interaction effects involving partition reflectance, one three-way and one two-way. The three-way interaction effect was medium-sized, and involved Reflectance X Print Size-linear X Time (see Table 48 and Figure 29). *Post hoc* tests revealed two aspects of the interaction (Table 49). Considered across Print Sizes by sessions (top panel), we see that all the trends are for increasing typing performance for larger Print Sizes in all sessions and for both reflectances, except for the dark partitions in session four. Considered across sessions by Print Sizes (lower panel), we see that all the trends are for increasing performance as the day progressed, except for the dark partitions at 16 point Print Size, which had no overall increase from session 2 to 4. One interpretation of this (given the absence of any between-groups main effect) is that increasing Print Size is a less important contributor to typing performance when the partition reflectance is low and people have become practiced. Perhaps the greater luminance contrast between the source text on the document holder and the partition contributes to making the source text more obvious or salient, so that there is no additional benefit to making the text very large.

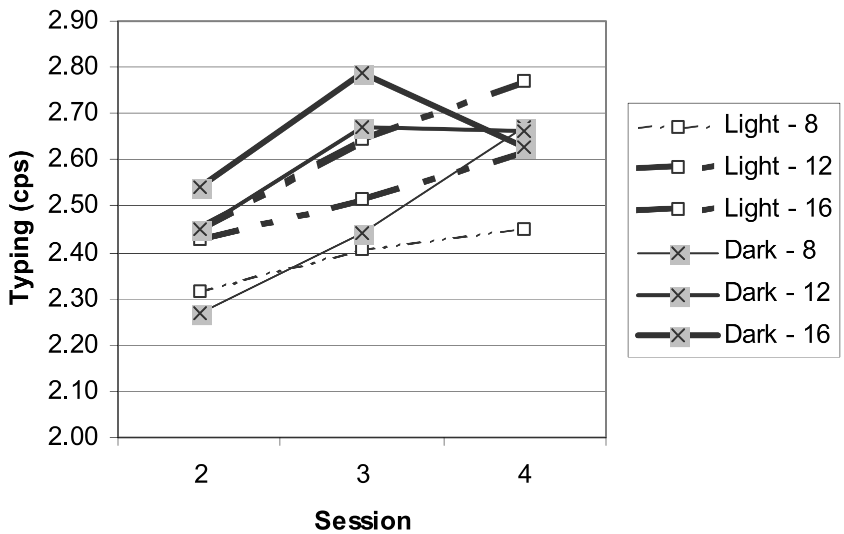
Table 48. Partition Reflectance by Print Size-L by Time-L interaction effect on typing performance.

	F	*df*	$\eta^2_{partial}$			
	19.64***	1, 173	.10			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Light – 8 point	2.31	0.96	2.41	1.02	2.45	1.05
Light – 12 point	2.43	1.05	2.51	1.06	2.62	1.12
Light – 16 point	2.44	1.04	2.65	1.12	2.77	1.12
Dark – 8 point	2.27	1.14	2.44	1.18	2.67	1.29
Dark – 12 point	2.45	1.18	2.67	1.28	2.66	1.26
Dark– 16 point	2.54	1.18	2.79	1.46	2.63	1.26

Note. ***p*<.01.

Figure 29. Two view of the Reflectance X Print Size-L x Time-L interaction effect on typing.





Note. Top panel shows the interaction effect plotted against Print Size. Bottom panel shows the same data reorganized to display against session.

Table 49. Post hoc tests for Partition Reflectance by Print Size-L by Time-L interaction effect on typing.

	F	*df*	$\eta^2_{partial}$
Light Print Size – L at Session 2	9.01***	1, 98	.08
Light Print Size – L at Session 3	21.29***	1, 98	.18
Light Print Size – L at Session 4	27.68***	1, 98	.22
Dark Print Size – L at Session 2	28.15***	1, 75	.27
Dark Print Size – L at Session 3	12.62***	1, 75	.14
Dark Print Size – L at Session 4	0.21	1, 75	
Light Time – L at 8 point	7.77**	1, 98	.07
Light Time – L at 12 point	6.82**	1, 98	.07
Light Time – L at 16 point	32.66***	1, 98	.25
Dark Time – L at 8 point	48.58***	1, 75	.39
Dark Time – L at 12 point	17.07***	1, 75	.19
Dark Time – L at 16 point	2.99	1, 75	

Note. ** p<.01. ***p<.001.

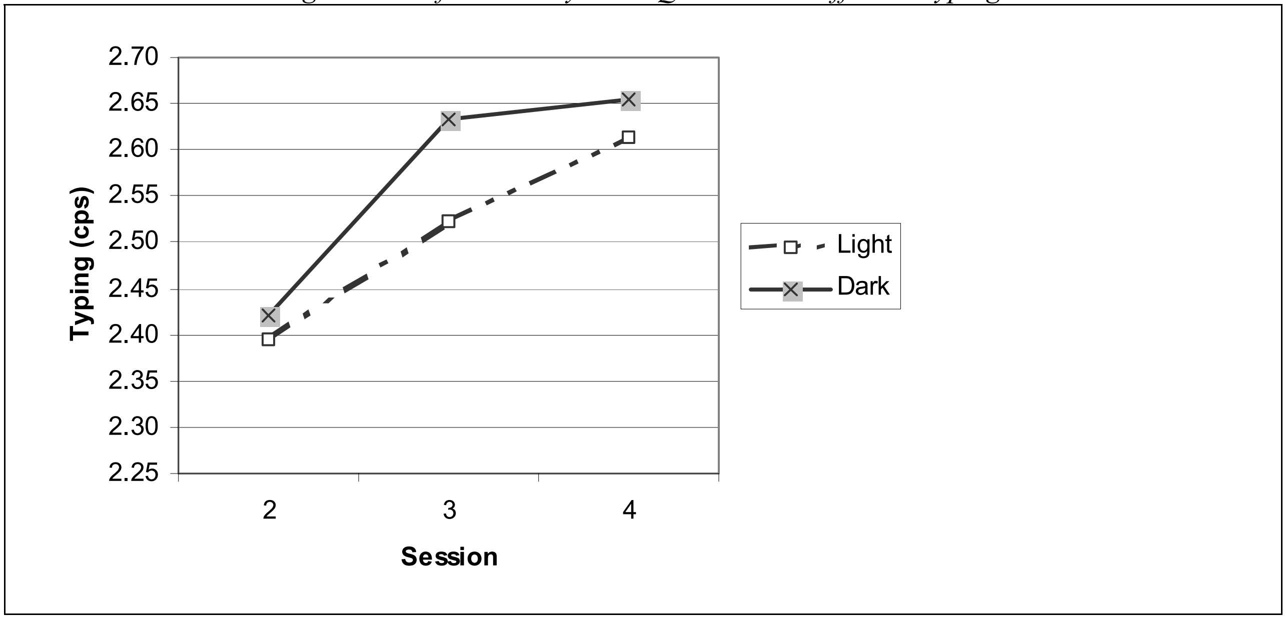
A two-way interaction of Reflectance by Time-Quadratic relates to the three-way interaction (Figure 30, Table 50). In this interaction, collapsed across all Print Sizes, we see that for the dark partitions there is a flattening of the curve between sessions 3 and 4, whereas there is a continuous increase for the light partitions. It appears that increasing practice has less effect for people with dark partitions, across all print sizes; however we know from the above three-way interaction that this effect is strongest for the 16-point print size, where visibility is not a factor. Note that the *post hoc* tests for this interaction effect reinforce the main effect tests; overall performance was not different for the two partition reflectance groups.

Table 50. Reflectance by Time-Q interaction effect on typing.

	F	*df*	$\eta^2_{partial}$				
	5.21*	1, 173	.03				
	Session 2		Session 3		Session 4		Time-Q *post hoc* tests
	M	*SD*	*M*	*SD*	*M*	*SD*	
Light Partitions	2.40	1.01	2.52	1.07	2.61	1.10	$F(1, 98) = 0.03$
Dark Partitions	2.42	1.17	2.63	1.31	2.65	1.26	$F(1, 75) = 7.98$**, $\eta^2_{partial} = .10$
Lighting Design *post hoc* tests	$F(1, 179) = 0.02$		$F(1, 179) = 0.44$		$F(1, 179) = 0.06$		

Note. *p<.05. **p<.01.

Figure 30. Reflectance by Time-Q interaction effect on typing.



4.4.4.5 Non-lighting effects.

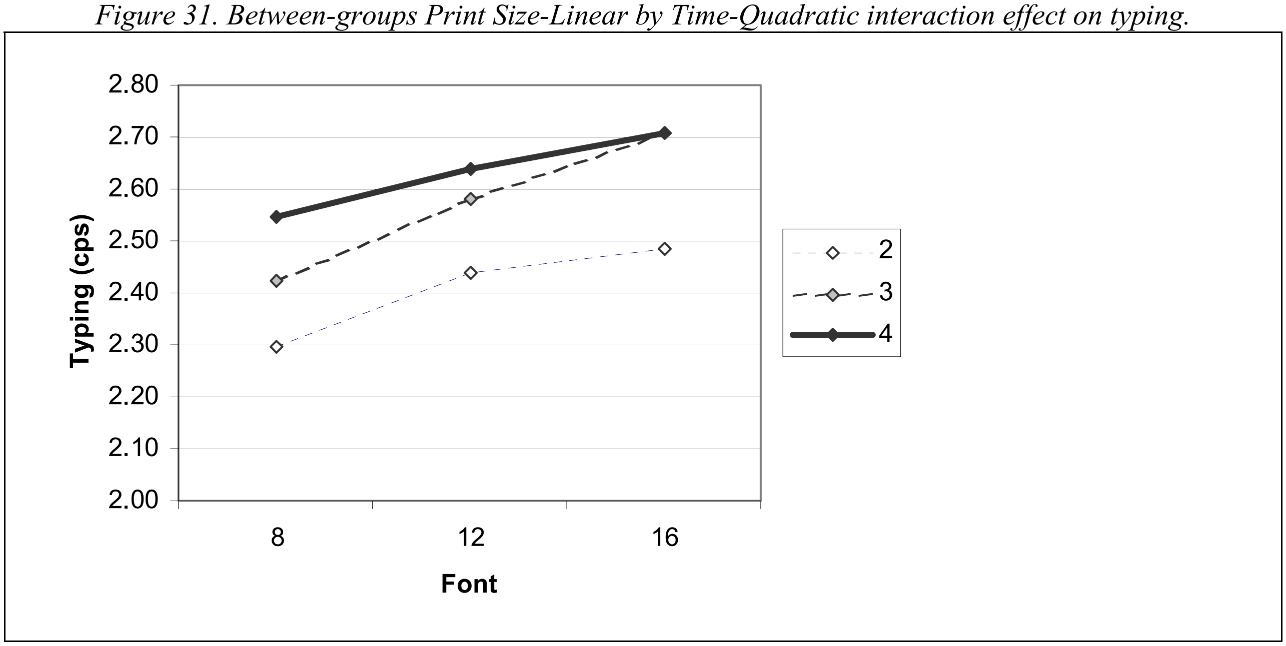
The non-lighting effects were consistent with predictions based on source text visibility and practice. The linear main effects are large enough that there is no question of their interpretation even in light of the various interaction effects. Typing performance increased with increasing source text Print Size and with increasing practice, in both the between-groups and repeated-measures experimental designs (Table 51). For the between-groups design only, there was a small interaction effect in the between-groups design in which Print Size-linear interacted with Time-Quadratic (Figure 31). The linear Print Size effect seems to have been steepest in Session 3. Overlaid on this is a small quadratic effect of time in which the performance did not increase quite so much between sessions 3 and 4 as it had between sessions 2 and 3. This could be the result of fatigue or boredom (a factor mentioned by several participants in their open-ended comments), or diminishing returns of practice.

Table 51. Time and Print Size effects on typing.

Between-groups design	*F*	*df*	$\eta^2_{partial}$					
Print Size – L X Time – Q Interaction	4.44*	1, 173	.03					
Print Size – L Main Effect	77.89***	1, 173	.31					
Time – L Main Effect	99.53***	1, 173	.37					
Time – Q Main Effect	6.15**	1, 173	.03					
	Session 2		Session 3		Session 4		Print Size Effect	
Print Size	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
8	2.29	1.04	2.42	1.09	2.55	1.16	2.42	1.10
12	2.44	1.10	2.58	1.16	2.64	1.18	2.55	1.15
16	2.49	1.10	2.71	1.27	2.71	1.18	2.63	1.19
Time Main Effect	2.41	1.08	2.57	1.18	2.63	1.17		
Repeated-measures design	*F*	*df*	$\eta^2_{partial}$					
Print Size -LMain Effect	23.82***	1, 44	.35					
Time-L Main Effect	39.32***	1, 44	.47					
	Session 2		Session 3		Session 4		Print size Effect	
Print Size	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
8							2.28	1.13
12							2.40	1.16
16							2.43	1.12
Time Main Effect	2.27	1.12	2.40	1.12	2.45	1.17		

Note. *$p<.05$. **$p<.01$. ***$p<.001$. Descriptive statistics are provided for statistically significant effects only.

Figure 31. Between-groups Print Size-Linear by Time-Quadratic interaction effect on typing.



4.4.5 Cognitive Judgments

4.4.5.1 Paper-based articles.

Midway through the afternoon occurred a group discussion task, which provided a background for social behavior measurements described below. The preparation for the task also provided information about cognitive judgments about material read from paper, rather than on the computer screen. Before congregating for the group discussion, each person read eight short articles that were the basis of the group discussion. They provided four judgments for each article: how easy it was to read (-3 to +3), how interesting they found it (-3 to +3), how relevant it might be to the (fictitious) magazine in which it was ostensibly intended to appear (-3 to +3), and how likely it would be that the group would choose that article to be the lead in the magazine (0-100 % probability). We created average scores for each of these four judgments (across the 8 articles). Internal consistency reliability for these scales, calculated for the between-groups experimental design, ranged from was just acceptable to good: For Interest, Cronbach's alpha = .57; for Relevance, α = .60; for Ease, α = .64; for Probability, α = .72.

We tested the hypothesis that lighting conditions might alter the judgments. For the between-groups design the model was a 4 x 2 (Lighting Design X Reflectance) MANOVA with the four judgments as dependent variables. For the repeated-measures design it was a MANOVA comparison between the two lighting conditions. There were no statistically significant differences in these cognitive judgments for either analysis.

4.4.5.2 Computer-based articles.

Participants spent much of each session evaluating short summaries and longer articles. We used the averages of articles judged in the first six trials in each session to create scores for the interest in reading the whole article based on a summary (0-100); the judgment that the summary includes the important facts in the article; the judgment that the summary is grammatically correct; and, the judgment that the summary is well-written (the latter three all on scales from 0 – 4). These scales had acceptable internal consistency reliability (for Interest, session 2 α = .76; session 3, α = .71, session 4, α = .63; for Important Facts, session 2 α = .67; session 3, α = .67, session 4, α = .65; for Grammatically Correct, session 2 α = .59; session 3, α = .63, session 4, α = .62; for Well-written, session 2 α = .67; session 3, α = .62, session 4, α = .67).

These four judgments were entered in MANOVA models for each experimental design. For the between-groups design the model was a 4 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA; for the repeated measures design it was a 2 x 3 (Lighting Design X Session) MANOVA.

4.4.5.2.1 Base Case vs Dimming Control.

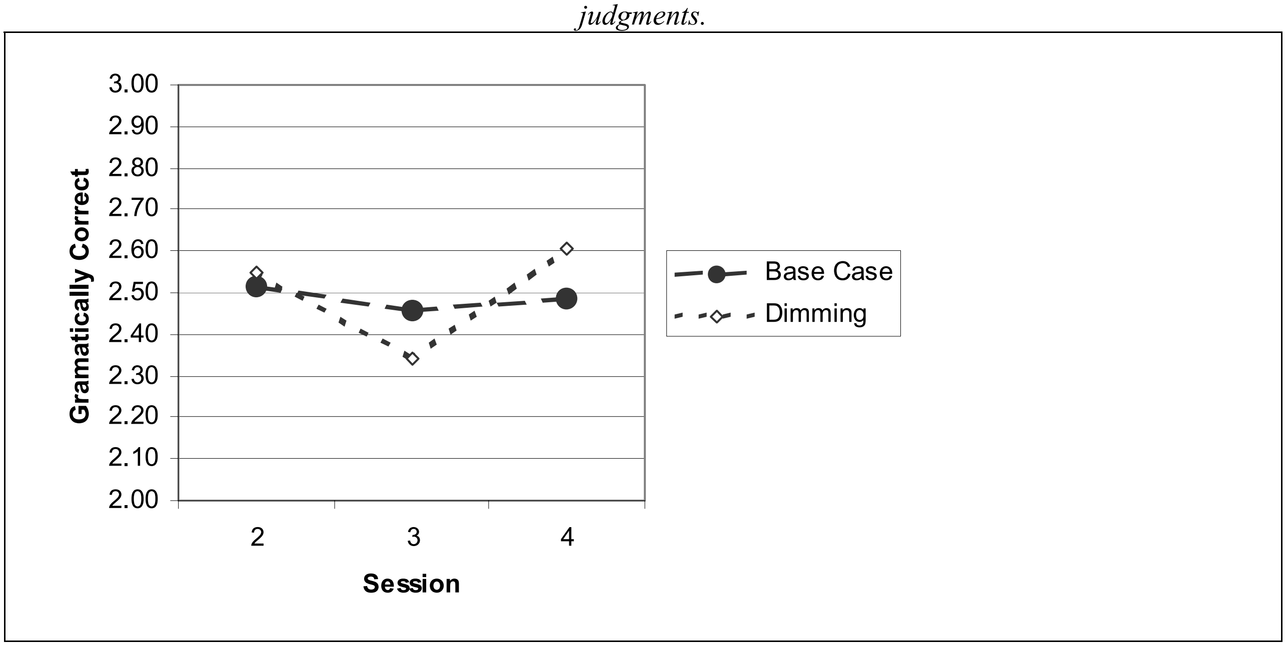
There was one lighting effect, an interaction in the repeated measures design of Base Case vs Dimming Control by Time-Quadratic. The significant multivariate test was accompanied by a significant univariate test for the judgment that the summary was grammatically correct (see Table 52 and Figure 32). Post hoc tests showed that although there was no quadratic trend for the Base Case condition, there was a statistically significant quadratic trend for the Dimming Control condition, in which ratings were higher in sessions 2 and 4 than in session 3. In addition, although there was no overall main effect of lighting design the two groups did differ in this judgment in Session 4, towards the end of the day.

Table 52. Base Case vs Dimming Control by Time-quadratic interaction effect on cognitive judgments.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$				
Multivariate	0.74	3.55**	4, 41	.05				
Interest		0.09	1, 44	.00				
Factual		0.98	1, 44	.02				
Well-written		0.36	1, 44	.01				
Grammatically Correct		7.73**	1, 44	.15				
		Session 2		Session 3		Session 4		
		M	*SD*	*M*	*SD*	*M*	*SD*	Quadratic Trends
Base Case		2.51	0.44	2.46	0.46	2.48	0.51	*F*(1, 44) = 0.59
Dimming Control		2.55	0.58	2.34	0.52	2.60	0.49	*F*(1, 44)= 21.30***, $\eta^2_{partial}$ = .33
Lighting Design *post hoc* tests		*F*(1, 44) = 0.61		*F*(1, 44) = 2.86		*F*(1, 44) = 4.17*, $\eta^2_{partial}$ = .09		

Note. *p<.05. **p<.01. ***p<.001. Group descriptive statistics and *post hoc* tests are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

Figure 32. Repeated measures Base Case vs Dimming Control by Time-Quadratic effect on cognitive judgments.



4.4.5.2.2 Non-lighting effects.

Both the between-groups and repeated measures designs showed linear and quadratic effects of time on the computer-based cognitive judgments. These are summarized in Table 53 and illustrated in Figure 33.

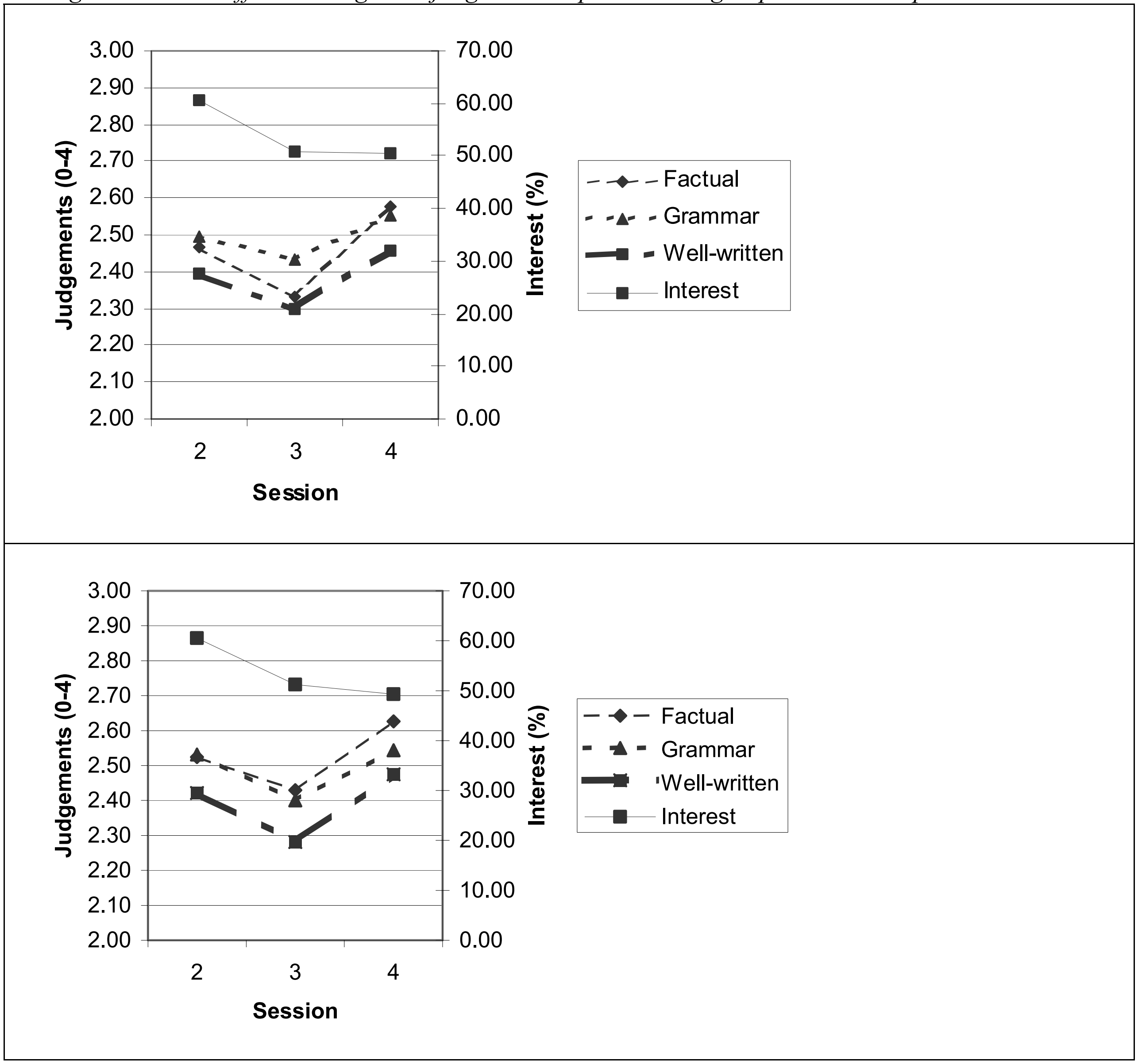
For the between-groups design, Interest dropped from session 2 to session 3 but then stayed steady at session 4 (the quadratic shape), giving an overall linear drop. The other ratings showed a decline from session 2 to session 3 and then an increase from session 3 to session 4 (the quadratic shape, found for two of the three)), for an overall increase for the linear effect (found for all three). Generally similar effects occurred in the repeated measures experimental design. We suspect that the drop in interest reflects fatigue, as it is consistent with the direction of the mood shift from morning to afternoon. The quadratic shape of the other ratings might reflect a change in strategy over the day: initially (relatively) higher ratings with less experience, more critical judgments in session 3 with practice, and finally less critical judgments as fatigue and boredom developed. This interpretation draws also on the patterns in the cognitive performance data discussed below.

Table 53. Time effects on cognitive judgments.

Between-groups design	Wilks' Λ		*F*		*df*	$\eta^2_{partial}$
Linear: Multivariate	0.73		15.48***		4, 170	.07
Interest			47.25***		1, 173	.21
Factual			4.24*		1, 173	.02
Grammar			4.24*		1, 173	.02
Well-written			2.00		1, 173	.01
Quadratic: Multivariate	0.81		9.72***		4, 170	.08
Interest			11.51***		1, 173	.06
Factual			28.86***		1, 173	.14
Grammar			7.47**		1, 173	.04
Well-written			15.36***		1, 173	.08
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Interest	60.74	18.60	50.86	18.50	50.58	16.82
Factual	2.47	0.64	2.33	0.63	2.58	0.58
Grammar	2.49	0.50	2.43	0.52	2.55	0.48
Well-written	2.39	0.58	2.30	0.56	2.46	0.55
Repeated-measures design	Wilks' Λ		*F*		*df*	$\eta^2_{partial}$
Linear: Multivariate	0.61		6.67***		4, 41	.11
Interest			23.24***		1, 44	.35
Factual			3.91*		1, 44	.08
Grammar			0.06		1, 44	.00
Well-written			0.79		1, 44	.02
Quadratic: Multivariate	0.70		4.41***		4, 41	.19
Interest			3.90*		1, 44	.08
Factual			13.11***		1, 44	.23
Grammar			13.55***		1, 44	.24
Well-written			12.48***		1, 44	.22
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Interest	60.50	16.88	51.21	18.06	49.27	16.11
Factual	2.52	0.67	2.43	0.71	2.63	0.62
Grammar	2.53	0.51	2.40	0.49	2.54	0.50
Well-written	2.42	0.62	2.28	0.62	2.47	0.60

Note. *p<.05. **p<.01. ***p<.001.

Figure 33. Time Effects on cognitive judgments. Top: Between-groups. Bottom: Repeated Measures



4.4.6 Cognitive Performance

4.4.6.1 Speed and accuracy.

There were four dependent measures associated with cognitive tasks. For the various parts of the work with summaries and articles, which is called the Categorization task above, there was an accuracy score for the assignment of summaries to content categories (# correct per session, squared to improve the shape of the distribution), and two speed measures: the time take to do the categorization, and the time required to make the four judgments reported above. In addition, we had a measurement of the time taken on the Summarization task, in which participants were given a full article and required to select sentences to form a summary of it. (There was no accuracy score available for this task.) Average scores for these dependent variables were calculated for the first six trials in each of three sessions (omitting session1 which was a training session). For the between-groups design the statistical model was thus a 4 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA, whereas for the repeated measures design it was a 2 x 3 (Lighting Design X Session) MANOVA. There was a lighting effect and several non-lighting effects.

4.4.6.1.1 Base Case vs Best Practice.

There was an overall main effect for this between-groups planned comparison; the significant multivariate test was accompanied by a small statistically significant univariate test for the time required to make the judgments. This is summarized in Table 54. Participants in the Best Practice condition took longer to make their judgments about the articles. This could mean that they took time to be more thoughtful; alternatively it could also mean that they were less efficient.

Table 54. Base Case vs Best Practice effect on cognitive performance.

					Base Case		Best Practice	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	0.93	3.24**	4, 170	.02				
Cat. Time		1.40	1, 173	.01				
Cat. Acc.		0.39	1, 173	.00				
Judge Time		9.96***	1, 173	.05	72.35	34.83	93.21	45.96
Summ. Time		2.60	1, 173	.02				

Note. ***p*<.01. ****p*<.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.4.6.1.2 Non-lighting effects.

Both the between-groups and repeated-measures experimental designs showed both linear and quadratic effects of time on measures of cognitive performance. These are summarized in Table 55 and illustrated in Figure 34. All of the linear effects are very large, and they appear in both experimental designs. The quadratic effects are smaller and do not occur on all dependent measures. In general the categorization task became both faster (requiring less time) and more accurate as the day progressed, clearly benefiting from practice. The time taken for judgments about the articles and to form summaries based on articles both dropped as the day progressed. It is less clear that these changes are the result of practice, although they might be. It is also possible that as the participants became less interested in the articles (as indicated by the interest ratings above), they became less thoughtful and worked faster to complete each trial.

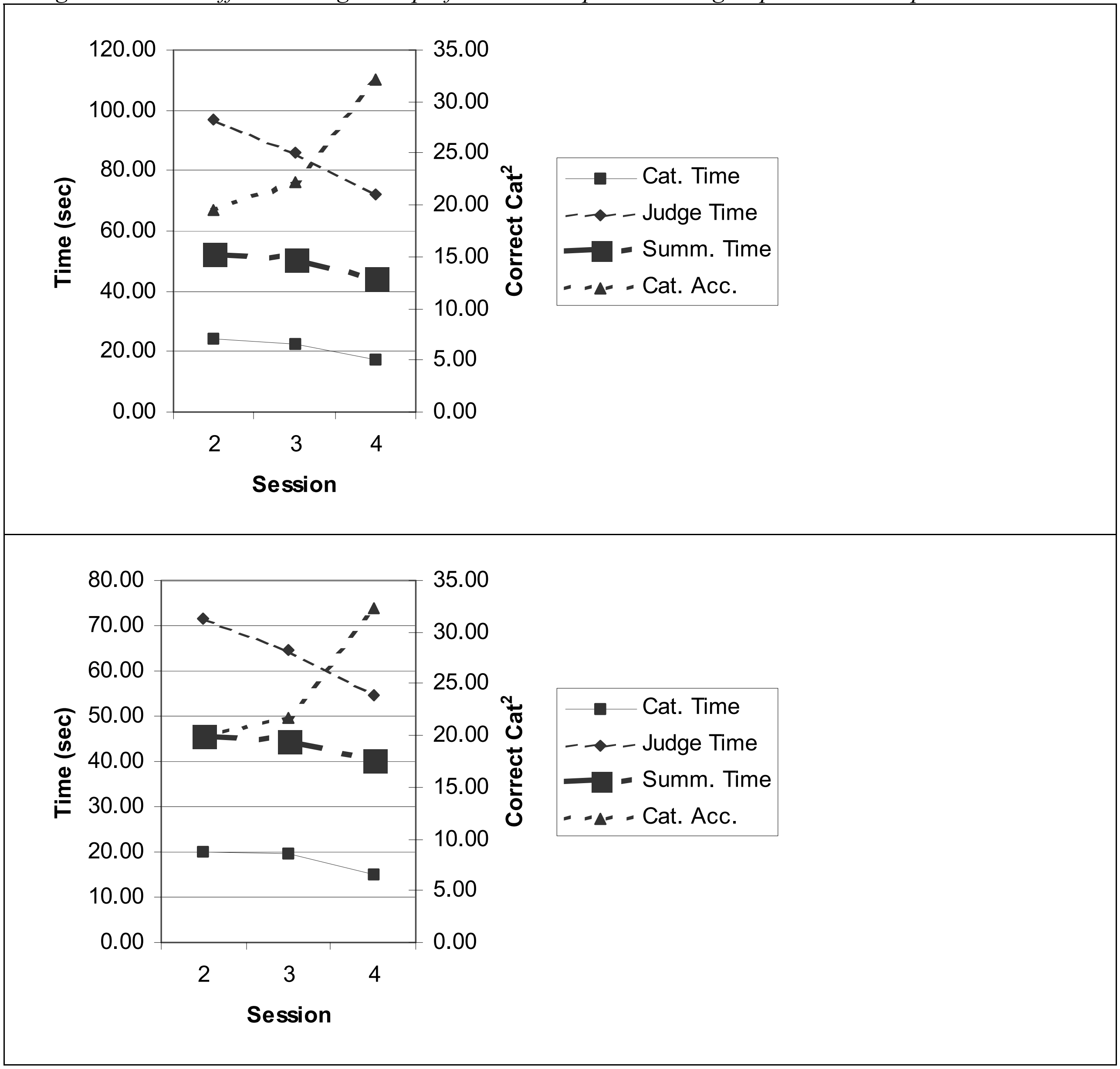
Table 55. Time effects on cognitive performance.

Between-groups design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Linear: Multivariate	0.27	114.86***	4, 170	.36		
Cat. Time		92.81***	1, 173	.35		
Cat. Acc.		349.90***	1, 173	.67		
Judge Time		91.00***	1, 173	.35		
Summ. Time		14.27***	1, 173	.08		
Quadratic: Multivariate	0.79	11.35***	4, 170	.06		
Cat. Time		11.43***	1, 173	.06		
Cat. Acc.		29.29***	1, 173	.14		
Judge Time		1.09	1, 173	.01		
Summ. Time		3.96*	1, 173	.02		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Cat. Time	23.99	13.09	22.46	10.26	17.41	8.55
Cat. Acc.	19.55	7.27	22.29	8.45	32.06	6.23
Judge Time	96.70	41.79	86.14	35.97	72.34	31.96
Summ. Time	52.61	29.20	51.04	33.21	44.50	32.73

Repeated-measures design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Linear: Multivariate	0.18	46.85***	4, 41	.49		
Cat. Time		123.10***	1, 44	.74		
Cat. Acc.		59.95***	1, 44	.58		
Judge Time		38.07***	1, 44	.46		
Summ. Time		9.54***	1, 44	.18		
Quadratic: Multivariate	0.50	10.25***	4, 41	.06		
Cat. Time		20.39***	1, 44	.32		
Cat. Acc.		1.00	1, 44	.02		
Judge Time		7.67**	1, 44	.15		
Summ. Time		1.78	1, 44	.04		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Cat. Time	19.90	10.58	19.81	11.56	15.07	7.10
Cat. Acc.	19.81	7.37	21.64	7.77	32.34	6.17
Judge Time	71.68	35.63	64.65	35.96	54.73	29.28
Summ. Time	45.70	23.27	44.68	27.73	40.21	21.81

Note. *p<.05. **p<.01. ***p<.001.

Figure 34. Time Effects on cognitive performance. Top: Between-groups. Bottom: Repeated Measures



4.4.6.2 Work structure.

Most of the tasks occurred in sets of several trials; for three tasks we have data on the length of time between trials. This is the period of time during which the introductory screen for the new trial appeared on the computer. For the typing and conveyor belt tasks this time amounted to a rest break, as there was no work to be done. For the summarizing task, the time combined a brief rest and the time required to read the article to be summarized in the new trial. In all three cases the measurement is in seconds, and average scores were calculated for each session (across three trials in each session for typing, four for conveyor, and six for summaries). These were transformed to log values to improve the shapes of the distributions. For the between-groups experimental design the statistical model was a 4 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA with the three break durations as dependent variables. For the repeated-measures design the statistical model was a 2 x 3 (Lighting Design X Session) MANOVA.

4.4.6.2.1 Base Case vs Best Practice.

There was no main effect for this planned comparison but there was an interaction of between it and reflectance (i.e., Base Case vs Best Practice X Reflectance). The multivariate effect was associated with a significant univariate effect for the break times in the summarization task. Table 56 summarizes it and Figure 35 illustrates this interaction. The figure and the post hoc tests reveal this interaction as one in which reflectance makes no difference to the Base Case group, but the break times are significantly longer for Best Practice people with dark partitions rather than light partitions. These breaks included time to read the articles to be summarized, which were printed on paper and were either on the desk beside the computer, or on the document holder.

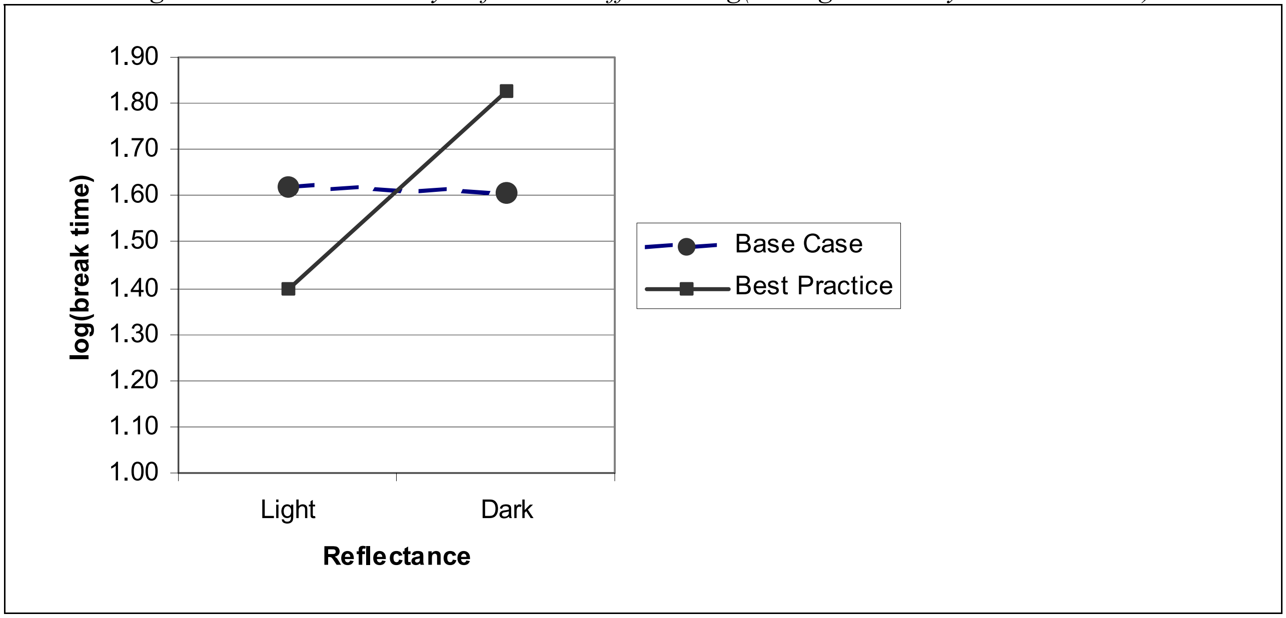
Table 56. Base Case vs Best Practice by Reflectance interaction effect on log(break durations).

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	
Multivariate	0.95	3.09	3, 171	.02	
Typing Breaks		0.18	1, 173	.00	
Conveyor Breaks		1.69	1, 173	.01	
Summary Breaks		6.30**	1, 173	.04	

	Light *M*	Light *SD*	Dark *M*	Dark *SD*	Reflectance *post hoc tests*
Base Case	1.62	0.45	1.61	0.53	$F(1, 49) = 0.01$
Best Practice	1.40	0.65	1.83	0.19	$F(1, 39) = 7.83$** $\eta^2_{partial} = .17$
Lighting Design *post hoc* tests	$F(1, 50) = 2.51$		$F(1, 38) = 3.19$		

Note. *p<.05. **p<.01. ***p<.001. Group descriptive statistics and *post hoc* tests are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

Figure 35. Base Case vs by Reflectance effect on log(average summary break duration).



4.4.6.2.2 Best Practice vs Best Practice + Switching Control.

This planned comparison showed a statistically significant multivariate main effect of lighting design on break times, with a significant univariate difference for the typing task breaks (Table 57). These breaks were shorter for the people with Switching Control than for the people in the Best Practice condition.

Table 57. Best Practice vs Switching Control effect on log(average break duration)

					Best Practice		Switching Control	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	0.95	3.31*	3, 171	.02				
Typing		5.30*	1, 173	.03	1.23	0.28	1.13	0.24
Conveyor		0.81	1, 173	.00				
Summaries		2.72	1, 173	.02				

Note. *p<.05. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

4.4.6.2.3 Non-lighting effects.

Both the between-groups and repeated-measures experimental designs showed large, statistically significant time-linear main effects (Table 58). In both cases all three break durations became shorter as the day progressed (i.e., from session 2 through Session 4). Note that the times are different for the three tasks in the expected direction: shortest for conveyor belt; intermediate for typing, when people were sorting the source text pages; and, longest for summaries, when they also read the source article. In general this is likely to be a practice effect, in which people move on more quickly as they become more experienced with the tasks. It is interesting to note that there was no quadratic component, as might have occurred if the effects of fatigue had offset practice. Perhaps participants kept moving on the tasks to offset boredom or decreasing interest.

Table 58. Time effects on log(average break duration).

Between-groups design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	
Linear: Multivariate	0.55	46.62***	3, 171	.24	
Typing		49.55***	1, 173	.22	
Conveyor		49.19***	1, 173	.22	
Summaries		65.40***	1, 173	.27	

	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Typing	1.27	0.24	1.19	0.23	1.10	0.25
Conveyor	0.70	0.27	0.61	0.32	0.52	0.33
Summaries	1.75	0.40	1.65	0.48	1.51	0.52

Repeated-measures design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Linear: Multivariate	.37	22.89***	3, 41	.34
Typing		13.47***	1, 43	.24
Conveyor		25.66***	1, 43	.38
Summaries		7.32***	1, 43	.40

	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Typing	1.16	0.28	1.12	0.27	1.02	0.28
Conveyor	0.68	0.27	0.57	0.35	0.47	0.33
Summaries	1.55	0.47	1.49	0.50	1.34	0.55

Note. ***p<.001.

4.4.7 Summary: Performance

The results for the various performance measures may be summarized thus:

- **Visibility, practice, and fatigue effects were found as expected.** The overall effect of target contrast on composite visual performance conformed to our knowledge of the visual system, as did the effect of print size on transcription typing performance. The tasks were novel at the start of the day, and showed performance improvements as the day went on. This is most clear in the speed and accuracy of the categorization task, and in the typing task, but also in the composite visual performance measure. Some of these linear effects were tempered over time as various processes interacted, such as declining interest leading to changes in the time taken to complete cognitive judgments, adding a quadratic shape to the time effects.

- **Indirect systems interact with partition reflectance.** There were few statistically significant effects involving the Base Case vs Best Practice comparison, and most were interactions with partition reflectance. These vary in their implications. The timed vision test showed that performance improvements over the day (which reflect learning this novel task) were smallest for people in the Best Practice design with dark blue partitions. The same group also took longer breaks between trials in the summarization task than Best Practice participants with light gray partitions. For the wholly computer-based task (the timed vision test) it appears that the combination of indirect lighting and a dark background might be detrimental, although the effect is small and did not occur on other similar tasks (e.g., the vigilance measures in the conveyor belt task). The meaning of the break durations for the summarization task is ambiguous. Longer breaks might reflect more time reading the article on paper before doing the summarization on the computer, which could mean a stronger task focus that might be desirable.

- **Individual control over lighting has performance benefits.** The several effects involving the comparisons of Best Practice with either Switching Control or Dimming Control are generally consistent in their directions. Dimming control participants showed steeper performance improvements over increasing contrast in the timed vision task and avoided motivation declines over the day. They also improved in vigilance performance over the day, whereas the Best Practice participants did not. There was additional evidence in interaction effects with Print size and time that typing performance also showed beneficial effects of having dimming control.

The results for the planned comparisons of light distribution are in keeping with other tests of lighting design effects on office work performance. Veitch and Newsham (1998a) found few performance effects involving planned comparisons between direct and indirect lighting systems, and the effects they observed were of comparable size to the ones reported here. Similarly, Eklund, et al. (2000) found no effects of light distribution on clerical work (data entry and proofreading).

These results are novel in their finding of beneficial effects of individual control on performance. Boyce, Eklund and Simpson (2000) found no clerical work performance effects associated with having control over lighting, but did not use tasks similar to those that in this experiment showed effects. Veitch and Newsham (2000a) did not demonstrate lighting control effects on performance using a matched design in which participants with control and those without experienced the same visual conditions. One unusual aspect of their methodology was that participants with control could not change the settings later during the day; their control was limited to the initial settings at the start of the day's work. In the present experiment, participants in the control conditions were free to alter the lighting at any time; even if they did not do so, they would have been always aware of the possibility to do so. Further exploration of the effects of control appears below, in analyses focusing on the participants in the Switching Control and Dimming Control groups specifically.

4.5 Social Behavior

4.5.1 Liking

Participants completed a five-question scale of liking for their fellow participants after spending time in a group discussion task. These were averaged to provide one score, on a scale of –3 through +3, which showed acceptable internal consistency reliability ($\alpha = .66$). The overall mean was 1.34 (*SD* = 0.76), indicating that participants in general felt somewhat positively towards their fellow employees.

Data from this scale score was analyzed in a 4 x 2 (Lighting Design by Reflectance) ANOVA for the between-groups experimental design, and a one-way ANOVA comparing two lighting conditions for the repeated measures design. There were no differences in either analysis. Liking for other group members was unrelated to lighting conditions.

4.5.2 Conflict Resolution

The conflict resolution task provided two types of scores, five ratings and five rankings. The ratings were the average (over five scenarios) likelihood that an individual would resolve conflicts using competition, collaboration, compromise, avoidance, or accommodation strategies, on a scale from –3 through +3, where higher scores indicated greater likelihood of using that strategy. The rankings were the average rank assigned to each of the five strategies over the various scenarios. Table 59 summarizes the characteristics of these two sets of variables. For some of the scales, internal consistency reliability is acceptable, but for others it is unacceptably low. Overall, this task requires further development and refinement as a measure of these tendencies.

Table 59. Conflict resolution scale characteristics

	Cronbach's alpha	M	SD
Ratings			
Competition	.46	0.26	0.91
Accommodation	.36	0.35	0.78
Avoidance	.37	-1.23	0.77
Collaboration	.58	1.33	0.79
Compromise	.38	1.78	0.62
Rankings			
Competition	.48	3.46	0.69
Accommodation	.16	3.23	0.51
Avoidance	.50	4.30	0.55
Collaboration	.38	2.16	0.56
Compromise	.20	1.86	0.50

Note. Each scale is the average score for that concept rated on 5 conflict scenarios. Ratings indicate the absolute likelihood that the individual will use that strategy to resolve conflict, and are on a scale from –3 through +3. Rankings indicate the relative likelihood of using that strategy in relation to the others, and are on a scale from 1 through 5.

The ratings and the rankings were analyzed in separate MANOVA models (e.g., either 5 ratings or 5 rankings were the dependent variables) for the between-groups and repeated measures designs. As usual, the between-groups design was a 4 x 2 (Lighting Design by Reflectance) MANOVA and the repeated measures design was a one-way MANOVA comparing two lighting conditions. There were no statistically significant effects in these analyses.

4.5.3 Willingness to Volunteer

At the end of the day, participants were asked two questions concerning their willingness to volunteer. One question concerned the number of unpaid hours (0 through 10) that the individual would be willing to put in at home to complete questionnaires as part of other research. The other question concerned their willingness to return for a second session if requested (a probability, 0 through 100).

These two scores were standardized and then averaged to form an overall score for their willingness to volunteer. As one would expect for a standardized variable, the mean was 0.00 and the standard deviation was 0.87 (i.e., close to 1). This single score was analyzed in a 4 x 2 (Lighting Design by Reflectance) ANOVA for the between-groups experimental design, and a one-way ANOVA comparing two lighting conditions for the repeated measures design. Neither analysis showed any effects of lighting conditions on willingness to volunteer.

4.5.4 Summary: Social Behavior

Social behaviors had been expected to show differences associated with positive affect (Baron, Rea, & Daniels, 1992). People in a state of positive affect should rate other people more positively, resolve conflicts with more collaborative or compromising strategies, and be more willing to volunteer. It appears that, just as there were no simple effects of lighting design on mood (pleasure), these lighting conditions were not different enough to result in changes on the measures of social behavior.

4.6 Lighting Choices

This section concerns only the participants who had one or another form of control: those in the Best Practice + Switching Control condition, and those in the Dimming Control condition. The latter group includes those who experienced Dimming Control on either their first or second visit (the second visit being for those people from the Base Case condition who also participated in the repeated-measures design). Lighting choice data were lost from some of the Dimming Control participants because of equipment failures, so that the total sample size for that group in this section is somewhat smaller than was available for the analyses reported above. Data from the two groups were kept separate because the two forms of control are very different; Switching Control only allowed the addition of light to the ambient level, and in one area of the workstation; Dimming Control allowed the user to increase and decrease the amount of light and, furthermore, affected the whole workstation.

We cleaned the data to ensure that only meaningful lighting choices were included. Events that occurred before the participants arrived, or after they had left, were deleted. Furthermore, when there occurred a rapid series of changes over a two-minute period, we retained only the last setting. We assumed that the rapid series of changes represented an attempt by the participant to hone in on the desired setting, represented by the final value.

4.6.1 Actions and Lighting Conditions

4.6.1.1 Frequency of use.

We examined the number of times that participants used the lighting controls available to them, as well as when they occurred (Table 60). The patterns are somewhat similar for the two forms of control. Some people never changed the setting, and most who did change the lighting did so only once, at the start of the day. The time of the first switching or dimming action was never later than 9:15 a.m., and the latest was at approximately 4:00 p.m. Given the low frequency of lighting changes, there was no detectable pattern that people changed settings in response to task demands.

Table 60. Frequency of switching or dimming during the day.

	0	1	2	3	4	5	6	7	Total
Switching									
N	11	12	3	2	2	1	1	1	33
%	33	36	9	6	6	3	3	3	
Dimming									
N	6	36	6	3	4	2			57
%	11	63	11	5	7	4			

For the further analyses we focused on the lighting conditions that prevailed during Session 4. We did so on the basis that these choices would be the most stable indicators of lighting preferences. In determining the prevailing Session 4 conditions, we looked for the conditions that were in place for most of the session (after 2:30 p.m.), and excluded any changes made less than 30 minutes before the participant completed the day. (There were several changes near the end of the day, which appeared to co-occur with a questionnaire about the ease of making lighting changes.) We also checked on when the Session 4 conditions had been chosen (Table 61). The lighting conditions prevailing in Session 4 were mostly chosen in Session 1.

Table 61. Session in which Session 4 lighting conditions had been set

	1	2	3	4	Total
Switching					
N	26	2	3	2	33
%	79	6	9	6	
Dimming					
N	47	3	5	2	57
%	82	5	9	4	

Note. Session 1, up to morning coffee break (about 10:30). Session 2, between break and lunch (10:30 – noon). Session 3, before afternoon coffee break (noon – 2:30). Session 4, after 2:30 p.m. but more than 30 minutes before the participant ended the day.

4.6.1.2 Chosen luminous conditions.

We characterized the lighting choices using data on the switching or dimming settings, several spot illuminance measurements, and two average illuminances. These are compiled in Table 62 and illustrated in Figure 36. The differences between the two strategies are clear in the conditions that were created. When given the opportunity to reduce ambient lighting levels below the initial design level, some people do so.[1] They had no possibility to do so with the Switching condition, which only allowed addition of light to the ambient level. However, some participants in the Switching condition did reduce the output of the task lamp from the initial, mid-range setting. Because of the location and type of desk lamp, it had the largest effect on the average horizontal illuminance (see Figure 36), and less effect on the vertical illuminance on the document holder. Looking to the illuminance on the face of a cube in the location of a seated occupant, we see that although the median and mean values are higher for the Switching Control condition, the range of achieved values is wider for the Dimming Control condition, with a lower minimum and a higher maximum.

[1] Only the maximum (100%) and minimum (0%) values in Table 62 have direct comparisons to design values presented in Table 3. In comparing the average chosen luminous conditions to the design values in section 3.5.1, bear in mind that the average chosen conditions are averaged across the workstations that were occupied (resulting in unequal numbers of each workstation in the sample). However, the averages for the design values are averaged across the nine workstations equally. The difference in the frequency of each workstation in the final sample accounts for differences between the achieved illuminance and the design illuminance statistics.

Table 62. Session 4 lighting choices.

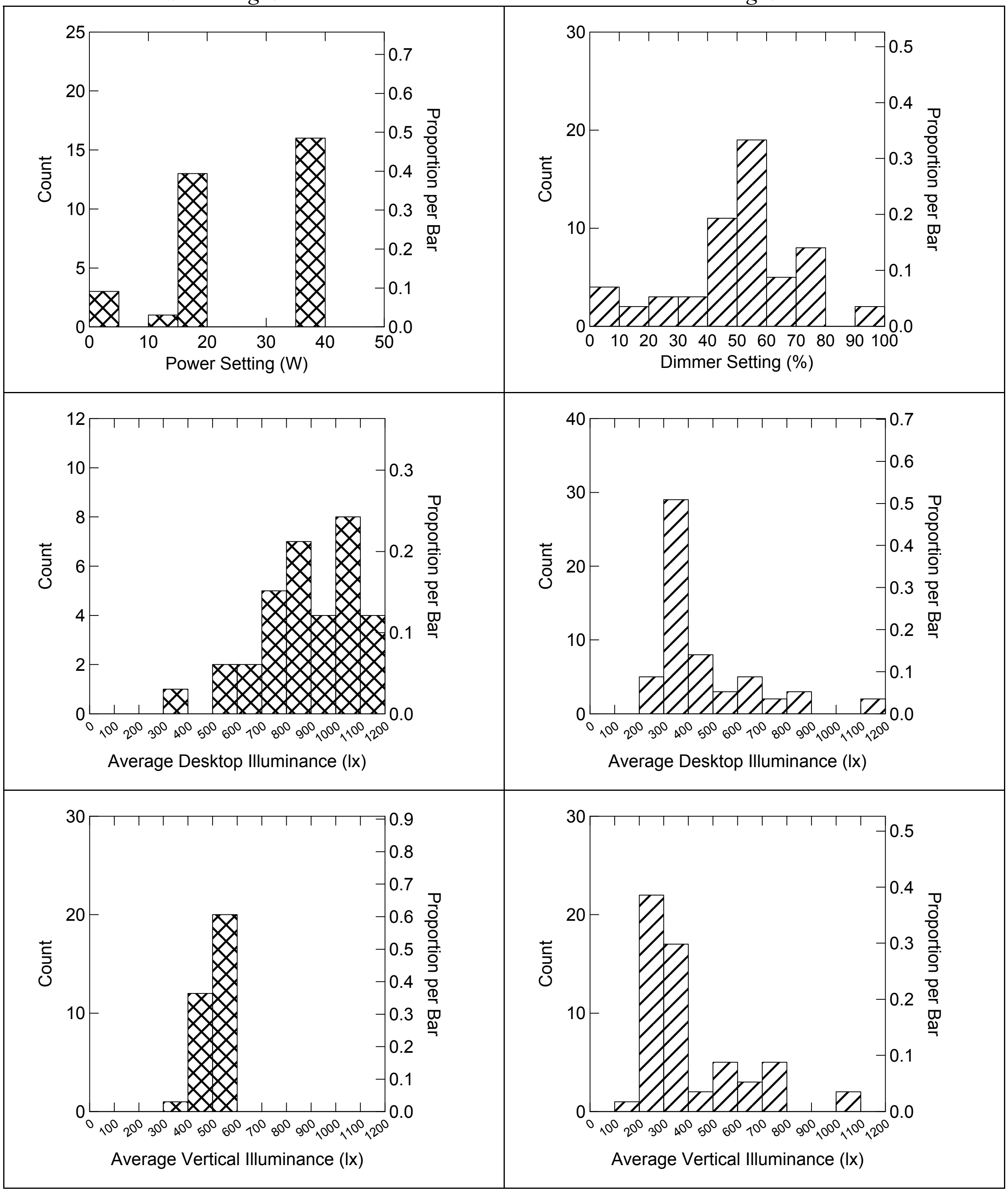
	Power (W)	Open Desk	Keyboard	Keyboard-Monitor	Monitor (vertical)	Face of Occupant
Switching						
Min	0	351	422	453	246	234
Max	35	2497	812	736	545	459
Median	19	1679	684	614	416	338
M	25	1701	691	616	408	349
SD	11	611	85	70	80	51
Dimming	Setting (%)	Open Desk	Keyboard	Keyboard-Monitor	Monitor (vertical)	Face of Occupant
Min	0	243	248	242	146	144
Max	100	1075	1279	1341	853	604
Median	51	384	390	394	258	251
M	50	435	485	489	327	278
SD	22	171	230	235	175	99

Switching	Document Holder – L	Document Holder – R	Under Bin	Average Desktop	Average Vertical
Min	345	366	371	399	356
Max	636	582	649	1131	585
Median	513	506	533	869	510
M	510	513	529	884	512
SD	66	48	74	187	54
Dimming	Document Holder – L	Document Holder – R	Under Bin	Average Desktop	Average Vertical
Min	185	179	257	252	191
Max	1129	1115	1010	1176	10900
Median	326	310	354	375	3110
M	416	392	424	458	404
SD	218	205	176	201	210

Note. Power is in watts. Dimmer setting is in %. All other values are illuminances in lux. Average Desktop illuminance is the average of the illuminances on the open desk, the keyboard, between the monitor and document holder, and under the binder bin. Average vertical illuminance is the average of the illuminances on the document holder in the right and left-hand positions.

Figure 36. Lighting conditions created by switching and dimming choices

Switching Control *Dimming Control*



Within each group (Switching and Dimming), we tested for partition reflectance differences in average desktop illuminance and average vertical illuminance. The conditions achieved in the two groups (light and dark) were not different. Therefore, only the aggregate data are presented here.

Among the most perplexing questions for behavioral lighting research is the choice of which measurement of the luminous conditions might be the best predictor of how people respond. To help us make that choice, we also examined how the various lighting measurements correlated (Table 63). The lower half of the table shows the data for the Switching Control condition. Power is relatively weakly related to the luminous conditions because the dominant source was the overhead Best Practice direct / indirect system. The correlations of Power with the vertical measurements are low, as expected. The desk lamp made the biggest contribution to the illuminance on the open desk, where it was located, as seen in the r=.97 between Power and the open desk measurement.

The upper half of the table shows the results for the Dimming Control choices. Setting is not perfectly related to the lighting conditions because there was always a fixed indirect component to provide minimal ambient illumination. Interestingly the values are very highly intercorrelated, some of them perfect correlations. The correlations are somewhat higher for this condition than for the Switching Control condition, and the difference serves as a reminder that these relationships for any given space will always depend on the lighting equipment. For the present experiment, these results led to the choice of average desktop illuminance as the value to use in analyses seeking to relate luminous conditions to behavioral outcomes, the topic of the next section.

Table 63. Intercorrelations among luminous conditions.

	Power (W) [Switch]	Open Desk	Keybd	Keybd-Monitor	Monitor (vert)	Face	Doc Holder – L	Doc Holder – R	Under Bin	Ave Vertical	Ave Desktop
Setting (%) [Dim]		0.80	0.82	0.83	0.81	0.78	0.80	0.81	0.80	0.81	0.82
Open Desk	0.97		0.98	0.97	0.95	0.89	0.94	0.97	0.95	0.96	0.98
Keybd	0.70	0.70		1.0	0.99	0.90	0.99	0.99	0.98	0.99	1.0
Keybd-Monitor	0.79	0.77	0.88		0.98	0.89	0.99	0.99	0.98	0.99	1.0
Monitor (vert)	0.45	0.44	0.82	0.72		0.87	0.99	0.97	0.96	0.99	0.98
Face	0.49	0.51	0.74	0.75	0.56		0.88	0.85	0.88	0.87	0.90
Doc Holder - L	0.43	0.47	0.88	0.75	0.72	0.60		0.97	0.97	0.99	0.98
Doc Holder - R	0.71	0.68	0.89	0.88	0.86	0.68	0.77		0.97	0.99	0.99
Under Bin	0.10	0.09	0.64	0.35	0.69	0.60	0.59	0.54		0.98	0.99
Ave Vertical	0.58	0.59	0.94	0.85	0.83	0.67	0.96	0.92	0.61		0.99
Ave Desktop	0.95	0.98	0.83	0.86	0.59	0.63	0.61	0.79	0.28	0.73	

Note. Top (shaded) half shows intercorrelations for the Dimming Control group. Bottom half (unshaded) shows intercorrelations for the Switching Control group.

4.6.2 Illuminance Effects on Behavioral Outcomes

We hypothesized that lighting levels might predict behavioral outcomes. In particular we expected that illuminance (either horizontal or vertical) would influence performance on visual tasks, such as transcription typing with a small (8-pt) print size. For these analyses we used multiple regression. For the Switching Control group, average desktop illuminance was the only predictor (independent) variable. For the Dimming Control group, we also entered a variable called visit, which marked whether the participant had experienced that condition on their first day, or their second (the latter occurred for the Base Case participants who repeated). The purpose of this was to control for any effects of practice or knowledge associated with having done the experiment previously.

We repeated the regression analyses for the same conceptual groups of dependent variables that were presented above, but only for the Session 4 scores. As previously, we required statistically significant multivariate effects for groups of variables, prior to interpreting univariate effects. Thus, for the performance measures there were tests for the timed vision test (6 contrasts in session 4); motivation (1 variable); typing task (3 print sizes in session 4); computer-based cognitive judgments (4 variables); cognitive performance (4 variables); vigilance (2 variables); and break durations (3 variables). None of these analyses showed any statistically significant effects of illuminance on performance for either Switching Control or Dimming Control.

For the perceptions and feelings measurements, we examined comfort (2 variables), lighting appraisals (2 lighting quality ratings), preference (4 room appearance ratings), satisfaction (3 ratings), and mood (3 scales). All of these outcomes were data provided in session 4. There were no statistically significant effects on any of these tests.

In addition, we tested two specific hypotheses using other illuminance predictors. We examined the effect of vertical illuminance on typing performance, and the effect of illuminance on the face on visual and physical comfort. These regressions were also not statistically significant for either the Switching Control or Dimming Control groups.

4.6.3 Switching/Dimming Frequency Effects

We hypothesized that people who used their controls more frequently might have lower scores on perceptions and feelings; that is, that they might use the controls to try to overcome problems with the lighting. The dependent variables were mood (three variables), lighting appraisals (one lighting quality rating), preferences (4 room appearance ratings), comfort (two ratings), and environmental competence (one rating). For the Dimming Control group we also controlled for visit. Frequency of switching or dimming did not predict any of these dependent variables for either group.

4.6.4 Satisfaction with Control Mechanism

Participants in the Switching Control and Dimming Control groups both completed an extra questionnaire concerning their satisfaction with the type of lighting control they had had. The responses are degree of agreement with various statements, where 0 = strongly disagree and 4 = strongly agree. Both the descriptive statistics and the test of between-groups differences are of interest. Table 64 summarizes responses for the two groups and includes the significance test between them. This table includes only participants from the between-groups experimental design, which keeps a more equal number of participants in the two groups and avoids any confounds associated with having participated more than once before completing this questionnaire.

Although in general the means all indicate some agreement with each statement (i.e., a degree of satisfaction with the lighting controls), the two groups differed significantly. In particular, the Switching Control participants showed lower agreement with the statement “The lighting control system allowed me to create the lighting conditions I wanted.” They also thought that the instructions for use were poorer, that the interface was less easy to use, and that the change did not occur rapidly enough in response to their commands. This is intriguing given that the interface in question was a three-level switch on the desktop. The Dimming Control interface was on the computer and might have been expected to be a more complex way to control lights, although it did offer continuous control rather than step increments.

Table 64. Satisfaction with the control interface.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Switching Control		Dimming Control	
Multivariate	.88	2.32*	5, 82	.06	M	SD	M	SD
The lighting control system allowed me to create the lighting conditions I wanted.		10.02**		.10	2.27	0.94	2.98	1.02
The lighting control system was easy to use.		3.66	1, 86	.04	2.76	0.79	3.13	0.92
The instructions on how to operate the lighting controls were adequate.		4.05*	1, 86	.05	2.70	0.95	3.13	0.97
The interface for the lighting controls was easy to use.		5.43*	1, 86	.06	2.81	0.81	3.21	0.76
When I used the lighting controls, the lighting changed rapidly enough in response to my commands.		4.68*	1, 86	.05	2.70	0.81	3.11	0.90

4.6.5 Summary: Switching and Dimming Control

Overall, participants used the controls in ways consistent with previous laboratory and field research: rarely, and mostly at the beginning of the day (e.g., Maniccia, Rutledge, Rea, & Morrow, 1999; Moore, Carter, & Slater, 2002). The range of illuminances in the Dimming Control condition has an average lower than the designed average, which both suggests the possibility of energy savings associated with individual control and is consistent with other studies (Boyce, et al., 2000; Veitch & Newsham, 2000b). These differences in illuminance did not predict performance, perceptions, or feelings. Although disappointing, this is not surprising given that none of the conditions substantially impaired the visibility of the high-contrast, sharply printed paper material, and the fact that the computer tasks were self-luminous and had high contrasts and large size.

The means of controlling the lighting did give rise to different responses. The Dimming Control condition was more satisfactory than the Switching Control condition. It seems likely that this is because it allowed the creation of a wider range of possible conditions, although the novelty effect of the computer interface cannot be ruled out.

4.7 Linked Mechanisms Tests

4.7.1 Analytic Strategy

We had initially planned to test the relationships in the linked mechanisms map (Figure 37, and 1) using structural equation modeling (SEM) (Kline, 1998). SEM is a statistical technique for solving sets of simultaneous linear regression relationships and including both the model of the manifest (measured) variables and the latent variables or constructs that give rise to the measurements. A simpler form of this is path analysis, which involves only the relationships between the manifest variables. However, both techniques are very sensitive to the violation of fundamental assumptions, particularly deviations from normal distributions. Several of our variables were not normally distributed. In addition, we had a very large model with many variables, which requires a very large number of participants for a robust test. Overall, it seemed unlikely that any meaningful results could be obtained from this approach.

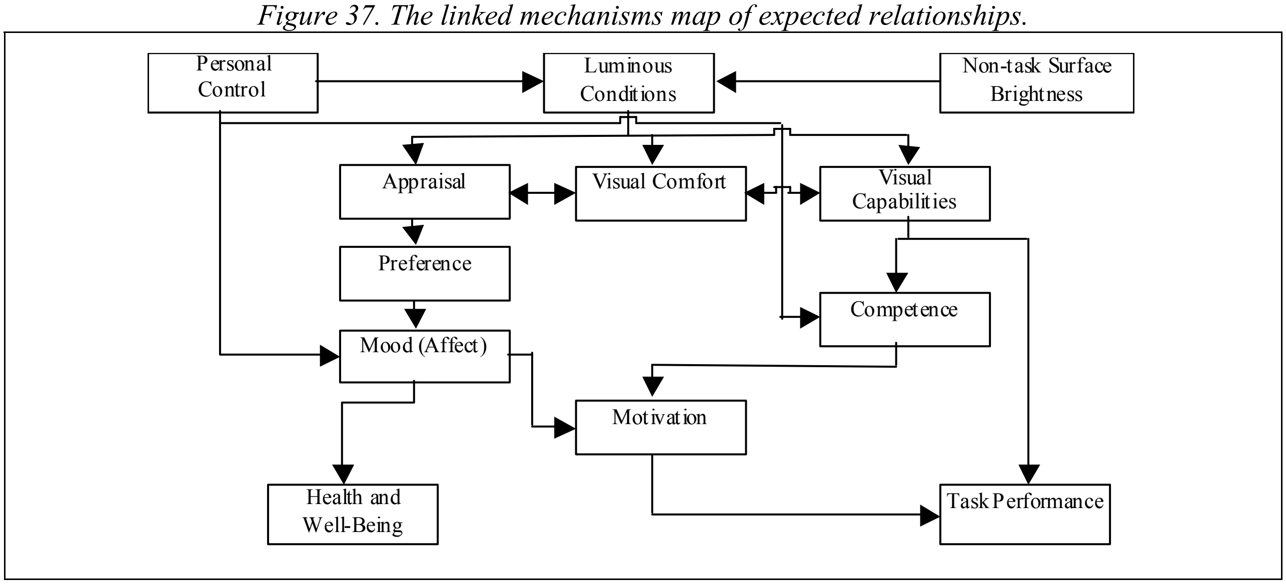
Instead, we decided to conduct a series of tests of each link using non-randomly formed groups in an analysis of variance approach. Thus, for each antecedent variable we formed two groups based on a split around the median (50^{th} percentile) value. We determined that the distribution of people into the above-and below-median groups was independent of lighting condition, using chi-squared tests of independence. The median split groups were equally distributed across lighting conditions. This is not surprising, given the results reported above.

We then tested the effect of being in the high or low group, on the outcomes linked to the consequent variable at that link. For example, do people who rate the lighting as being of higher quality rate the room as being more attractive? Do the people with the highest pleasure scores report greater environmental satisfaction? These tests were conducted using only data from session 4, when there had been the greatest exposure to the experimental conditions. We conducted separate tests for data from the between-groups experimental design (designated BG) and from the second visit of the repeated-measures design (designated RM_V2). (Data from the first visit had been included in the between-groups analysis).

There are important caveats to this approach. Although we used analysis of variance to test the links, causal inferences are weak at best. We are able to argue that there are meaningful associations between variables on the basis of these significant tests, but because we did not randomly assign individuals to the above- and below-median groups on each variable we cannot assert that we have controlled for all the reasonable extraneous variables. For example, it is possible that people who reported higher pleasure might be higher in conscientiousness, a personality trait, and might have reported greater environmental satisfaction in order to fulfill the experimenters' expectations as they understood them. The personality variable, which we did not measure, would in that case be the causal agent, rather than our explanation of mood effects.

In addition, the second visit of the repeated-measures group is an imperfect replication for the between-groups data because the sample size is smaller (which reduces the statistical power considerably, making it less likely that real effects might be detected) and because the RM_V2 sample is a non-random subset of the people in the BG group.

Figure 37. The linked mechanisms map of expected relationships.



The links at the top level of the path are logically, rather than statistically, established. The four lighting designs and variations in task conditions (e.g., print size or luminance contrast) created a variety of luminous conditions, which were further varied for participants with Switching or Dimming control and by the use of two partition reflectances. Furthermore, we had already tested the links between luminous conditions, and appraisal, visual comfort, and visual capability. It was clear that the luminous conditions did result in discriminating appraisals (the LRC Office Lighting Survey results) and that luminous conditions did result in differential visual performance (variations in typing performance by print size, and in timed vision test performance by contrast level). We had not demonstrated that our range of luminous conditions influenced visual comfort, but this relationship has been well-established previously with other lighting systems; indeed, this is the basis for predictive tools for discomfort glare such as Visual Comfort Probability and the Unified Glare Rating (CIE, 1995; Committee on Recommendations, 1966). We did not test the relationships at this level any further.

The relationships at the next level down were the focus of our analyses. As is evident from Figure 37, broadly speaking there are two parallel paths through the map, one leading to through appraisal (the 'Appraisal Path') and the other through vision (the 'Vision Path'). Although this structure is an oversimplification, we use it here to organize the results for easier comprehension.

4.7.2 Appraisal Path

4.7.2.1 Median-split groups.

The first step in the analysis was the formation of non-random groups for each variable. These groups were used as two levels of one independent variable when that variable served as the antecedent part of a link. When a given conceptual group served as the consequent part of a link, we used the same dependent variables as had been used for the previously-reported analyses. Thus, for some models we examined ANOVA tests with one dependent variable, if there had only been one (e.g., for motivation), and in other cases we used MANOVA tests with multiple dependent variables.

Table 65 describes the median-split groups for both the between-groups and repeated-measures sets. As is evident, it was in some instances difficult to form groups that were equal in size. This was particularly true for the more skewed variables. When a large number of cases fell exactly on the median, separating them was impossible. We tried various options in these cases and chose the one that gave the closest to an even split. The total number of cases varies for each variable because of missing data.

Table 65. Median split groups for appraisal path analyses.

Variable		Low Group			High Group		
Set	Split	N	*M*	*SD*	N	*M*	*SD*
Appraisal (Lighting Quality)							
BG	<3 or >=3	63	1.93	0.67	117	3.44	.40
RM_V2	< 3.25 or >= 3.25	19	2.30	0.75	26	3.75	0.33
Preference (Attractiveness)*							
BG	< 48.89 or => 48.89	92	36.20	10.47	89	62.18	11.75
RM_V2	< 51.78 or => 51.78	23	40.57	10.58	22	65.03	12.08
Mood (Pleasure)							
BG	< 4.17 or => 4.17	89	3.03	0.91	81	5.74	1.01
RM_V2	< 4.67 or => 4.67	22	3.49	0.92	20	6.38	1.11
Mood (Arousal)							
BG	< 3.33 or => 3.33	74	2.32	0.71	96	4.27	0.91
RM_V2	< 3.67 or => 3.67	24	2.57	0.80	18	4.54	0.77
Mood (Dominance)							
BG	< 4 or => 4	71	3.28	0.62	105	4.63	0.73
RM_V2	< 4 or => 4	16	3.29	0.46	27	4.66	0.73
Motivation (Spd. Random Resp.) **							
BG	< 729 or = 729	58	336.90	205.04	123	729	0
RM_V2	< 729 or = 729	16	354.97	212.67	29	729	0
Environmental Competence							
BG	< 6 or => 6	76	4.81	1.00	105	6.65	0.66
RM_V2	< 6 or => 6	17	4.88	0.46	28	6.98	0.78

Note. BG = between-groups data (N=181). RM_V2 = repeated-measures data from second visit (N=45). * Attractiveness was chosen as the only Preference variable for this analysis because it was the strongest of the four components; in the principal components analysis it explained the largest percentage of variance, and its scale scores showed high internal consistency reliability. ** The maximum possible value for this variable was 729.

4.7.2.2 Appraisal → Preference link.

The first link to be tested was between Appraisal and Preference, with its four dependent variables. Table 66 shows the summary for this analysis. Appraisal, represented by low and high lighting quality ratings, was related to preference in both sets of data. Higher appraisal values (above-median lighting quality ratings) are associated with higher ratings of Attractiveness, Comfort, Visibility and Spaciousness in the BG sample and higher Comfort and Visibility in the RM_V2 sample. The effects are small to medium in size.

Table 66. Test of Appraisal → Preference link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Appraisal → Preference	.80	11.11***	4, 175	.10	.78	2.88*	4, 40	.08
Attractiveness		33.77***	1, 178	.16		0.63	1, 43	.01
Comfort		21.38***	1, 178	.11		7.95**	1, 43	.16
Visibility		13.18***	1, 178	.07		6.74**	1, 43	.14
Spaciousness		13.69***	1, 178	.07		1.14	1, 43	.03
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Attractiveness	39.77	13.62	54.06	16.76				
Comfort	59.55	14.53	69.30	12.91	65.12	13.76	75.40	10.73
Visibility	57.51	13.67	65.84	15.22	58.17	10.98	70.23	17.53
Spaciousness	57.08	15.48	65.58	14.26				

Note. ****p*<.001. ***p*<.01. **p*<.05.

4.7.2.3 Preference → Mood link.

Next we tested the link from preference to mood, with its three dependent variables (Table 67). Here too we found statistically significant effects in which higher preference (above-median ratings of attractiveness) was associated with higher scores on all three aspects of mood: pleasure, arousal, and dominance, for both sets of data. The effect sizes were large for the RM_V2 sample and small to medium for the BG sample.

Table 67. Test of Preference → Mood link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Preference→ Mood	.88	7.85***	3, 166	.07	.67	6.21**	3, 38	.18
Pleasure		17.66***	1, 168	.10		11.61**	1, 40	.22
Arousal		9.80***	1, 168	.06		11.51**	1, 40	.22
Dominance		11,62***	1, 168	.06		4.64*	1, 40	.10
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pleasure	3.83	1.61	4.85	1.55	4.07	1.53	5.74	1.64
Arousal	3.14	1.19	3.73	1.19	2.86	1.10	4.04	1.14
Dominance	3.86	0.94	4.33	0.85	3.86	0.84	4.46	0.95

Note. ****p*<.001. ***p*<.01. **p*<.05.

4.7.2.4 Mood → Health and Well-being link.

We tested median split groups for all three mood measures in examining the relationship between mood and health and well-being outcomes. For health and well-being there are three separate conceptual groups: health outcomes (visual and physical comfort), satisfaction outcomes (environmental satisfaction, performance satisfaction, and self-rated productivity), and task difficulty ratings (7 separate ratings). Table 68 summarizes the results for the health outcomes. For both sets, pleasure was associated with health outcomes, with more and stronger symptoms reported among people with lower pleasure. These effects are medium-sized. Smaller effects but in the same direction were found in the BG group for both arousal and dominance, but these effects did not occur in the RM_V2 group.

Table 68. Tests of Mood → Health link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Pleasure) → Health	.85	14.58***	2, 167	.12	.75	6.64**	2, 39	.14
Visual Comfort		18.00***	1, 168	.10		2.52	1, 40	.06
Physical Comfort		27.45***	1, 168	.14		11.45**	1, 40	.22
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Visual Comfort	6.35	5.59	3.25	3.63	5.05	7.65	2.15	3.01
Physical Comfort	6.19	4.73	2.85	3.40	6.68	6.36	1.55	2.44
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Arousal) → Health	.96	3.92*	2, 167	.04	.94	1.33	2, 39	
Visual Comfort		7.47**	1, 168	.04				
Physical Comfort		4.97*	1, 168	.03				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Visual Comfort	6.55	6.61	3.97	4.34				
Physical Comfort	5.55	5.47	3.94	4.55				
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Dominance) → Health	.96	3.30	2, 173	.03	.93	1.45	2, 40	
Visual Comfort		6.41**	1, 174	.04				
Physical Comfort		3.80*	1, 174	.02				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Visual Comfort	6.03	5.14	4.10	4.81				
Physical Comfort	5.30	4.61	3.98	4.23				

Note. ***$p<.001$. **$p<.01$. *$p<.05$.

Table 69 summarizes the results for the link between mood and measures of satisfaction. Pleasure was a statistically significant predictor of all three measures of satisfaction. In all cases, satisfaction was higher for the above-median pleasure group. The effects were large for the repeated measures set and small to medium for the between-groups set. Arousal and dominance were not significant in multivariate tests, so we did not interpret significant univariate tests.

Table 69. Tests of Mood → Satisfaction link.

Link Dependent Variable	BG Data Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	RM_V2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Pleasure) → Satisfaction	.85	9.50***	3, 163	.09	.63	7.12***	3, 36	.26
Environ. Satis.		16.52***	1, 165	.09		15.51***	1, 38	.29
Perform. Satis.		22.13***	1, 165	.12		11.93***	1, 38	.24
Self-Rated Prod.		12.39***	1, 165	.07		12.11***	1, 38	.24
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Environ. Satis.	2.55	0.85	3.03	0.66	2.57	0.62	3.39	0.57
Perform. Satis.	2.43	0.85	3.00	0.70	2.43	0.68	3.29	0.80
Self-Rated Prod.	0.18	1.82	1.19	1.89	0.32	1.19	2.35	1.67
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Arousal) → Satisfaction	.96	2.49	3, 163		.95	0.70	3, 36	
Mood (Dominance) → Satisfaction	.96	2.36	3, 169		.93	0.95	3, 37	

Note. ****p*<.001. ***p*<.01. **p*<.05.

The final set of tests for the health and well-being outcomes concerned ratings of task difficulty. None of the mood measurements predicted task difficulty ratings (Table 70) for either data set.

Table 70. Tests of Mood → Task Difficulty link.

Link Dependent Variable	BG Data Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	RM_V2 Wilks' Λ	*F*	*Df*	$\eta^2_{partial}$
Mood (Pleasure) → Task Difficulty	.98	0.56	7, 158		.93	0.33	7, 32	
Mood (Arousal) → Task Difficulty	.96	1.00	7, 159		.89	0.58	7,32	
Mood (Dominance) → Task Difficulty	.93	1.79	7, 163		.88	0.66	7, 33	

4.7.2.5 Mood → Motivation link.

We had predicted that mood would influence motivation, and tested this link using all three mood measurements. The results are summarized in Table 71. Because there was only one dependent variable for motivation, these are ANOVA rather than MANOVA results. There was one statistically significant result, a small relationship between dominance and motivation for the BG data set. It was in the opposite direction to what we would have expected, with higher motivation scores for the people with below-median dominance scores. One possibility might be that people who felt more in control (higher dominance scores) exercised this control by not persisting so long when the task became extremely difficult, as it did inexorably and automatically.

Table 71. Tests of Mood → Motivation link.

Link	BG Data				RM_V2			
Dependent Variable	*F*		*df*	$\eta^2_{partial}$	*F*		*df*	$\eta^2_{partial}$
Mood (Pleasure) → Motivation	2.78		1, 168		0.04		1, 40	
Mood (Arousal) → Motivation	3.20		1, 168		0.00		1, 40	
Mood (Dominance) → Motivation	3.93*		1, 174	.02	2.70		1, 41	
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Motivation	650.31	179.57	587.40	222.92				

Note. *$p<.05$.

4.7.2.6 Competence → Motivation link.

We had also predicted that environmental competence would relate to motivation, and tested this link, with results shown in Table 72. Neither data set showed a statistically significant relationship.

Table 72. Test of Environmental Competence → Motivation link.

Link	BG Data			RM_V2		
Dependent Variable	*F*	*df*	$\eta^2_{partial}$	*F*	*df*	$\eta^2_{partial}$
Environmental Competence → Motivation	0.22	1, 179		0.67	1, 43	

4.7.2.7 Motivation → Performance link.

We tested each set of performance measurements in their appropriate MANOVA model, to see whether or not motivation was associated with them. These are summarized in Table 73. Some, but not all, of the performance measures were related to motivation. The least ambiguous effects occurred in the BG data, in which more motivated individuals had higher hit rate scores and were more accurate in categorizing summaries. In both cases the RM_V2 data did not show statistically significant multivariate tests, so we did not report the univariate tests; however, we note that some of the univariate tests in the RM_V2 data were statistically significant. (We await the analysis of the new data from the supplementary experiment to provide a better replication.) RM_V2 showed a significant effect on cognitive judgments in which more motivated individuals gave lower ratings for the quality of the summaries they read. People in both sets showed differences in break durations in relation to motivation; people with higher motivation paused slightly longer.

Table 73. Tests of Motivation → Performance link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Attention)	.76	25.46***	2, 162	.10	0.88	2.54	2, 36	
Hit Rate		38.00***	1, 163	.19				
Envelope Speed		1.54	1, 163	.01				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Hit Rate	0.17	0.26	0.43	0.25				
Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Typing)	.97	1.80	3, 177		0.91	1.35	3, 41	
Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Cognitive)	0.88	5.89***	4, 176	.01	0.86	1.66	4, 40	
Cat. Time		2.25	1, 179	.01				
Cat. Acc.		20.34***	1, 179	.01				
Judge Time		3.41	1, 179	.02				
Summ. Time		0.04	1, 179	.00				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Cat. Acc.	29.17	7.74	33.42	4.83				
Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Cognitive Judgments)	0.98	0.71	4, 176		0.80	2.56*	4, 40	.11
Interest						2.28	1, 43	.05
Factual						4.81*	1, 43	.10
Grammar						6.10*	1, 43	.12
Well-written						8.36**	1, 43	.16
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Interest					45.29	18.96	52.94	14.59
Factual					2.89	0.45	2.46	0.70
Grammar					2.80	0.47	2.42	0.50
Well-written					2.77	0.47	2.25	0.64

Link Dependent Variable	BG Data Wilks' Λ	 *F*	 *df*	 $\eta^2_{partial}$	RM_V2 Wilks' Λ	 *F*	 *df*	 $\eta^2_{partial}$
Motivation → Performance (Break Durations)	0.94	3.91**	3, 177	.02	0.71	5.38***	3, 40	.08
Typing Breaks		0.56	1, 179	.00		0.48	1, 42	.01
Conveyor Breaks		4.05*	1, 179	.02		0.94	1, 42	.02
Summary Breaks		8.20**	1, 179	.04		10.50***	1, 42	.20
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Typing Breaks								
Conveyor Breaks	0.45	0.38	0.56	0.29				
Summary Breaks	1.35	0.67	1.58	0.42	0.88	0.62	1.40	0.44

Note. $^{***}p<.001$. $^{**}p<.01$. $^{*}p<.05$.

4.7.2.8 Summary: appraisal path.

The appraisal path received good support in these analyses. People with more favorable appraisals about lighting showed higher scores on preference ratings for the space. People with higher preference ratings showed higher mood scores. Those with higher mood scores showed better outcomes for health and satisfaction, although they were not different in terms of perceived task difficulty. The link between mood and motivation was not clearly made, although there was a hint of an unusual relationship that is worthy of further examination. Motivation did influence performance on some, but not all, tasks. Additional support for the idea that motivation and performance are linked comes from the pattern of results for the Best Practice vs Switching Control and Best Practice vs Dimming Control planned comparisons described above. This is discussed in detail in the General Discussion, below.

4.7.3 Vision Path

4.7.3.1 Median split groups.

As for the appraisal path variables, we formed median-split groups for the antecedent variables involved in this path. Their characteristics are found in Table 74.

Table 74. Median split groups for vision path analyses.

Variable Set	 Split	Low Group N	 *M*	 *SD*	High Group N	 *M*	 *SD*
Visual Comfort							
BG	<3 or => 3	73	0.86	0.90	108	7.75	4.89
RM_V2	< 1 or => 1	18	0	0	26	6.04	6.68
Visual Capability (Composite visual performance at contrast = 16)							
BG	< 1.48 or => 1.48	91	1.21	0.26	89	1.74	0.22
RM_V2	< 1.60 or => 1.60	22	1.21	0.35	22	1.83	0.21

Note. BG = between-groups data (N=181). RM_V2 = repeated-measures data from second visit (N=45).

4.7.3.2 Appraisal → Visual Comfort link.

We used the Appraisal split (see above) to test whether appraisals of the lighting were associated with visual comfort directly, in addition to the indirect linkage we tested above. This link was statistically significant for both groups (Table 75), with medium-sized effects in which higher appraisal scores are associated with fewer or less severe comfort problems.

Table 75. Test of Appraisal → Visual Comfort link.

Link Dependent Variable	BG Data *F*	*df*	$\eta^2_{partial}$	RM_V2 *F*	*df*	$\eta^2_{partial}$
Appraisal → visual comfort	28.43***	1, 178	.14	5.15*	1, 42	.11

	BG Data Low *M*	Low *SD*	High *M*	High *SD*	RM_V2 Low *M*	Low *SD*	High *M*	High *SD*
Visual comfort	7.51	5.71	3.56	4.13	5.89	8.34	1.96	2.51

Note. ***$p<.001$. *$p<.05$.

4.7.3.3 Visual Comfort → Visual Capability link.

We tested this link using MANOVA, in which composite visual performance scores at each contrast level on the timed vision test were the six dependent variables. The overall multivariate test did not reach statistical significance (Table 76). Given the generally low incidence of visual comfort problems, this is not surprising. It might be possible to detect such effects using more extreme lighting or task conditions, or by influencing visual comfort in some other way to provide a more extreme contrast.

Table 76. Test of Visual Comfort → Visual Capability link.

Link Dependent Variable	BG Data Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	RM_V2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual comfort → visual capability	0.95	1.41	6, 173		0.85	1.12	6, 37	

4.7.3.4 Visual Capability → Competence link.

Although we had measured task competence in session 1, we decided to include it in a MANOVA model along with environmental competence to see whether or not visual capability related to competence overall. This proved not to be the case (Table 77).

Table 77. Test of Visual Capability → Competence link.

Link Dependent Variable	BG Data Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	RM_V2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → competence	0.98	1.41	2, 177		0.92	1.89	2, 41	

4.7.3.5 Visual Capability → Task Performance link.

As above, we tested each conceptual group of performance tasks separately in a MANOVA to see whether or not visual capability would influence them. The results are summarized in Table 78.

There were no statistically significant effects for the RM_V2 data, but several small- to medium-sized effects in the BG data, which follow the expected directions and occur on the most visually demanding tasks. People higher in visual capability showed better attention both in terms of hit rate on the conveyor belt task and in speed of responding to the envelope prompt. They also performed better on the typing task. The pattern of effects for the three print sizes was also as expected: the size of the effect declined as the task became visually easier (print size increased).

Table 78. Tests of Visual Capability → Performance link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Attention)	.93	6.38	2, 162	.05	.92	1.52	2, 35	
Hit Rate		10.74***	1, 163	.06				
Envelope Speed		4.82*	1, 163	.03				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Hit Rate	0.28	0.26	0.42	0.28				
Envelope Speed	0.34	0.13	0.38	0.13				

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*Df*	$\eta^2_{partial}$
Visual capability → Performance (Typing)	0.89	7.18***	3, 176	.08	.91	1.39	3, 40	
8 point		20.61***	1, 178	.10				
12 point		16.61***	1, 178	.09				
16 point		13.82***	1, 178	.04				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
8 point	2.18	0.95	2.93	1.24				
12 point	2.30	0.98	2.99	1.26				
16 point	2.40	0.99	3.93	1.28				

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Cognitive)	0.97	1.18	4, 175		0.94	0.59	4, 39	

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Cognitive Judgments)	0.99	0.45	4, 175		0.83	1.98	4, 39	

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Break Durations)	0.98	1.50	3, 176		0.85	2.37	3, 40	

Note. ***p<.001. **p<.01. *p<.05.

4.7.3.6 Summary: vision path.

This path received partial support. Lighting appraisals predicted visual comfort, but visual comfort (at least at the levels achieved here) did not related to visual capability. Visual capability was not related to competence. Visual capability did relate to task performance on the tasks where such relationships were to be expected, at least in the BG data. One cannot rule out sample size and restricted range in the RM_V2 data as reasons for not finding visual capability effects; that is, the fact of not observing the effects does not rule out their existence, especially given the literature on visibility.

4.7.4 Exploratory Tests

In addition to the planned tests, we considered several new questions that arose as we thought further about how the variables might inter-relate.

4.7.4.1 Motivation → Health and Well-being?

Although we had not proposed one, it seemed plausible to consider that people who are more motivated might show higher levels of health and well-being. We tested this hypothesis as we had for mood → health and well-being, described above, using separate MANOVA models for health, satisfaction, and task difficulty. None of these analyses returned statistically significant results.

4.7.4.2 Satisfaction → Task Performance?

It is a perennial question, "Are happier workers better workers?" We did not design this experiment in order to test this hypothesis directly. The measurements of satisfaction occurred temporally following all of the performance measurements, and therefore should depend on the participants' experiences during those tasks. This breaks the temporal chain of causality. Nonetheless, in order to address the inevitable question about the connection between the end points of the appraisal and vision paths, we formed three median-split groups to examine it. The groups are described in Table 79.

Table 79. Median split groups for satisfaction analyses.

Variable		Low Group			High Group		
Set	Split	N	*M*	*SD*	N	*M*	*SD*
Environmental Satisfaction							
BG	< 3 or => 3	82	2.12	0.64	99	3.32	0.39
RM_V2	<= 3 or > 3	28	2.59	0.51	17	3.69	0.35
Performance Satisfaction							
BG	< 2.75 or >= 2.75	78	1.96	0.64	100	3.25	0.47
RM_V2	< 3 or => 3	21	2.23	0.61	23	3.51	0.41
Self-Rated Productivity							
BG	<= 0 or > 0	101	-0.75	1.03	80	2.50	1.11
RM_V2	<= 1.5 or > 1.5	22	-0.27	1.12	22	3.00	0.87

Note. BG = between-groups data (N=181). RM_V2 = repeated-measures data from second visit (N=45).

We tested the links from satisfaction to performance using the same set of performance outcomes as have been previously described. Only the tests with statistically significant outcomes are reported here (Table 80). Of the 30 tests (15 within each sample, x two samples), only three showed a statistically significant multivariate effect. For the BG sample, higher environmental satisfaction was associated with higher ratings of how well-written the summaries were, and higher self-rated productivity was associated with faster work in making judgments about summaries and in creating summaries from articles. For the RM_V2 sample, higher performance satisfaction was associated with lower accuracy in categorizing summaries.

Table 80. Tests of Satisfaction → Task Performance link.

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Environmental Satisfaction → Performance (Cognitive Judgments)	.93	3.15**	4, 176	.02	0.94	0.69	4, 40	
Interest		0.26	1,179	.00				
Factual		2.54	1,179	.04				
Grammar		1.09	1,179	.01				
Well-written		9.84**	1,179	.05				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Well-written	2.32	0.50	2.57	0.56				

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Performance Satisfaction → Performance (Cognitive)	.96	1.65	4, 173		.79	2.61*	4,39	.04
Cat. Time						0.70	1, 42	.02
Cat. Acc.						4.95*	1, 42	.11
Judge Time						0.05	1, 42	.00
Summ. Time						1.00	1, 42	.02
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Cat. Acc.					33.91	4.43	29.65	7.66

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Self-Rated Productivity → Performance (Cognitive)	0.95	2.47*	4, 176	.02	0.86	1.55	4, 39	
Cat. Time		1.40	1, 179	.01				
Cat. Acc.		2.77	1, 179	.02				
Judge Time		5.05*	1, 179	.03				
Summ. Time		5.02*	1, 179	.03				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Judge Time	77.04	31.71	66.40	31.47				
Summ. Time	49.30	37.85	38.44	23.68				

Note. ****p*<.001. ***p*<.01. **p*<.05.

4.7.4.3 Performance → Satisfaction?

Given the ambiguous directionality of the link between our measurements of satisfaction and the performance measurements, we decided also to test the reverse link. We created three median-split variables, choosing variables that either had been sensitive in the earlier analyses (e.g., break durations on the summarization task, interest in reading articles), or that made logical sense (typing at 8-point print size). We then conducted 18 MANOVA tests to determine whether groups based on these splits predicted any of the three health and well-being outcomes in either sample. The median split groups are described in Table 81 and the results of the analyses in Table 82.

Table 81. Median split groups for performance analyses.

Variable		Low Group			High Group		
Set	Split	N	*M*	*SD*	N	*M*	*SD*
Typing (8 pt)							
BG	<= 2.39 or > 2.39	91	1.64	0.47	90	3.46	0.90
RM_V2	<= 2.22 or > 2.22	23	1.59	0.57	22	3.33	1.11
Summary Breaks							
BG	<= 1.71 or >1.71	90	1.16	0.55	91	1.84	0.10
RM_V2	<= 1.38 or > 1.38	22	0.79	0.49	22	1.64	0.19
Interest in Article							
BG	<= 50.83 or > 50.83	90	38.06	11.49	91	62.96	11.11
RM_V2	<= 50 or > 50	24	39.31	11.55	21	62.69	11.75

For the BG data there was one test that showed statistical significance: typing performance was associated with task difficulty ratings, and the significant multivariate test was explained by a significant univariate test for the rating of typing task difficulty. Poorer typists found the typing task more difficult. This is an unsurprising finding, but one that does demonstrate discriminant validity for the task difficulty ratings.

For the RM_V2 data there was a statistically significant relationship between the duration of breaks during the summarization task and satisfaction, with the multivariate test being accompanied by significant univariate tests for environmental satisfaction and self-rated productivity. Those who took longer breaks during this task tended to be less satisfied with the environment and to rate it as having a poorer effect on their own productivity. The direction of this effect is consistent with the inverse effect reported above, an effect of self-rated productivity on cognitive performance, reported for the BG data set.

Table 82 Tests of Task Performance → Satisfaction link..

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Performance (Typing) → Task Difficulty	.87	3.68***	7, 168	.02	0.78	1.40	7, 35	
Typing		21.25***	1, 174	.10				
Categorizing		0.88	1, 174	.01				
Evaluating summ.		0.04	1, 174	.00				
Extracting summ.		0.05	1, 174	.00				
Questionnaires		0.93	1, 174	.01				
Vision tests		0.08	1, 174	.00				
All tasks		0.60	1, 174	.00				
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Typing	1.64	1.09	0.94	0.96				

Link	BG Data				RM_V2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Performance (Summ. Breaks) → Satisfaction	0.99	0.39	3, 174		0.78	3.57*	3, 38	.11
Perform. Sat.						0.87	1, 40	.02
Env. Sat.						10.48***	1, 40	.21
Self-rated Prod.						4.85*	1, 40	.11
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Env. Sat.					3.34	0.64	2.72	0.67
Self-rated Prod.					2.05	1.61	0.77	2.11

4.7.4.4 Mood affects Social Behavior?

Social behavior was not included on the linked mechanisms map, but had been included in the experimental design because of suggestions in the literature that positive affect might influence behavior in desirable ways (Baron et al., 1992), and because of a request to examine it. We did not find lighting effects on social behavior but wondered if the positive affect hypothesis would hold, that people who were happier or more satisfied would show more desirable social behaviors. We approached this question with two analyses, the first examining pleasure as a predictor and the second involving environmental satisfaction. The dependent variables were: liking for other group members, willingness to volunteer, and the two types of responses to the conflict resolution task (ratings and rankings).

Pleasure did not relate to any of the social behavior outcomes. However, environmental satisfaction did relate to several outcomes, with small effects in the direction expected (Table 83). People in the above-median group for environmental satisfaction for the BG sample were more willing to volunteer and showed higher liking for other group members. For the RM_V2 sample (which was made up of people willing to return, one of the questions in the willingness-to-volunteer measurement), there was no effect on willingness to volunteer but there was a significant effect for liking for other group members.

Table 83. Tests of Environmental Satisfaction → Social Behavior link.

Link	BG Data				RM_V2			
Dependent Variable	*F*		*df*	$\eta^2_{partial}$	*F*	*df*		$\eta^2_{partial}$
Environmental Satisfaction → Willingness to Volunteer	4.21*		1, 178	.02	2.68	1, 42		
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Willingness to Volunteer	-0.14	0.87	0.12	0.85				
Link	BG Data				RM_V2			
Dependent Variable	*F*		*df*	$\eta^2_{partial}$	*F*	*df*		$\eta^2_{partial}$
Environmental Satisfaction → Liking	10.06***		1, 178	.05	4.08*	1, 43		.09
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Liking	1.15	0.71	1.50	0.76	1.39	0.80	1.86	0.69

Note. ***$p<.001$. *$p<.05$.

For the conflict resolution task, there were no effects when the scores were rankings of the relative likelihood that the individual would use each of the various conflict resolution strategies but there were effects when the dependent variables were individual ratings of their likelihood of using each strategy (Table 84). There was a multivariate effect for both the BG and RM_V2 samples, although the univariate effects differed. For the BG group, people high in environmental satisfaction were more likely to use more co-operative styles of conflict resolution (accommodation, collaboration, or compromise) than people low in environmental satisfaction. For the RM_V2 group, the effect for accommodation held, and people high in environmental satisfaction were less likely to use avoidance, which is an uncooperative, unassertive form of conflict resolution (Kilmann & Thomas, 1977; Thomas, 1976).

The relationships between environmental satisfaction and social behavior are suggestive but not conclusive. The same temporal causality problem applies here as in the examination of satisfaction → performance linkages: environmental satisfaction was measured at the end of the day. Moreover, it is puzzling that pleasure did not predict any of the social behaviors, as it ought to have done according to the positive affect theory. Although pleasure did predict environmental satisfaction (see above, under "Appraisal Path"), mean pleasure scores were not as widely different in the low- and high- environmental satisfaction groups (3.32 and 4.67 for BG, and 4.27 and 5.82 for RM_V2) and as they were in the low- and high-pleasure groups (3.03 and 5.74 for BG and 3.49 and 6.38 for RM_V2). Thus, pleasure as an indicator of positive affect does not fully account for the effect. Nonetheless, the effect does suggest that there be further study of the possibility that conditions supportive of environmental satisfaction might encourage more pro-social behavior. Given the costs of conflict management to organizations, there could be important implications for organizational outcomes (Baron, 1990a).

Table 84. Test of Environmental Satisfaction → Conflict Resolution.

Link Dependent Variable	BG Data Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	RM_V2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Environmental Satisfaction → Conflict Resolution (Ratings)	.92	3.17	5, 172	.03	.73	2.83*	5, 38	.06
Competition		0.82	1, 176	.00		2.87	1, 42	.06
Accommodation		9.10**	1, 176	.05		6.03*	1, 42	.13
Collaboration		6.43**	1, 176	.04		1.29	1, 42	.03
Compromise		7.91**	1, 176	.04		0.00	1, 42	.00
Avoidance		2.81	1, 176	.02		4.56*	1, 42	.10
	Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Competition								
Accommodation	0.16	0.75	.51	.78	0.15	0.58	0.71	0.94
Collaboration	1.16	0.81	1.46	0.76				
Compromise	1.64	0.61	1.90	0.61				
Avoidance					-1.22	0.80	-1.72	0.66

4.7.5 Summary: Linked Mechanisms Tests

The linked mechanisms tests, despite their weaknesses, provided useful information about the relationships between the concepts in the linked mechanisms maps. Although some links were not supported, most were supported at least in part, and in the anticipated direction. This approach to making the connections does not rule out all possible alternative explanations. Nonetheless, most links have support from other investigations and from supporting theories, which can give us more confidence in the map as a means of understanding how employees experience and respond to their working environments. Although the median –split groups were independent of lighting condition, we can learn from the other analyses and from other research how lighting conditions influence lighting appraisals, and thereby influence the entire set of linked mechanisms. This overall interpretation of the linked mechanisms together with the other analyses, is the subject of the General Discussion.

5. Method - Experiment 2

Experiment 2 was undertaken for two reasons. First, to provide another test of the hypothesis that better quality lighting will enhance the health, well-being, and task performance of office workers. Second, to provide a comparison between a different form of Base Case lighting and a variant of the Best Practice lighting used in Experiment 1. The Base Case 2 lighting used in Experiment 2 was a regular array of recessed prismatic luminaires with the same luminaire layout as the Base Case parabolic luminaire lighting used in Experiment 1. The Best Practice 2 lighting used in Experiment 2 was the same linear array of the direct / indirect luminaires and the wall washing luminaires used in Experiment 1, but with the direct / indirect luminaires dimmed to provide a lower illuminance on the desks. It is a common practice to reduce the design illuminance when using indirect lighting to compensate, in terms of energy consumption, for the reduced efficiency of indirect lighting relative to direct lighting.

5.1 Setting

The setting used for experiment 2 was identical to that used in experiment 1 (see Section 3.1).

5.2 Finishes and Furnishings

The finishes and furnishings used in Experiment 2 were identical to those used in Experiment 1 (see Section 3.2)

5.3 Lighting Manipulations

Two different lighting installations were used in Experiment 2. The two lighting installations were both designed to be representative of current office lighting practice.

The two lighting installations were:

1. Base Case 2: This installation was designed to represent another common approach to office lighting. It consisted of a regular array of three lamp, recessed prismatic luminaires, mounted on 8 ft by 10 ft centers in the suspended ceiling. Each luminaire contained three T8, 32W, fluorescent lamps of correlated color temperature (CCT) of 3,500K and a CIE General Color Rendering Index (CRI) of 82. These three fluorescent lamps were operated from an electronic dimming ballast. This lighting was installed in the areas frequently used by the subjects, i.e., the cubicle area, hallway and conference room. Figures 38 and 39 show the space lit with this installation. Other areas i.e., restrooms, kitchen, storage rooms, research office, and vision screening room, were lit by the parabolic luminaire Base Case 1 installation throughout both experiments. This installation provided an adequate illuminance on the working plane but the luminaire could be seen as bright from a wide range of positions. This made it more likely that the installation would create high luminance reflections visible on the screen of the computer monitor, although not every workstation had this characteristic (Figure 40).

2. Best Practice 2: This installation was designed to provide a lower illuminance on the working plane than the Base Case 2 installation and to produce brighter walls and ceiling. The luminaires, lamps and layout were identical to those used for Best Practice 1, but the direct / indirect luminaires in the cubicle area were dimmed to produce a lower illuminance on the working plane, more typical of contemporary installations with direct / indirect systems. All other areas were lit in exactly the same manner as for Best Practice 1 (see section 3.3, Figure 10). The photographs in Figures 11-12 show this installation, although at the higher illuminance setting used in Experiment 1.

Figure 38. General view of workstations with Base Case 2 lighting installation.

Figure 39. View to the break area under Base Case 2 lighting installation.

Figure 40. Two workstation views from the Base Case 2 lighting installation

5.4 Lighting Monitoring

Both lighting installations were operated through a control system. This system allowed the luminaires forming the installation to be put into functional groups, e.g., for the Best Practice 2 installation, the direct / indirect luminaires over the cubicle area formed one functional group and the wall washing luminaires formed another. Each functional group could be dimmed as desired. The power demand for each functional group of luminaires was recorded every five minutes.

5.5 Photometric Conditions

5.5.1 Illuminances

Table 85 shows the mean illuminances provided by the two lighting installations on the work surface, on the monitor screen, on the keyboard, and at the subject's eyes, in the cubicles with the gray and navy blue panels. A comparison of Table 85 with Table 3 shows that the Base Case 2 lighting provides slightly lower horizontal illuminances and slightly higher vertical illuminances than those given by the Base Case 1 lighting. However, the Best Practice 2 lighting provides illuminances about 25 to 30 percent lower than does the Best Practice 1 lighting, for both vertical and horizontal illuminances.

Table 85. Means of the illuminances (lx) provided on the work surface, on the monitor, on the keyboard, and at the subject's eye, by the two lighting installations in the cubicles with the gray and dark blue panels.

Lighting installation and panel type	Illuminance on work surface (lx)	Illuminance on monitor (lx)	Illuminance on keyboard (lx)	Illuminance at subject's eye (lx)
Base Case 2				
Gray	567	364	586	321
Blue	514	352	547	294
Best Practice 2				
Gray	431	283	443	269
Blue	398	242	446	235

Illuminances were also measured at other salient locations in the space, i.e., the reception desk, the sofa near the reception desk, on the filing cabinets in the hallway, and on the tables at the end of the hall way that were used for meeting. Table 86 shows the mean illuminances provided by the two lighting installations at these locations.

Table 86. Means (and standard deviations) of the illuminances (lx) provided on the reception desk, on the sofa near the reception desk, on the filing cabinets in the hallway, and on the tables at the end of the hallway, by the two lighting installations.

	Illuminance on reception desk (lx)	Illuminance on sofa (lx)	Illuminance on filing cabinets in hallway (lx)	Illuminance on tables at end of hallway (lx)
Base Case 2	425 (69.5)	536 (24.6)	485 (51.9)	427 (61.5
Best Practice 2	385 (62.9)	419 (35.4)	398 (42.8)	419 (29.4)

As for the conference room, the illuminance on the table in front of each chair and on the faces of people sitting in each chair were measured. The mean illuminances for the two lighting installations used in the conference room are given in Table 87.

Table 87. Means (and standard deviations) of the illuminances (lx) provided on the conference room table, and at the eyes of subjects sitting around the table, for the two lighting installations used in the conference room.

Lighting installation	Illuminance on table (lx)	Illuminance at subject's eye (lx)
Base Case 2	565 (40.3)	395 (56.6)
Best Practice 2	567 (35.9)	440 (48.4)

5.5.2 Luminances

5.5.2.1 Room surfaces

The luminances of the all the surfaces in the cubicle area, the hallway and the conference room, under each lighting installation, were measured using a digital imaging photometer, and the appearance of the space was recorded with a digital camera.

5.5.2.2 Computer monitor

The computer monitors used extensively throughout the study were all set up to have a similar background color. The screen luminance of a blank white screen display, in the absence of any lighting in the room, was 120 cd/m^2.

5.6 Thermal and Acoustic Conditions

The air temperature and relative humidity conditions provided by the air-conditioning in the office were typical of offices and showed little variation over the time of the experiment. The sound pressure levels occurring in the office were less than 50 dB(A). Such sound pressure levels are typical of a quiet office.

5.7 Experimental Design

5.7.1 Outline

There were two independent variables for Experiment 2, one being the two different lighting installations; the other being the two reflectances of the panels forming the background to the computer in the cubicle. Experiment 2 was a between groups experiment only.

5.7.2 Sample Size

The target number of subjects per condition used in Experiment 2 was the same as for the between groups part of Experiment 1, i.e., about 50. In the between-groups experiment, each participant is randomly assigned to one of four experimental conditions (two lighting conditions crossed by two cubicle panel reflectances). Table 88 summarizes the experimental design. The random assignment of participants ensures that individual differences are equally distributed between the groups, and allows the inference that if any differences exist between the groups on any dependent measure, then the treatment is the cause of the effect (Kerlinger & Lee, 2000).

Table 88. Schematic Experimental Design: Experiment 2

	Base Case 2	Best Practice 2
Gray Panel	Between groups N = 25	Between groups N = 25
Blue Panel	Between groups N = 25	Between groups N=25
	N = 50 participants randomly assigned to the 2 above conditions	

5.7.3. Participant Characteristics

As before, the participants in Experiment 2 were recruited from an office temporary services agency; and paid for at the agency rate for each day they worked. They were all over 18 years of age and were required to have experience with Windows-based word processing and spreadsheet software and a minimum typing speed of 30 words per minute, and have normal or corrected to normal vision and hearing. None of the participants taking part in Experiment 2 had taken part in Experiment 1. Table 89 shows how the 107 participants in Experiment 2 were allocated to the different conditions.

Table 89. Experiment 2 participants by experimental condition

	Base Case 2	Best Practice 2	Total
Light (Gray)	29	29	58
Dark (Blue)	23	26	49
Total	52	55	107

The demographic characteristics of the Experiment 2 participants are summarized in Table 90.

Table 90. Experiment 2 sample characteristics

Sex	42 Male			65 Female		
Age	60 – 18-29	22 – 30-39	13 – 40-49	9 – 50-59	2 – 60-69	1 – 70-79
Corrective Lenses	40 – None	13 – Reading	31 – Distance	6 – Bi- or Tri-focal	4 – Multi-focus	13 – Contact lenses
Education	22 – HS	12 – CC	31 – UC	31 – UD	11 – GD	

Note. HS = high school diploma. CC = community college diploma. UC = undergraduate courses. UD = undergraduate degree. GD = graduate degree.

We inquired about the frequency of hearing impairment or use of a hearing aid on the testing day. Five participants reported having at some time being diagnosed with a hearing impairment, but none wore a hearing aid at the test. In addition, no one reported any functional problems with hearing that prevented them from communicating with others during the session.

In the separate vision testing (section 3.7.3), participants' color vision was tested: Two people failed the Ishihara color test. All participants had a full visual field.

The contrast sensitivity test consisted of five rows of rings, from high to low spatial frequency; in each row were nine targets of diminishing luminance contrast. Scores on this test were the sum of correctly-identified rings, from 0 to 45. The mean contrast sensitivity score was 33.38 (*SD* = 4.63). Near-field visual acuity was generally very high; the distribution is shown in Table 91.

Table 91. Experiment 2 participants' near-field visual acuity

	Visual Acuity 20/							
	20	25	30	35	40	50	60	80
Total frequency	49	34	11		11	2		

We also asked participants to report the duration of their time in the paid work force and as a temporary office worker. Overall, participants had been in the workforce for an average of 12.2 years (*SD* = 10.42) and as a temporary worker for an average of 1.2 years (*SD* = 2.09).

5.8 Independent Variables

Experiment 2 had two independent variables, lighting design and panel reflectance, that are applicable to all the dependent variables. As in Experiment 1, time was an additional independent variable for some outcomes. Print size was an additional variable for the typing task, and contrast was an additional variable for the timed vision test.

5.9 Dependent Variables

5.9.1 Daily Schedule

The daily schedule adopted for Experiment 2 was the same as for Experiment 1 (see Section 3.9.1 and Table 15).

5.9.2 Perceptions and Feelings

The instruments used to measure the participants' perceptions and feelings in Experiment 2 were the same as those used in Experiment 1 (see Section 3.9.2).

5.9.3 Performance Measures

The measurements taken to quantify the participants' performance at the various tasks in Experiment 2 were the same as those used in Experiment 1 (see Section 3.9.3).

5.9.4. Social Behavior

The questions used to assess the participants' social behavior in Experiment 2 were the same as those used in Experiment 1 (see Section 3.9.4).

5.10 Daily Procedure

The daily procedure followed in Experiment 2 was the same as that followed in Experiment 1 (see Section 3.10).

5.11 Ethical Issues

A revised experiment protocol for Experiment 2 was submitted for approval to the Ottawa Research Ethics Board at the National Research Council of Canada, and to the Ethics Review Board of the Pacific Northwest National Laboratory, prior to the commencement of data collection (The Institute Review Board of Rensselaer Polytechnic Institute did not examine the protocol of Experiment 2 because it was considered as a simple extension of Experiment 1 and not a new experiment). Both Boards approved the protocol. The informed consent procedure and forms were as in Experiment 1.

6.0 Results - Experiment 2

6.1 Manipulation Checks

6.1.1 Demographics

We checked the random assignment assumption against several demographic characteristics. There were no differences between the groups (2 Lighting Designs or 2 Reflectances) for sex (Lighting Design, $\chi^2(1) = 1.83$, $p>.05$; Reflectance, $\chi^2(1) = 0.09$, $p>.05$) , age (Lighting Design, $\chi^2(5) = 3.71$, $p>.05$; Reflectance, $\chi^2(5) = 4.93$, $p>.05$), education (Lighting Design, $\chi^2(4) = 1.97$, $p>.05$; Reflectance, $\chi^2(4) = 5.08$, $p>.05$), or type of corrective lenses (Lighting Design, $\chi^2(5) = 2.56$, $p>.05$; Reflectance, $\chi^2(5) = 9.05$, $p>.05$). In addition, the distributions did not differ for hearing impairment (Lighting Design, $\chi^2(1) = 2.07$, $p>.05$; Reflectance, $\chi^2(1) = 0.07$, $p>.05$). The frequency of impaired color vision did not differ between the lighting conditions, with one in each of the two groups ($\chi^2(1) = 0$, $p>.05$), nor across Reflectances ($\chi^2(1) = 2.41$, $p>.05$). For contrast sensitivity, we conducted a 2 x 2 Lighting Design X Reflectance ANOVA on the overall scores and found no main effects or interactions. Similarly, there were no main effects or interactions for a 2 x 2 Lighting Design X Reflectance ANOVA for visual acuity. We also asked participants to report the duration of their time in the paid work force and as a temporary office worker. These continuous variables we tested in a 2 x 2 (Lighting Design by Reflectance) MANOVA design that also showed no differences. Overall, random assignment appears to have successfully created two broadly equivalent groups on these demographic characteristics.

6.1.2 Participant Expectancies

At the end of each testing day, participants were asked to report their beliefs about the purpose of the experiment and their feelings about what might have influenced their performance or responses. These questions we ask routinely to probe the extent to which participants had expectancies about the expected outcome of the investigation, or beliefs about how the researchers had wanted them to act. In reading these responses we looked particularly for indications that participants had identified lighting as the focus of the experiment.

A large percentage of the participants (53% overall) identified the effects of lighting on performance, vision, or well-being as the purpose of the experiment. This percentage was higher for participants in the

Best Practice 2 condition (62%) than the Base Case 2 condition (44%). The pattern is the same as it was in Experiment 1, but the overall percentages are higher.

Participants were also asked about what, in their opinion, might have influenced them during the day. Overall, 20% said that lighting had affected them, 3% positively and 17% negatively. All of the positive responses were from people in the Best Practice 2 condition. The negative responses were split across conditions, with 19% of Base Case 2 participants, and 15% of Best Practice 2 participants, reporting negatively about the lighting. Of the Best Practice 2 complaints, some stated that the lighting was too bright and others too dim.

In addition, 23% percent of the participants reported boredom and fatigue (21% of Base Case 2, and 25% of Best Practice 2). Conversely, 25% reported having had a good experience and having enjoyed the day (23% of Base Case 2 and 27% of Best Practice 2). Ten per cent expressed interest in the experiment and in knowing more, with a small tendency for people in Best Practice 2 to express this more frequently (13%, versus 8% for Base Case 2).

Overall, the pattern does not suggest that participant expectancies about the expected outcomes biased the results. If that were the case we would expect a higher percentage of lighting-positive comments in Best Practice 2 and lighting-negative comments in Base Case 2.

6.2 Perceptions and Feelings

In this section we again organized the results by the conceptual groups of outcomes, and used the same analytic strategies as for Experiment 1.

6.2.1 Lighting Appraisal

6.2.1.1 LRC's Office Lighting Survey.

The results of the non-parametric analysis of the afternoon responses to this questionnaire are in Table 92. Here, too, the results were compared to the normative data for the OLS (Eklund & Boyce, 1996). In general, the results follow expected patterns. Base Case 2 (recessed prismatic luminaires) were not different in comfort from the normative sample, with 69% agreeing that "the lighting is comfortable". Noticeably more people in the Best Practice 2 condition (direct / indirect) agreed with the statement (81%), which is comparable to the level achieved by the Best Practice condition in Experiment 1 (85%, section 4.3.1.1). Given the other responses we view this as a practically significant difference between the lighting conditions.

For instance, the Base Case 2 condition was judged to be uncomfortably bright by a higher-than-expected number of people (question 2). This question also showed a significant difference between the two groups. Similarly, Base Case 2 was also more likely to be judged to cause reflections that hindered work (question 6) and to have light fixtures that are too bright (question 7). Conversely, Best Practice 2 showed a lower-than-expected percentage agreement that "the lighting is poorly distributed here" (question 4). Overall it appears that people judged the Best Practice 2 condition to be better than Base Case 2.

Interestingly, however, that did not translate into between-group differences on question 10. There were no differences between the groups, nor between either group and the normative sample, on the question of whether the lighting installation was worse, the same, or better than in other places. This finding differs from the assessment of Best Practice in Experiment 1; in that case the same system (at a higher illuminance level) was judged to be better than in other places. It could be that the lower illuminance setting made this lighting design equivalent to other places, particularly for people in the group whose personal preferences would be for higher illuminances.

Table 92. Experiment 2 OLS Results

1. Overall, the lighting is comfortable.		
Norm: 69% Agree		Between-Groups $X^2(1) = 2.00$
Base Case 2	Count of total	36 of 52
	% Agree	69
	$X^2(1)$	0.00
Best Practice 2	Count of total	43 of 53
	% Agree	81
	$X^2(1)$	3.65
2. The lighting is uncomfortably bright for the tasks that I perform.		
Norm: 16% Agree		Between-Groups $X^2(1) = 5.85$*
Base Case 2	Count of total	18 of 52
	% Agree	35
	$X^2(1)$	13.41***
Best Practice 2	Count of total	8 of 55
	% Agree	15
	$X^2(1)$	0.09
3. The lighting is uncomfortably dim for the tasks that I perform		
Norm: 14% Agree		Between-Groups $X^2(1) = 0.85$
Base Case 2	Count of total	6 of 51
	% Agree	12
	$X^2(1)$	0.21
Best Practice 2	Count of total	10 of 55
	% Agree	18
	$X^2(1)$	0.8
4. The lighting is poorly distributed here.		
Norm: 25% Agree		Between-Groups $X^2(1) = 3.91$*
Base Case 2	Count of total	12 of 52
	% Agree	23
	$X^2(1)$	0.1
Best Practice 2	Count of total	5 of 55
	% Agree	9
	$X^2(1)$	7.42**
5. The lighting causes deep shadows		
Norm: 15% Agree		Between-Groups $X^2(1) = 0.22$
Base Case 2	Count of total	4 of 52
	% Agree	8
	$X^2(1)$	2.18
Best Practice 2	Count of total	3 of 55
	% Agree	5
	$X^2(1)$	3.93*

6. Reflections from the light fixtures hinder my work.		
Norm: 19% Agree		Between-Groups $X^2(1) = 6.76^{**}$
Base Case 2	Count of total	21 of 51
	% Agree	41
	$X^2(1)$	16.3***
Best Practice 2	Count of total	10 of 55
	% Agree	18
	$X^2(1)$	0.02
7. The light fixtures are too bright.		
Norm: 14% Agree		Between-Groups $X^2(1) = 14.94^{***}$
Base Case 2	Count of total	20 of 52
	% Agree	38
	$X^2(1)$	25.84***
Best Practice 2	Count of total	4 of 55
	% Agree	7
	$X^2(1)$	2.07
8. My skin is an unnatural tone under the lighting.		
Norm: 9% Agree		Between-Groups $X^2(1) = 0.69$
Base Case 2	Count of total	12 of 52
	% Agree	23
	$X^2(1)$	12.58***
Best Practice 2	Count of total	9 of 54
	% Agree	17
	$X^2(1)$	3.88*
9. The lights flicker throughout the day.		
Norm: 4% Agree		Between-Groups $X^2(1) = 0.00$
Base Case 2	Count of total	3 of 52
	% Agree	6
	$X^2(1)$	0.42
Best Practice 2	Count of total	3 of 53
	% Agree	6
	$X^2(1)$	0.38

10. How does the lighting compare to similar workplaces in other buildings?.				
Norm: 19% worse – 60% same – 22% better		Between-Groups $X^2(2) = 1.46$		
		Worse	Same	Better
Base Case 2	Count	5 of 52	36 of 52	11 of 52
	%	10	69	21
	$X^2(2)$		3.17	
Best Practice 2	Count of total	8 of 55	32 of 55	15 of 55
	%	15	58	27
	$X^2(2)$		1.30	

Note. * $p<=.05$. ** $p<=.01$. *** $p<=.001$.

6.2.1.2 NRC's lighting quality scale.

As before, we analyzed the lighting quality scale with two dependent variables, an overall assessment of lighting quality and a specific scale for bothersome glare, both as reported in the afternoon. Internal consistency reliability was excellent for Lighting Quality, Cronbach's $\alpha = .90$. For Bothersome Glare, internal consistency was acceptable, Cronbach's $\alpha = .65$. For Base Case 2, the mean Lighting Quality rating was 2.84 ($SD = 0.82$); for Best Practice 2, the mean Lighting Quality rating was 2.83 ($SD = 0.77$). The respective ratings for Bothersome Glare were: Base Case 2 $M = 1.33$ ($SD = 1.00$), Best Practice 2 $M = 0.98$ ($SD = 0.96$). Thus, in general the lighting in both designs was rated as being of good quality, and glare was not a very bothersome problem. The statistical model for the analysis was a 2 x 2 Lighting Design x Reflectance MANOVA. There were no statistically significant effects on either dependent variable.

However, responses to open-ended questions about the sources of glare revealed that people in the Base Case 2 condition were more likely to say that the lighting was a source of glare or reflections on the computer screen. Thirty-eight per cent (38%) of Base Case 2 participants reported this, versus 16% of Best Practice 2 participants. These percentages confirm the pattern seen in the responses to question 6 of the Office Lighting Survey (Table 92, above).

6.2.2 Preference

The dependent variables, as before, were the four scales derived from the 27 semantic differential (SD) items in the Room Appearance questionnaire in Experiment 1: Attractiveness, Comfort, Visibility, and Spaciousness. Internal consistency reliability for these scales was not as strong as it had been in Experiment 1, except for the Attractiveness scale (Cronbach's α scores: Attractiveness, AM = .89, PM = .88; Comfort, AM = .39, PM = .46; Visibility, AM = .56, PM = .64; Spaciousness, AM = .51, PM = .51). (Compare to values in table 20).

The statistical analysis model was a 2 x 2 x 2 (Lighting Design X Reflectance X Time) MANOVA. We examined all the main effects and interactions. As for Experiment 1, only statistically significant effects are presented in the text.

6.2.2.1 Base Case 2 versus Best Practice 2.

There was a small, but statistically significant multivariate interaction of Lighting Design X Reflectance X Time, accompanied by a univariate effect for the rating of Comfort. The statistical tests are summarized in Table 93 and the interaction is shown graphically in Figure 41. *Post hoc* tests are summarized in Table 94. The *post hoc* tests show that in the morning, the Base Case 2 - Light was rated as more comfortable than the Best Practice 2 - Light. The Base Case 2 - Light and Best Practice 2 - Dark ratings both dropped significantly by the afternoon. Base Case 2 - Dark and Best Practice 2 - Light stayed the same.

We interpret this pattern as indicating that although Base Case 2 - Light and Best Practice 2 - Dark received higher initial ratings, they could not sustain them over time. The other two conditions received sustained ratings over time, which we interpret as a more favorable outcome in view of the main effect of time on this variable (see next section). We speculate that the Base Case 2 - Light appeared comfortable in the morning because of its overall lightness, but that it dropped in comfort by the afternoon because of its potential for glare problems. Best Practice 2 - Dark, by comparison, might be have been rated as comfortable in the morning because of the combination of direct / indirect lighting and variety in surface luminances, but that this degree of variety appears not to have been suitable for sustained work periods (see below, for Performance effects involving this interaction).

Table 93. Lighting Design x Reflectance x Time interaction effect on preference.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Multivariate	.91	2.64*	4, 100	.02
Attractive		0.20	1, 103	.00
Comfort		7.88**	1, 103	.07
Visibility		0.63	1, 103	.01
Spacious		0.39	1, 103	.00
	AM		PM	
Comfort:	*M*	*SD*	*M*	*SD*
Base Case 2 – Light	72.90	10.90	65.50	9.93
Base Case 2 – Dark	67.17	10.52	64.51	11.58
Best Practice 2 – Light	66.77	9.39	66.54	10.00
Best Practice 2 - Dark	72.35	12.19	67.09	10.87

Note. *p<=.05. **p<=.01.

Figure 41. Lighting Design x Reflectance x Time interaction effect on Preference- Comfort.

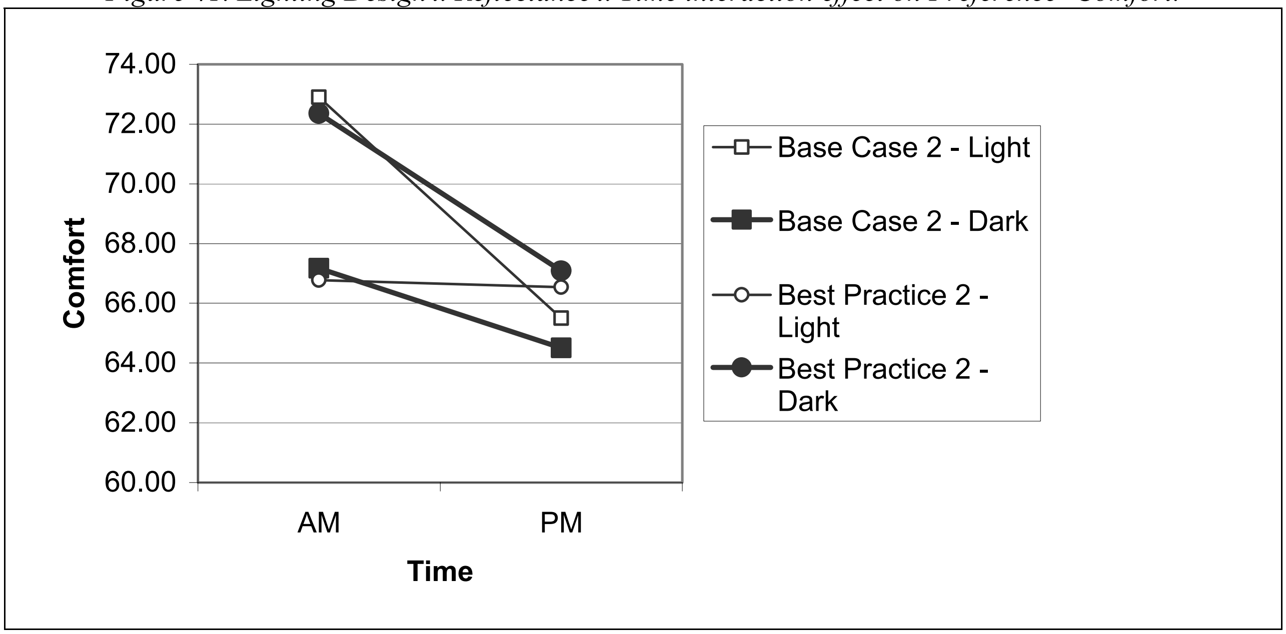


Table 94. Post hoc tests for Lighting Design by Reflectance by Time effect on Preference - Comfort.

	F	*df*	$\eta^2_{partial}$
Base Case 2 – Light – Time-linear	24.91***	1, 28	.47
Base Case 2 – Dark – Time-linear	3.22	1, 22	
Best Practice 2 – Light – Time-linear	0.01	1, 28	
Best Practice 2 – Dark – Time-linear	7.97**	1, 25	.24
Base Case 2 – Light vs Best Practice 2 – Light - AM	5.28*	1, 56	.09
Base Case 2 – Light vs Best Practice 2 – Dark - AM	0.03	1, 53	
Base Case 2 – Dark vs Best Practice 2 – Light – AM	0.02	1, 50	
Base Case 2 – Dark vs Best Practice 2 – Dark – AM	2.51	1, 47	
Base Case 2 – Light vs Base Case 2 – Dark - AM	3.66	1, 50	
Best Practice 2 – Light vs Best Practice 2 – Dark - AM	3.67	1, 53	
Base Case 2 – Light vs Best Practice 2 – Light - PM	0.16	1, 56	
Base Case 2 – Light vs Best Practice 2 – Dark - PM	0.32	1, 53	
Base Case 2 – Dark vs Best Practice 2 – Light – PM	0.46	1, 50	
Base Case 2 – Dark vs Best Practice 2 – Dark – PM	0.65	1, 47	
Base Case 2 – Light vs Base Case 2 – Dark - PM	0.11	1, 50	
Best Practice 2 – Light vs Best Practice 2 – Dark - PM	0.04	1, 53	

Note. * $p<=.05$. ** $p<=.01$. *** $p<=.001$.

6.2.2.2 Non-lighting effects.

The results for Experiment 2 replicated those of Experiment 1 in the effect of time on Preference ratings. Ratings of both Comfort and Visibility dropped to a statistically significant degree in both experiments (Table 95). The effect sizes were also comparable in both cases (see Table 21 for comparison).

Table 95. Time main effect on preference

					AM		PM	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	.82	5.66***	4, 100	.06				
Attractiveness		0.92	1, 103	.01				
Comfort		19.852***	1, 103	.16	69.87	11.00	65.96	10.44
Visibility		8.47**	1, 103	.08	64.31	11.26	61.17	11.62
Spaciousness		0.38	1, 103	.00				

Note. * $p<=.05$. ** $p<=.01$. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

6.2.3 Mood

The three mood scales were scored as in Experiment 1, on scales ranging from 0 – 8, and higher scores indicated stronger feelings (i.e., more pleasure, more arousal, or more dominance). Internal consistency reliability, as indicated by Cronbach's alpha, was typical of previously-obtained values, although more consistent from morning to afternoon (for Pleasure, $\alpha = .88$ AM, $\alpha = .88$ PM; for Arousal, $\alpha = .71$ AM, $\alpha = .76$ PM; for Dominance, $\alpha = .53$ AM, $\alpha = .58$ PM). These scores were analyzed in a 2 x 2 x 2 Lighting Design X Reflectance X Time MANOVA, with all main effects and interactions.

Table 96 shows the only statistically significant result, a small main effect of Time. It partially replicated the results of Experiment 1. The multivariate effect was accompanied by a statistically significant drop in Pleasure that was smaller than the one found in Experiment 1 (see Table 23 for comparison). There was also a statistically significant, small effect on Dominance, in which scores increased in the afternoon. This effect is difficult to explain as it has not been found in previous research.

In addition, these participants did not show a drop in Arousal scores in the afternoon, which is typical of previous research (e.g., Veitch & Newsham, 1998a; Veitch & Newsham, 2000a).

Table 96. Time main effect on mood.

					AM		PM	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	.82	6.72***	3, 91	.04				
Pleasure		7.90**	1, 93	.08	4.78	1.38	4.38	1.44
Arousal		0.43	1, 93	.00				
Dominance		4.65*	1, 93	.05	3.97	0.81	4.14	0.76

Note. **p*<=.05. ***p*<=.01.****p*<=.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

6.2.4 Competence

6.2.4.1 Task competence.

As in Experiment 1, participants reported in the morning a goal for the number of summaries they would process in a session; this was the only measurement of task competence. It was analyzed in a 2 x 2 (Lighting Design X Reflectance) factorial ANOVA. There were no statistically significant differences in this analysis.

6.2.4.2 Environmental competence.

Scores on the Environmental Competence scale have a range from 0 to 8, with higher scores indicating feelings of greater environmental competence (the ability to successfully alter environmental conditions). Internal consistency reliability for this four-item scale was good, although slightly lower than in Experiment 1 (Cronbach's α = .69 in the AM and α = .78 in the PM). The values for both the morning and afternoon scores on this variable were slightly negatively skewed; the median scores for both times were 6 of 8, indicating that most participants felt somewhat environmentally competent. These scores were analyzed in a 2 x 2 x 2 (Lighting Design X Reflectance x Time) factorial ANOVA. No statistically significant differences were found.

6.2.5 Health & Well-Being

6.2.5.1 Satisfaction.

There were three indicators of satisfaction measured at the end of the session. Performance Satisfaction (the individual's sense of having worked effectively) and Environmental Satisfaction (feelings that the workplace had been suitable for their work) were scored on scales from 0 –4, with higher scores indicating greater satisfaction. For Environmental Satisfaction, internal consistency reliability was very good, with Cronbach's α = .86. Overall, most participants were satisfied with the environment: Average Environmental Satisfaction was 2.75 (*SD* = 0.67), and the median was 3. For Performance Satisfaction, Cronbach's α was .86, and the overall mean was 2.86 (*SD* = 0.64).

Self-rated productivity was a single-item measure on a scale from –4 through +4. Negative scores indicate that the participant felt that the environment had reduced his or her productivity relative to other workplaces; positive scores indicate that the participant felt the environment had improved productivity relative to other workplaces. The overall mean was 0.47 (*SD* = 1.64), indicating than on average the participants believed that the Office Solutions environment improved their ability to work a small amount relative to most places where they usually work.

These three dependent variables were analyzed in a 2 x 2 (Lighting Design X Reflectance) MANOVA model. There were no statistically significant differences between the experimental conditions in these scores.

6.2.5.2 Comfort.

Recall that there were two measures of comfort, one specifically visual comfort and the other more general physical comfort. Each was the sum of severity ratings (from 0 = none, to 4 = extremely severe) on each of a list of related symptoms. Thus, high scores are more uncomfortable. For visual comfort the maximum possible score was 32, and for physical comfort it was 36. Participants completed these scales in both the morning and afternoon. Internal consistency reliability was acceptable to very good for visual comfort (Cronbach's α = .67 AM, α = .87 PM) and good for physical comfort (Cronbach's α = .75 AM, α = .79 PM).

The analysis used a 2 x 2 x 2 (Lighting Design X Reflectance X Time) MANOVA model. There were no effects involving lighting design or reflectance; the only statistically significant effect involved time (Table 97). Both visual and physical comfort got worse over the course of the day, as they had in Experiment 1 (Table 27) and previous research (Amick et al., 2002). These are large effects in terms of explained variance and of the same magnitude as in Experiment 1. The scores always remained low in relation to the possible scale maxima, indicating that in absolute terms people were never very uncomfortable.

Table 97. Time main effect on comfort.

					AM		PM	
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$	*M*	*SD*	*M*	*SD*
Multivariate	.72	20.04***	2, 102	.24				
Visual Comfort		32.74***	1, 103	.24	2.06	2.37	4.42	5.13
Physical Comfort		31.90***	1, 103	.24	2.09	2.75	3.69	4.14

Note. ***p<=.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

6.2.5.3 Task difficulty ratings.

There were six ratings of the difficulty of the tasks and one overall rating of the difficulty of all the tasks. These were analyzed in a 2 x 2 (Lighting Design X Reflectance) MANOVA. The analysis did not show any differences in the rated difficulty of the tasks.

6.2.6 Summary: Perceptions and Feelings

The results for the various measurements of Perceptions and Feelings may be summarized thus:

- **Appraisals of the lighting differentiated between the lighting designs.** This conclusion was also made following Experiment 1. The Office Lighting Survey and the open-ended responses from the NRC Lighting Quality questionnaire both showed that the Base Case 2 condition was perceived as causing more glare than the Best Practice 2, which would be expected for a recessed prismatic luminaire installation. The Office Lighting Survey results furthermore indicated that overall the Best Practice 2 installation was perceived as better than the Base Case 2.

- **The experiment was sensitive to known effects of fatigue.** Experiment 1 findings also led to this conclusion. Pleasure was lower in the afternoon, as in previous research and Experiment 1 (e.g., Veitch & Newsham, 1998a; Veitch & Newsham, 2000a). Visual and physical comfort also dropped from morning to afternoon, as expected from Experiment 1 and as consistent with the literature (Amick et al., 2002; Wibom & Carlsson, 1987). The decline in room appearance ratings over the day followed the Experiment 1 pattern.

- **Lighting and interior design choices interact in influencing room appearance.** Although the interpretation of the interaction remains speculative, the presence of an interaction between lighting design and partition reflectance and time on preference (room appearance) ratings warrants greater attention. This issue is discussed further below, where other similar interactions were also found.

The absence of simple main effects of lighting design or reflectance on mood, health, or well-being, although disappointing, is also typical of previous research in this area, including Experiment 1 above (e.g., Eklund, et al., 2000; Veitch & Newsham, 1998a). The encouraging aspect of these findings is the replication of so many statistically significant non-lighting effects in similar directions and with similar effect sizes. Only the mood scores showed a pattern that was not consistent with previous research (Veitch & Newsham, 1998a; Veitch & Newsham, 2000a), with no effect of time on Arousal, and an increase in Dominance in the afternoon. Given the use of the same experimental protocol as had been used in Experiment 1, and the broadly similar demographics of the new sample, this change in mood response is a mystery.

6.3 Performance

The results are organized, as before, by their conceptual groups, which relate to constructs in the linked mechanisms map. All analyses proceeded using the MANOVA and ANOVA strategy outlined above (section 4.1).

6.3.1 Timed Vision Test

As in Experiment 1, twice during the day, participants completed a computer-based contrast sensitivity test that was scored for both speed and accuracy. Response times and accuracy scores were combined to form a composite visual performance score: total correct/total time (summed across three target sizes) for all trials in each of six contrast levels. Thus, time and contrast were two additional within-subjects variables for the vision test analyses. The general analytic model was a 2 x 2 x 2 x 6 x 2 (Lighting Design X Reflectance X Contrast X Time) ANOVA, in which we examined main effects and interactions.

6.3.1.1 Overall contrast sensitivity.

The overall effect of target contrast on composite visual performance followed the same pattern as in Experiment 1, as seen in Figure 42 (compare to Figure 17 for Experiment 1). Performance increased as target contrast increased, but was lowest when the target was present but only faintly (gray level 1). We conducted an ANOVA for all the polynomials making up this curve; the results are shown in Table 98. As expected, and as in Experiment 1, the overall effect of contrast on composite visual performance was statistically significant, as were all of the polynomial components. All of the effects were very large, and are approximately of the same magnitude as for Experiment 1. The largest effect was the linear trend, and for further analyses of contrast level effects we focused on this effect alone, as we had done in Experiment 1. We further simplified the analyses by treating the gray-level values as being equally spaced.

Figure 42. Composite visual performance by target contrast level.

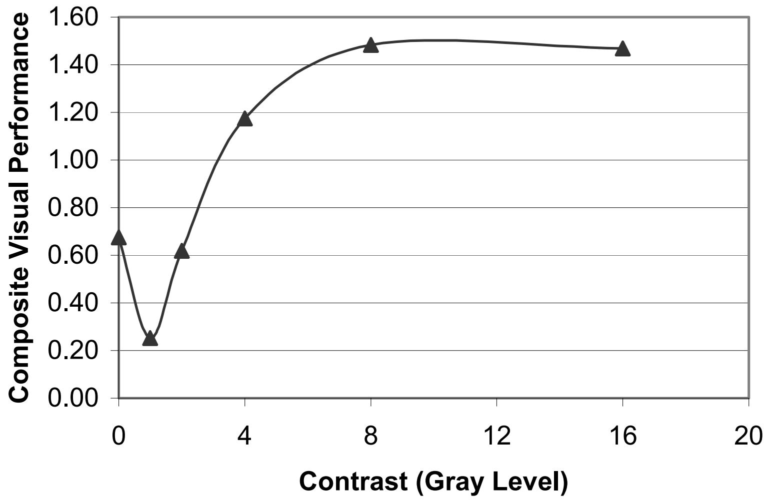


Table 98. Contrast effect on composite visual performance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Multivariate	.053	334.14***	5, 94	.69
Linear trend		1151.55***	1, 98	.92
Quadratic trend		365.95***	1, 98	.79
Cubic trend		52.19***	1, 98	.35
Quartic trend		350.69***	1, 98	.78
Quintic trend		154.41***	1, 98	.61

Note. ****p*<.001. Lighting condition, reflectance, and time were controlled for in this analysis. Contrast is weighted by gray level value.

6.3.1.2 Base Case 2 vs Best Practice 2.

There was one statistically significant, small effect involving the contrast of Base Case 2 versus Best Practice 2. It was a four-way interaction effect of Base Case vs Best Practice X Reflectance X Contrast X Time (Figure 43 and Table 99). We probed this complex effect with *post hoc* tests (Table 100). From the results of Experiment 1, we know that the contrast effect should get stronger with practice, and performance should improve with practice (see Section 4.4.1.4); that is the reading of the Contrast x Time interaction in Experiment 1. Conditions that enhance those effects will lead to better performance at high contrasts.

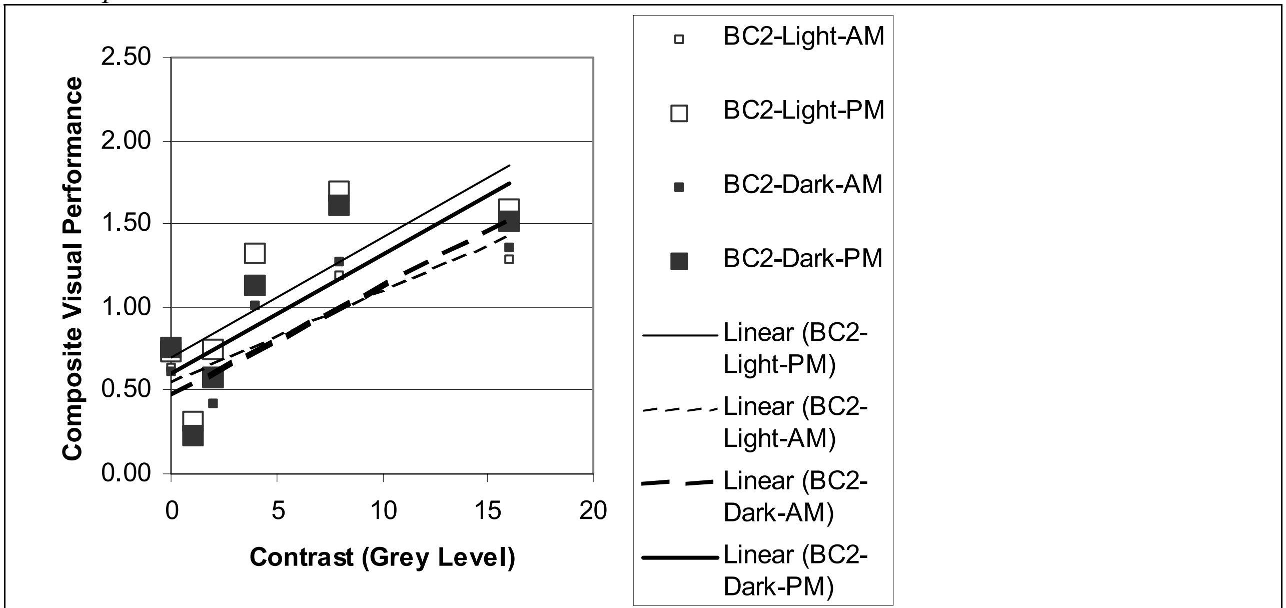
Reading the graphs in Figure 43, it appears that for Best Practice 2-Light the Contrast X Time interaction effect was smaller than for Best Practice 2 - Dark (i.e., the difference in slope from AM-PM is bigger for Best Practice 2-Dark). For Base Case 2 the effect was reversed (the difference in slope from AM-PM is bigger for light partitions). Thus, it appears that for the timed vision task, there was a small benefit to having dark partitions with direct / indirect lighting, but light partitions when the lighting system used recessed prismatic luminaires. The mean differences and the percentage of variance explained, however, are very small.

Table 99. Base Case 2 vs Best Practice 2 by Reflectance by Contrast by Time interaction effect on timed vision test.

		F	df	$\eta^2_{partial}$	
		5.02*	1, 98	.05	
		AM		PM	
		M	*SD*	*M*	*SD*
Base Case 2 – Light	0	0.64	0.38	0.73	0.35
	1	0.22	0.14	0.31	0.19
	2	0.58	0.28	0.75	0.34
	4	1.12	0.43	1.32	0.33
	8	1.19	0.38	1.69	0.32
	16	1.28	0.37	1.59	0.28
Base Case 2 – Dark	0	0.62	0.33	0.76	0.39
	1	0.20	0.16	0.23	0.15
	2	0.42	0.24	0.57	0.32
	4	1.01	0.43	1.13	0.43
	8	1.28	0.39	1.61	0.44
	16	1.36	0.41	1.51	0.35
Best Practice 2 – Light	0	0.66	0.38	0.70	0.32
	1	0.29	0.14	0.32	0.21
	2	0.67	0.26	0.84	0.32
	4	1.18	0.44	1.39	0.37
	8	1.38	0.37	1.67	0.45
	16	1.39	0.35	1.57	0.41
Best Practice 2 – Dark	0	0.57	0.31	0.73	0.35
	1	0.22	0.15	0.22	0.12
	2	0.47	0.24	0.63	0.34
	4	1.02	0.45	1.20	0.43
	8	1.26	0.39	1.78	0.46
	16	1.35	0.41	1.68	0.36

Note. *p<=.05.

Figure 43. Lighting Design X Reflectance X Contrast X Time interaction plot, with linear trend lines. Top: Base Case 2 X Time X Contrast-linear. Bottom: Best Practice 2 X Time X Contrast-linear.



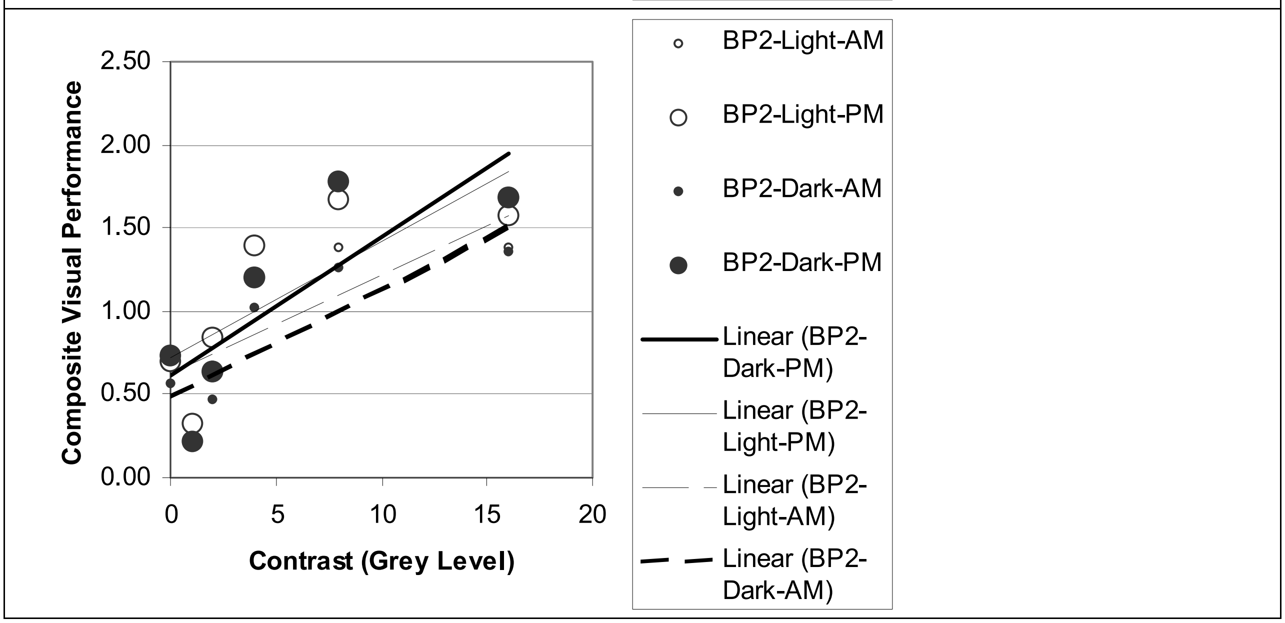


Table 100. Post hoc tests for Lighting Design X Partition Reflectance by Time-L by Contrast-L interaction effect on timed vision test.

	Slope	*F*	*df*	$\eta^2_{partial}$
Contrast - L for Base Case 2 - Light at AM		213.59***	1. 26	.89
Contrast - L for Base Case 2 - Dark at AM		150.52***	1, 22	.87
Contrast - L for Base Case 2 - Light at PM		312.00***	1, 26	.92
Contrast - L for Base Case 2 - Dark at PM		230.14***	1, 22	.91
Contrast - L for Best Practice 2 - Light at AM		164.25***	1, 25	.87
Contrast - L for Best Practice 2 - Dark at AM		172.95***	1, 25	.87
Contrast - L for Best Practice 2 - Light at PM		251.87***	1, 25	.91
Contrast - L for Best Practice 2 - Dark at PM		238.23***	1, 15	.91

Note. ** p<=.01. ***p<=.001.

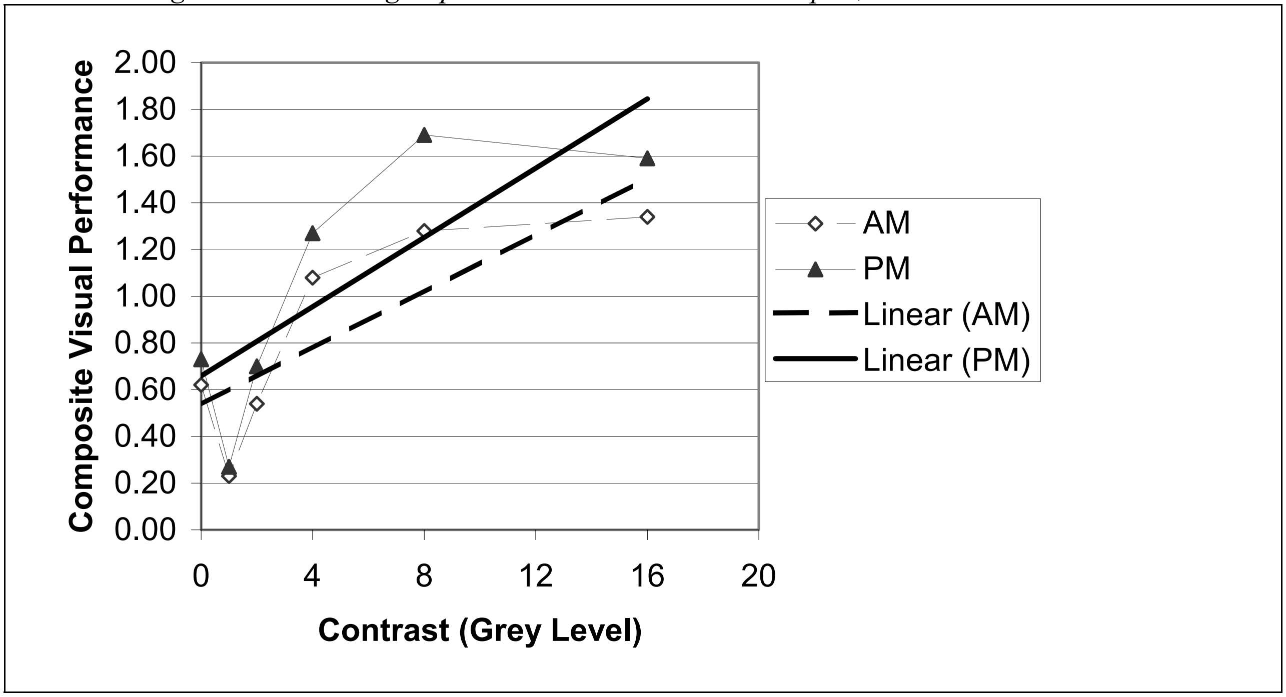
6.3.1.3 Non-lighting effects.

There were three statistically significant effects that did not involve lighting variables – that is, these effects apply regardless of lighting condition. As in Experiment 1, there were large main effects of Contrast and Time, and a large interaction effect of Contrast X Time. These effects replicated the effects of Experiment 1 in both direction and size. Table 101 summarizes the statistics for these effects, and Figure 44 illustrates them graphically.

Table 101. Contrast -Linear and Time effects on timed vision test.

Between-groups design	*F*	*df*	$\eta^2_{partial}$			
Contrast X Time Interaction	52.93***	1, 98	.35			
Contrast Main Effect	1043.45***	1, 98	.91			
Time Main Effect	101.15***	1, 98	.51			
	AM		PM		Contrast Effect	
Contrast (Grey Level)	*M*	*SD*	*M*	*SD*	*M*	*SD*
0	0.62	0.35	0.73	0.35	0.68	0.35
1	0.23	0.15	0.27	0.18	0.25	0.16
2	0.54	0.27	0.70	0.34	0.62	0.32
4	1.08	0.44	1.27	0.40	1.18	0.43
8	1.28	0.38	1.69	0.42	1.48	0.45
16	1.34	0.38	1.59	0.35	1.47	0.39
Time Main Effect	0.85	0.53	1.04	0.62		

Figure 44. Between-groups Contrast X Time interaction plot, with linear trend lines.



Note. Solid line links the average performance at each contrast level. Trend lines (dotted) show the linear effect of contrast = 0 through contrast = 16 in the morning and afternoon.

6.3.2 Motivation

Motivation, as in Experiment 1, was measured using the Speed of Random Responding on the persistence task (see section 4.4.2 for a full description of this variable). The distribution of scores was very skewed, as expected from the Experiment 1 results. Over all times and lighting and reflectance conditions, the mean score was 620.97 (*SD* = 197.64), and the median was 729, indicating that half of the scores were the maximum value. These values are almost identical to the descriptive statistics for Experiment 1.

A 2 x 2 x 3 (Lighting Design X Reflectance X Time) ANOVA was conducted. There were no statistically significant effects. The absence of effects involving lighting design and reflectance was not surprising, given that in Experiment 1 the only effects involved comparisons in which one group had some form of control over the lighting. However, the Time-Quadratic main effect seen in the between-groups design in Experiment 1 did not replicate. In Experiment 2, persistence scores were consistent from one session to another.

6.3.3 Vigilance

Vigilance was measured in the same way as it had been in Experiment 1, with two dependent measures: the hit rate on the conveyor belt task at intermediate speeds, and the speed of responding to an envelope prompt. Both measures occurred in three sessions, making the analysis a 2 x 2 x 3 Lighting Design X Reflectance X Time MANOVA. There were no statistically significant effects involving lighting design or reflectance. This was not surprising, given that the only such effects in Experiment 1 involved conditions with individual control.

There were two statistically significant multivariate effects of Time, the linear and the quadratic trend. Both were associated with statistically significant effects on the Envelope Speed variable, as had been the case in Experiment 1. Table 102 summarizes these effects, which are quite large. The size of the quadratic effect is the same as its Experiment 1 counterpart; the linear effect is larger. Overall, there was a drop in the speed of responding to the envelope prompt, but there is evidence of recovery at the end of the day following a post-prandial dip. This pattern is the same as it was in Experiment 1.

Table 102. Time effects on vigilance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Time – Linear Multivariate	.76	12.77***	2, 79	.12		
Hit Rate		0.05	1, 80	.00		
Envelope Speed		25.21***	1, 80	.24		
Time-Quadratic Multivariate	.54	33.52	2, 79	.23		
Hit Rate		1.87	1, 80	.02		
Envelope Speed		59.79***	1, 80	.43		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Envelope Speed	0.41	0.13	0.28	0.11	0.33	0.14

Note. ***p*<=.01. ****p*<=.001. Group descriptive statistics are provided only for the statistically-significant univariate tests. For multivariate tests, the $\eta^2_{partial}$ indicator of effect size is the average of all the univariate values. It is the proportion of variance explained, with a range from 0 – 1.

6.3.4 Typing

Recall that there was one dependent variable for the transcription typing task: correct characters per second. In addition to the lighting design and time variables (three sessions), there were three print sizes (8, 12, and 16 point) for the source text. Thus, the statistical model was a 2 x 2 x 3 x 3 Lighting Design X Reflectance X Print Size X Time ANOVA.

6.3.4.1 Base Case 2 vs Best Practice 2.

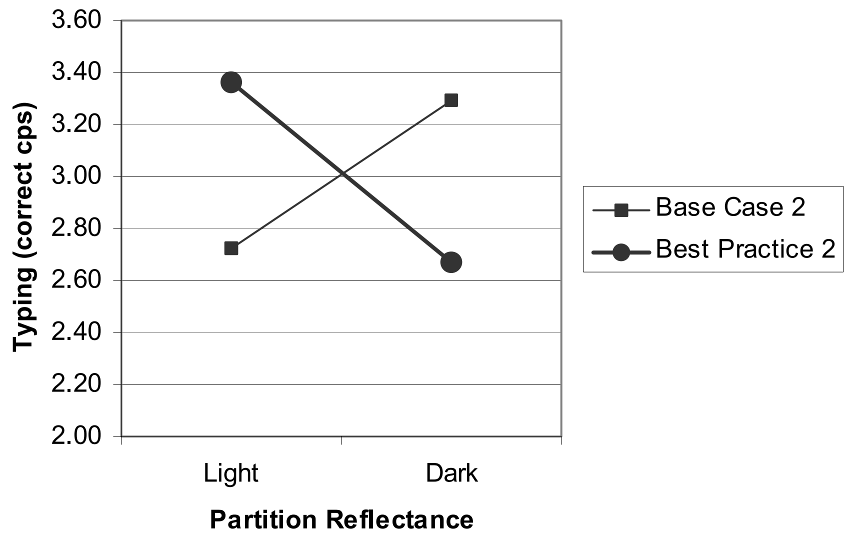
There were two, separate, statistically significant interactions involving the comparison between Base Case 2 and Best Practice 2, one involving partition reflectance and the other involving time. Table 103 summarizes the interaction of Base Case 2 vs Best Practice 2 by Reflectance, and Figure 45 illustrates it. *Post hoc* tests are included in Table 103. The effect sizes are small to medium. The *post hoc* tests revealed that for Base Case 2, there was no difference between performance with light and dark partitions; whereas, for Best Practice 2, performance was better with light partitions than dark ones. Interestingly, with light partitions, performance was better under Best Practice 2, but there was no difference between the lighting conditions when the partitions were dark.

Table 103. Base Case 2 vs Best Practice 2 by Reflectance interaction effect on typing.

	F	*df*	$\eta^2_{partial}$		
	7.79**	1, 97	.07		
	Light Partitions		Dark Partitions		Reflectance *post hoc* tests
	M	*SD*	*M*	*SD*	
Base Case 2	2.72	1.17	3.29	1.26	$F(1, 47) = 2.87$
Best Practice 2	3.36	1.24	2.67	0.98	$F(1, 50) = 5.23$*, $\eta^2_{partial} = .09$
Lighting Design *post hoc* tests	$F(1, 51) = 3.98$*, $\eta^2_{partial} = .07$		$F(1, 46) = 3.88$		

Note. **p*<=.05. ***p*<=.01. ****p*<=.001

Figure 45. Base Case 2 vs Best Practice 2 by Reflectance interaction effect on typing.



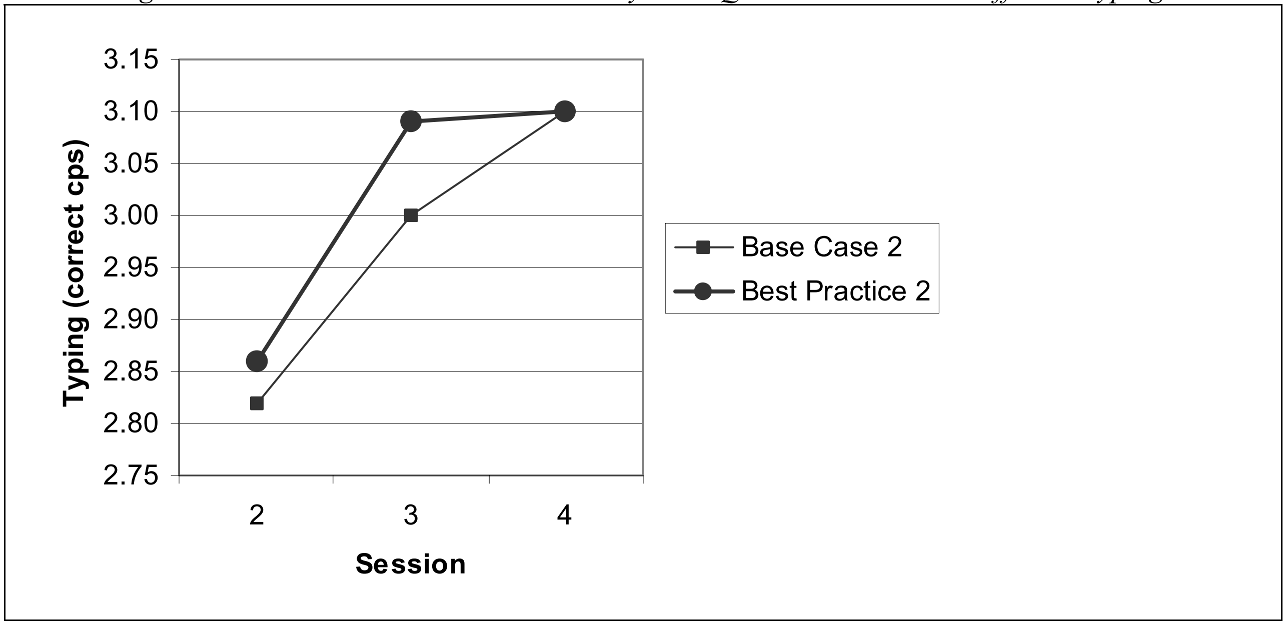
The other statistically significant interaction was the Base Case 2 vs Best Practice 2 X Time-Quadratic interaction, summarized in Table 104 and Figure 46. The *post hoc* tests revealed that there was no quadratic effect for Base Case 2, but there was a large quadratic effect for Best Practice 2. The two groups ended up at the same place, and indeed did not show statistically significant between-groups differences at any session. One interpretation of this pattern might be that learning took place faster for Best Practice 2; they appear to have reached a plateau at Session 3 that the others did not reach until Session 4.

Table 104. Base Case 2 vs Best Practice 2 by Time-Quadratic interaction effect on typing.

	F	*df*	$\eta^2_{partial}$				
	4.58*	1, 97	.05				
	Session 2		Session 3		Session 4		Time-Q *post hoc* tests
	M	*SD*	*M*	*SD*	*M*	*SD*	
Base Case 2	2.82	1.22	3.00	1.23	3.10	1.27	$F(1, 48) = 2.74$
Best Practice 2	2.86	1.12	3.09	1.17	3.10	1.21	$F(1, 51) = 18.09$***, $\eta^2_{partial} = .26$
Lighting Design *post hoc* tests	$F(1, 99) = 0.03$		$F(1, 99) = 0.16$		$F(1, 99) = 0.00$		

Note. *p<=.05. **p<=.01. ***p<=.001

Figure 46. Base Case 2 vs Best Practice 2 by Time-Quadratic interaction effect on typing.



6.3.4.2 Reflectance effects.

There were also two statistically significant interactions involving Reflectance X Print Size-Linear X Time (both Time-Linear and Time-Quadratic). These are summarized in Table 105 and illustrated in Figure 47. The Reflectance X Print Size-Linear X Time - Linear interaction is medium-sized; the Reflectance X Print Size -Linear X Time - Quadratic effect is small.

Tables 106 and 107 summarize the *post hoc* tests conducted for these interactions. For Light partitions, the linear effect of Print Size was statistically significant at all three times and became larger as the day went on; however, for Dark partitions there was no linear effect of Print Size in Session 4, although there had been large ones in Sessions 2 and 3. Turning the data around, for Light partitions there was no linear effect of Time for 8 point source text, whereas there were large linear effects of Time for 12 point and 16 point source text. For Dark partitions, there were linear effects of Time for all three print sizes, with larger effect sizes for smaller print. From Figure 47, we see that the performance of people with Light partitions with 8 point font at session 4 was lower than we might expect (there is no continued increase from session 3 to session 4), whereas for Dark partitions the odd point out is the mean for 16 point source text at session 4. This accounts for the significant quadratic trends for these two lines, seen in table 107.

This is a puzzling pair of interactions, for which we have only a speculative interpretation. Perhaps, when the visual task was demanding (as it was for the 8 point source text), and when fatigue was developing at the end of a day's work, light partitions behind the white paper did not provide sufficient contrast to focus attention on the visual task. Conversely, perhaps when the contrast was greater, as between the dark partitions and the light paper, it created a distraction that was evident when the task was less demanding. This would argue for a balance to be struck between surrounds that are very low or very high in contrast to the visual task (which is usually very light).

Table 105. Partition Reflectance by Print Size by Time interaction effect on typing performance.

	F	df	$\eta^2_{partial}$			
Reflectance X Print Size-L X Time - Linear	8.56**	1, 97	.08			
Reflectance X Print Size -L X Time - Quadratic	4.52*	1, 97	.04			
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Light – 8 point	2.80	1.14	3.02	1.28	2.90	1.18
Light – 12 point	2.90	1.28	3.08	1.24	3.21	1.28
Light – 16 point	2.92	1.24	3.17	1.25	3.29	1.36
Dark – 8 point	2.71	1.12	2.86	1.10	3.05	1.33
Dark – 12 point	2.79	1.10	3.01	1.18	3.05	1.18
Dark– 16 point	2.92	1.16	3.12	1.12	3.09	1.11

Note. **p<=.01.

Table 106. Post hoc tests for Partition Reflectance by Print Size-L by Time-L interaction effect on typing.

	F	df	$\eta^2_{partial}$
Light Print Size – L at Session 2	3.91*	1, 52	.07
Light Print Size – L at Session 3	4.81*	1, 52	.08
Light Print Size – L at Session 4	21.81***	1, 52	.30
Dark Print Size – L at Session 2	15.63***	1, 47	.25
Dark Print Size – L at Session 3	22.74***	1, 47	.33
Dark Print Size – L at Session 4	0.19	1, 47	
Light Time – L at 8 point	1.86	1, 52	
Light Time – L at 12 point	21.13***	1, 52	.29
Light Time – L at 16 point	27.50***	1, 52	.35
Dark Time – L at 8 point	21.50***	1, 47	.31
Dark Time – L at 12 point	10.55**	1, 47	.18
Dark Time – L at 16 point	5.51*	1, 47	.10

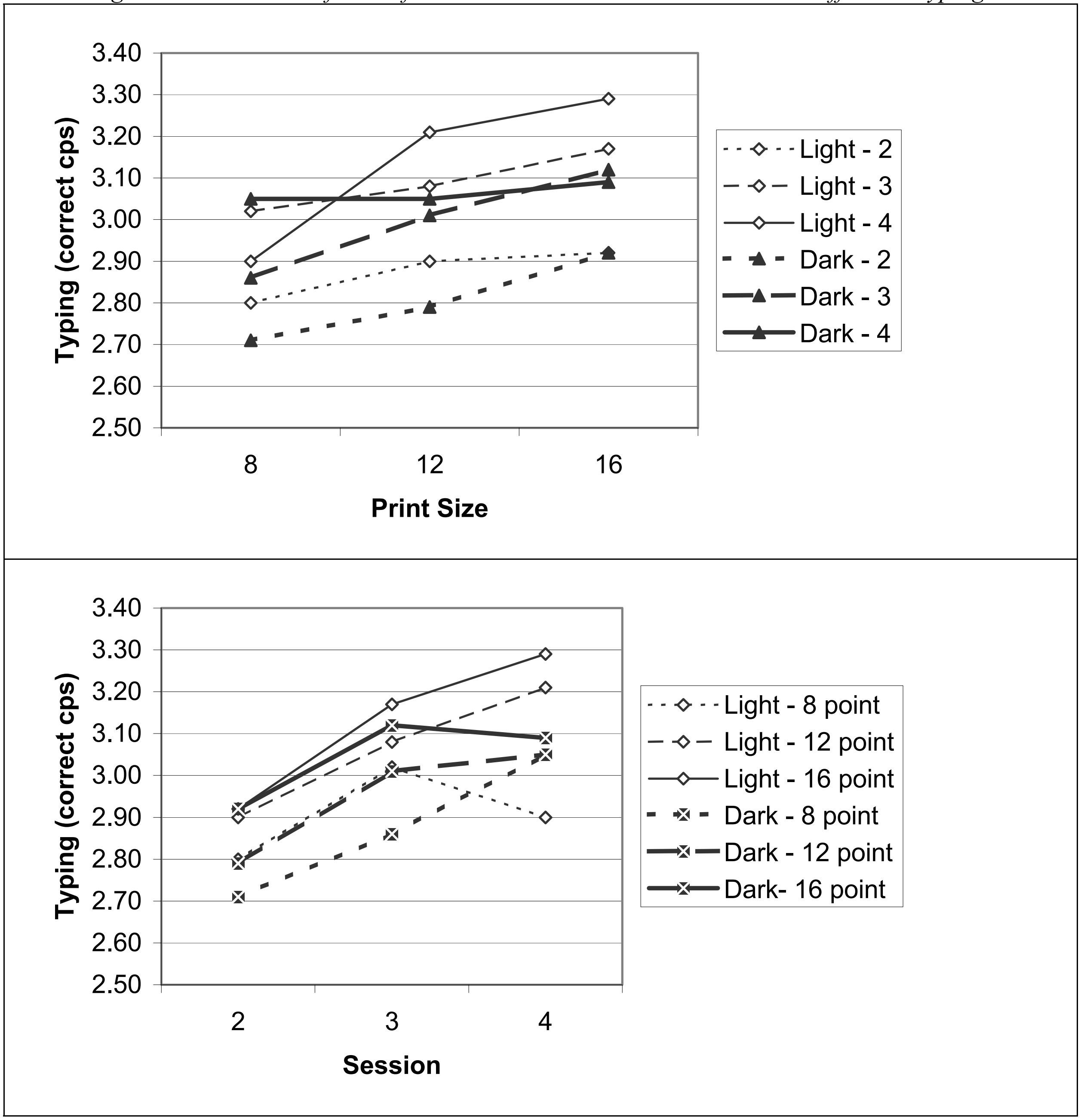
Note. ** p<=.01. ***p<=.001.

Table 107. Post hoc tests for Partition Reflectance by Print Size-L by Time-Q interaction effect on typing.

	F	df	$\eta^2_{partial}$
Light Time – Q at 8 point	7.60**	1, 52	.13
Light Time – Q at 12 point	0.41	1, 52	
Light Time – Q at 16 point	1.54	1, 52	
Dark Time – Q at 8 point	0.09	1, 47	
Dark Time – Q at 12 point	3.27	1, 47	
Dark Time – Q at 16 point	7.68**	1, 47	.14

Note. ** p<=.01. ***p<=.001.

Figure 47. Two views of the Reflectance X Print Size x Time interaction effects on typing.



Note. Top panel shows the interaction effect plotted against Print Size. Bottom panel shows the same data reorganized to display against session.

6.3.4.3 Non-lighting effects.

There were three statistically significant main effects that did not involve lighting: Time-Linear and Time-Quadratic, and Print Size-Linear. These were similar to their counterparts in Experiment 1. Both effects of Time were larger than they were in Experiment 1. The effect of Print Size was almost exactly the same size. Experiment 2 did not show the Print Size-Quadratic X Time-Quadratic effect that had been seen in Experiment 1.

The effects are summarized in Table 108. Typing performance increased over the day, but faster between Sessions 2 and 3 than between 3 and 4. Typing performance was better for larger print sizes. These effects were expected, based on existing knowledge of learning and visibility.

Table 108. Time and Print Size effects on typing.

	F		df	$\eta^2_{partial}$		
Print Size – L Main Effect	49.80***		1, 97	.34		
Time – L Main Effect	79.69***		1, 97	.45		
Time – Q Main Effect	18.59***		1, 97	.16		
	Session 2		Session 3		Session 4	
	M	SD	M	SD	M	SD
Main Effect	2.84	1.17	3.05	1.19	3.10	1.24
	8 pt		12 pt		16 pt	
	M	SD	M	SD	M	SD
Main Effect	2.89	1.19	3.01	1.21	3.08	1.21

Note. *p<=.05. **p<=.01. ***p<=.001. Descriptive statistics are provided for statistically significant effects only.

6.3.5 Cognitive Judgments

6.3.5.1 Paper-based articles.

We calculated average scores for the ratings of articles read prior to the group discussion task: how easy it was to read (-3 to +3), how interesting the participant found it (-3 to +3), how relevant it might be to the (fictitious) magazine in which it was ostensibly intended to appear (-3 to +3), and how likely it would be that the group would choose that article to be the lead in the magazine (0-100 % probability). Internal consistency reliability for these scales ranged from was just acceptable to good: For Interest, Cronbach's alpha = .57; for Relevance, α = .62; for Ease, α = .72; for Probability, α = .66.

We tested the hypothesis that lighting conditions might alter the judgments. The model was a 2 x 2 (Lighting Design X Reflectance) MANOVA with the four judgments as dependent variables. There were no statistically significant differences in these cognitive judgments for either analysis.

6.3.5.2 Computer-based articles.

These dependent variables were derived from the categorizing and summarizing tasks. The averages of articles judged in the first six trials in each session were used to create scores for the interest in reading the whole article based on a summary (0-100); the judgment that the summary includes the important facts in the article; the judgment that the summary is grammatically correct; and, the judgment that the summary is well-written (the latter three all on scales from 0 – 4). Internal consistency reliability ranged from poor (Cronbach's α < .50) to good (Cronbach's α > .70) (for Interest, session 2 α = .74; session 3, α = .66, session 4, α = .54; for Important Facts, session 2 α = .70; session 3, α = .49, session 4, α = .57; for Grammatically Correct, session 2 α = .66; session 3, α = .71, session 4, α = .64; for Well-written, session 2 α = .67; session 3, α = .57, session 4, α = .56).

These four judgments were entered in a 2 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA. There were no statistically significant differences involving lighting design or reflectance.

There were both Time-Linear and Time-Quadratic effects on cognitive judgments. These are summarized in Table 109 and illustrated in Figure 48. They mirrored the effects seen in Experiment 1 in both direction and size. Interest in the articles declined over the day, but the biggest drop was between sessions 2 and 3. Ratings concerning how factual and well-written were the summaries tended to increase slightly over the day, but, as importantly, they showed a quadratic shape, dropping in session 3 right after lunch and then rebounding in session 4. The rating of the summaries' grammar had no linear effect, only the quadratic shape.

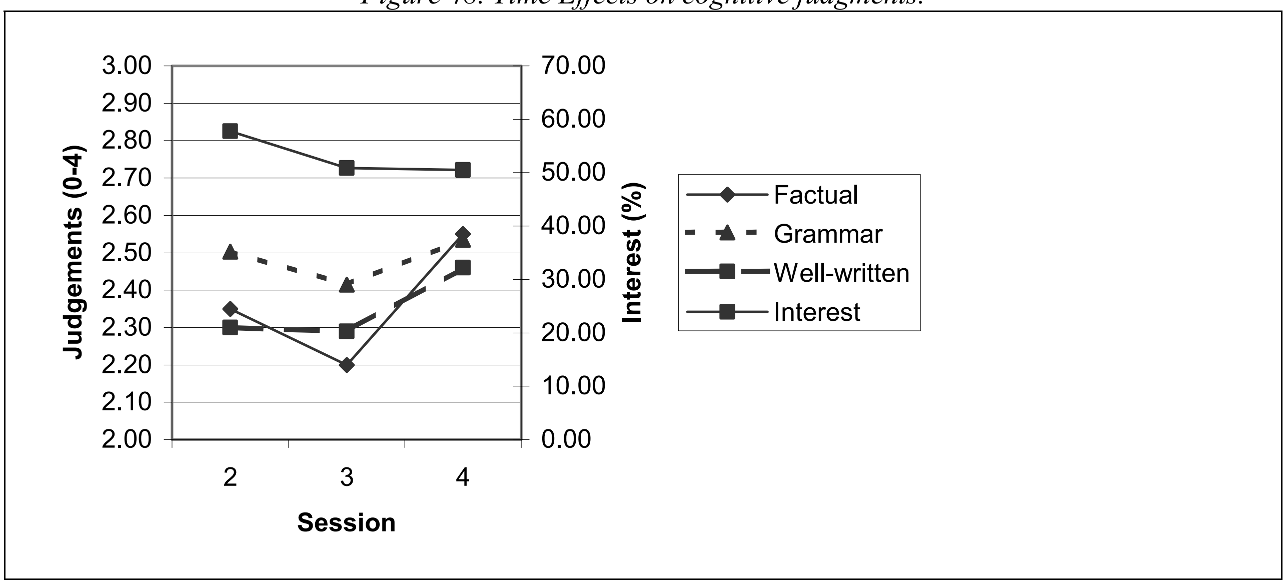
Table 109. Time effects on cognitive judgments.

Between-groups design	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Linear: Multivariate	.74	9.01	4, 100	.09
Interest		21.43***	1, 103	.17
Factual		10.68***	1, 103	.09
Grammar		0.45	1, 103	.00
Well-written		9.60**	1, 103	.09
Quadratic: Multivariate	.73	9.45***	4, 100	.10
Interest		5.30*	1, 103	.05
Factual		35.60***	1, 103	.26
Grammar		5.72*	1, 103	.05
Well-written		5.26*	1, 103	.05

	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Interest	57.77	17.74	50.88	17.60	50.52	14.98
Factual	2.35	0.66	2.20	0.53	2.55	0.51
Grammar	2.50	0.55	2.41	0.58	2.53	0.51
Well-written	2.30	0.60	2.29	0.52	2.46	0.48

Note. *p<=.05. **p<=.01. ***p<=.001.

Figure 48. Time Effects on cognitive judgments.



6.3.6 Cognitive Performance

6.3.6.1 Speed and accuracy.

There were four dependent measures associated with cognitive tasks: an accuracy score for the assignment of summaries to content categories (# correct per session, squared to improve the shape of the distribution), two speed measures from the categorization task: the time taken to do the categorization, and the time required to make the four judgments reported above, and the time taken on the Summarization task. Average scores for these dependent variables were calculated for the first six trials in each of three sessions. The statistical model was thus a 2 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA. There was a lighting effect and several non-lighting effects.

6.3.6.1.1 Base Case 2 vs Best Practice 2.

The statistically significant lighting effect was an interaction of Base Case 2 vs Best Practice 2 X Reflectance X Time-Quadratic. The multivariate effect was small (see Table 110), and was associated with a significant univariate effect on categorization accuracy. This interaction effect is illustrated in Figure 49. *Post hoc* tests were conducted, and are reported in Table 111.

In general, categorization accuracy increased over time, but for the Best Practice 2 - Dark condition there was a marked quadratic effect. Performance in this group dropped from Session 2 to Session 3, and then rebounded in Session 4. Moreover, the *post hoc* tests revealed that in Session 2, the Best Practice 2 - Dark condition did significantly better than the Base Case 2 - Dark, but the reverse was true in Session 3. By Session 4, all groups performed equally well. Only the Best Practice 2 - Dark condition showed a post-prandial dip in categorization accuracy.

Table 110. Lighting Design x Reflectance x Time-Quadratic interaction effect on cognitive performance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Multivariate	.91	2.58*	4, 100	.03		
Categorization Time		2.81	1, 103	.03		
Categorization Accuracy		7.91**	1, 103	.07		
Judgment Time		1.61	1, 103	.02		
Summarization Time		0.57	1, 103	.01		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Base Case 2 – Light	17.79	7.29	22.69	7.96	33.41	5.38
Base Case 2 – Dark	16.22	6.52	24.74	7.58	31.30	6.26
Best Practice 2 – Light	19.90	7.50	24.21	9.03	32.97	5.00
Best Practice 2 - Dark	20.62	7.52	19.65	8.04	33.96	5.12

Note. *p<=.05. **p<=.01.

Figure 49. Lighting Design by Reflectance by Time-Quadratic effect on Categorization Accuracy.

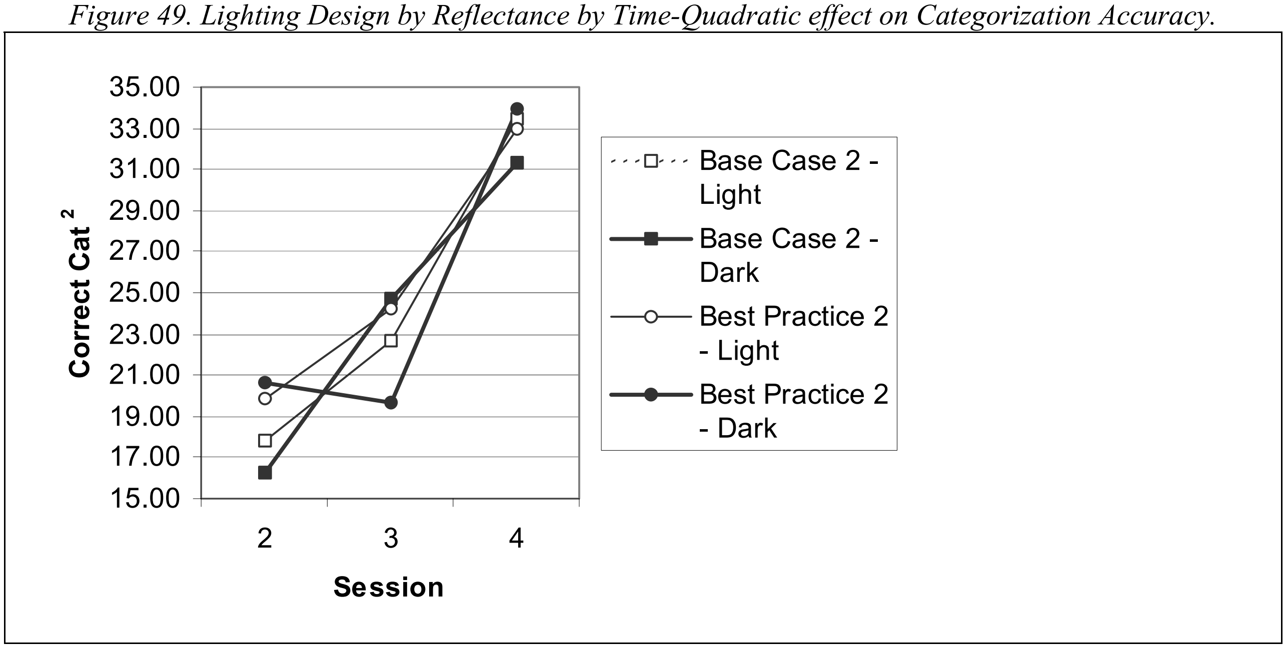


Table 111. Post hoc tests for Lighting Design by Reflectance by Time effect on Categorization Accuracy

	F	*df*	$\eta^2_{partial}$
Base Case 2 – Light – Time-quadratic	3.34	1, 28	
Base Case 2 – Dark – Time-quadratic	0.43	1, 22	
Best Practice 2 – Light – Time-quadratic	1.56	1, 28	
Best Practice 2 – Dark – Time-quadratic	22.32***	1, 25	.47
Base Case 2 – Light vs Best Practice 2 – Light - Session 2	1.17	1, 56	
Base Case 2 – Light vs Best Practice 2 – Dark - Session 2	2.00	1, 53	
Base Case 2 – Dark vs Best Practice 2 – Light – Session 2	3.46	1, 50	
Base Case 2 – Dark vs Best Practice 2 – Dark – Session 2	4.72*	1, 47	.09
Base Case 2 – Light vs Base Case 2 – Dark - Session 2	0.66	1, 50	
Best Practice 2 – Light vs Best Practice 2 – Dark - Session 2	0.13	1, 53	
Base Case 2 – Light vs Best Practice 2 – Light - Session 3	0.46	1, 56	
Base Case 2 – Light vs Best Practice 2 – Dark - Session 3	1.98	1, 53	
Base Case 2 – Dark vs Best Practice 2 – Light – Session 3	0.05	1, 50	
Base Case 2 – Dark vs Best Practice 2 – Dark – Session 3	5.15*	1, 47	.10
Base Case 2 – Light vs Base Case 2 – Dark - Session 3	0.89	1, 50	
Best Practice 2 – Light vs Best Practice 2 – Dark - Session 3	3.86	1, 53	
Base Case 2 – Light vs Best Practice 2 – Light - Session 4	0.11	1, 56	
Base Case 2 – Light vs Best Practice 2 – Dark - Session 4	0.15	1, 53	
Base Case 2 – Dark vs Best Practice 2 – Light – Session 4	1.13	1, 50	
Base Case 2 – Dark vs Best Practice 2 – Dark – Session 4	2.67	1, 47	
Base Case 2 – Light vs Base Case 2 – Dark - Session 4	1.71	1, 50	
Best Practice 2 – Light vs Best Practice 2 – Dark - Session 4	0.53	1, 53	

Note. * $p<=.05$. ** $p<=.01$. *** $p<=.001$.

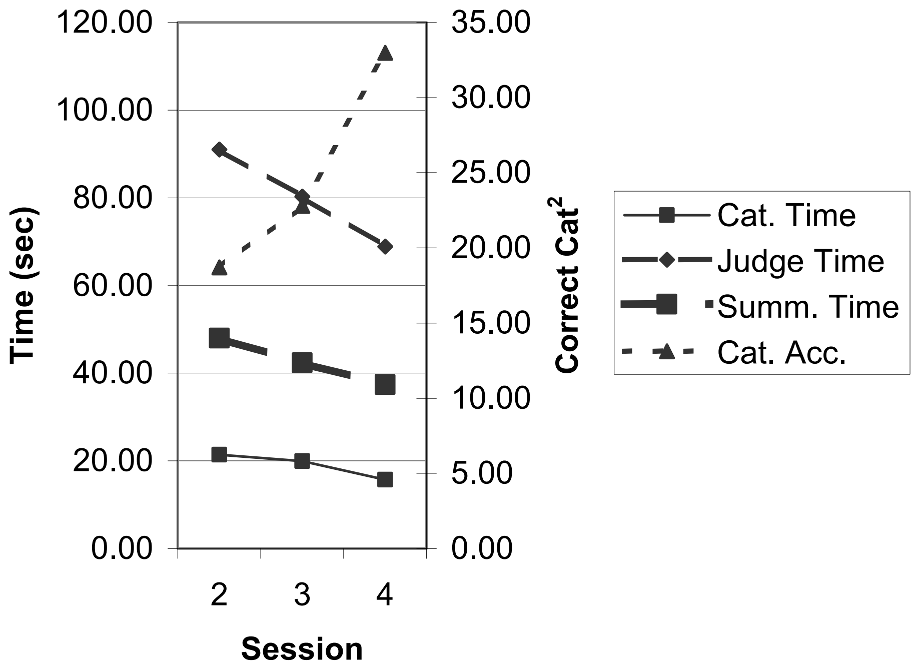
6.3.6.1.2 Non-lighting effects.

The non-lighting effects on cognitive performance were the same in Experiment 2 as they had been in Experiment 1 (see Section 4.4.6.1.2, Figure 34, top panel). The Time-Linear effect was statistically significant and large for all four dependent measures (Table 112 and Figure 50). Participants became more accurate at categorizing and faster on all measures, as the day wore on. Two outcomes also showed Time-Quadratic effects consistent with a post-prandial dip. Categorization accuracy increased less from Session 2 to Session 3 than it did from Session 3 to Session 4, and there was a similar, although smaller, interruption in the change in the time it took to do the categorization.

Table 112. Time effects on cognitive performance.

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$		
Linear: Multivariate	.18	113.41***	4, 100	.51		
Cat. Time		99.04***	1, 103	.49		
Cat. Acc.		335.29***	1, 103	.77		
Judge Time		108.98***	1, 103	.51		
Summ. Time		34.98***	1, 103	.25		
Quadratic: Multivariate	.80	6.44***	4, 100	.04		
Cat. Time		7.05**	1, 103	.06		
Cat. Acc.		12.71***	1, 103	.11		
Judge Time		0.00	1, 103	.00		
Summ. Time		0.80	1, 103	.00		
	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Cat. Time	21.41	9.68	19.96	8.63	15.69	7.81
Cat. Acc.	18.71	7.34	22.80	8.32	32.97	5.43
Judge Time	90.99	35.53	80.27	31.12	68.87	28.14
Summ. Time	47.99	24.10	42.37	20.04	37.43	17.99

Figure 50. Time Effects on cognitive performance.



6.3.6.2 Work structure.

Recall that most of the tasks occurred in sets of several trials, and that for three tasks we collected data on the length of time between trials. This was the period of time during which the introductory screen for the new trial appeared on the computer. For the typing and conveyor belt tasks this time amounted to a rest break, as there was no work to be done. For the summarizing task, the time combined a brief rest and the time required to read the article to be summarized in the new trial. In all three cases the measurement was in seconds, and average scores were calculated for each session (across three trials in each session for typing, four for conveyor, and six for summaries). These were transformed to log values to improve the shapes of the distributions. For Experiment 2 the statistical model was a 2 x 2 x 3 (Lighting Design X Reflectance X Session) MANOVA with the three break durations as dependent variables. There were no lighting or reflectance effects on these outcomes in Experiment 2.

The only statistically significant effect on work structure was a Time-Linear effect. The significant multivariate effect was accompanied by univariate effects on all three dependent measures. The break times declined from Session 2 through Session 4 in all cases (Table 113). The effects were large for the Typing and Summaries breaks, and medium-sized for the Conveyor Belt breaks. In general this replicated the non-lighting effects seen in Experiment 1, although the Conveyor Belt size was smaller (see Section 4.4.6.2.3).

Table 113. Time-Linear effects on log(average break duration).

	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Linear: Multivariate	.52	31.76***	3, 101	.23
Typing		55.14***	1, 103	.35
Conveyor		9.45**	1, 103	.08
Summaries		35.16***	1, 103	.25

	Session 2		Session 3		Session 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
Typing	1.24	0.22	1.19	0.35	1.05	0.23
Conveyor	0.60	0.30	0.52	0.34	0.47	0.40
Summaries	1.75	0.31	1.69	0.44	1.52	0.50

Note. ***p<=.001.

6.3.7 Summary: Performance

The results for the various performance measures may be summarized thus:

- **Visibility, practice, and fatigue effects were found as expected.** With few exceptions, the non-lighting effects in Experiment 2 replicated those of Experiment 1. The effect of target contrast on composite visual performance conformed to our knowledge of the visual system, as did the effect of print size on transcription typing performance. Performance improvements over the day were seen, as expected for the tasks involving novel features, including the speed and accuracy of the categorization task, the typing task, and the composite visual performance measure. As in Experiment 1, some of these linear effects were tempered over time as various processes interacted, adding a quadratic shape to the time effects.

- **Indirect systems interact with partition reflectance.** Several statistically significant effects were found involving interactions of lighting design and partition reflectance, but their interpretation is challenging. The four-way interaction effect on the timed vision test was very difficult to interpret, but appeared to suggest that dark partitions are more desirable with Best Practice 2 that light partitions, whereas light partitions are more desirable with Base Case 2 than dark. Conversely, however, the two-way interaction effect on typing performance, showed a fairly clear benefit for Best Practice 2 with light partitions, rather than dark. The interaction effect with time on the categorization accuracy further suggested that the Best Practice 2 with dark partitions could create undesirable conditions at least some of the time.

Overall, the results for the performance measures are consistent with the published literature and with Experiment 1. In Experiment 1, the results showed few statistically significant comparisons between Base Case and Best Practice for measures of performance, and when there were, their interpretation was ambiguous (see section 4.4.7). Looking to the scientific literature, there are few simple main effects of lighting distribution on performance measures of this kind (Veitch, 2001). For example, Veitch and Newsham (1998a) found no effects of lighting distribution on typing performance that involved direct / indirect or indirect-only systems; nor did Eklund, et al. (2000). However, the present experiments had more statistical power because of their larger sample size, and also differed from the others in including the partition reflectance variable, which allowed interactions to be revealed.

6.4 Social Behavior

6.4.1 Liking

The scores on the five-question scale of liking for their fellow participants were averaged to provide one score, on a scale of –3 through +3, which showed acceptable internal consistency reliability ($\alpha = .71$). The overall mean was 1.28 ($SD = 0.75$), indicating that participants in general felt somewhat positively towards their fellow employees.

Data from this scale score was analyzed in a 2 x 2 (Lighting Design by Reflectance) ANOVA. There were no statistically significant effects; liking for other group members was unrelated to lighting or partition reflectance conditions.

6.4.2 Conflict Resolution

Recall that the conflict resolution task provided two types of scores, five ratings and five rankings. The ratings were the average (over five scenarios) likelihood that an individual would resolve conflicts using competition, collaboration, compromise, avoidance, or accommodation strategies, on a scale from –3 through +3, where higher scores indicated greater likelihood of using that strategy. The rankings were the average rank assigned to each of the five strategies over the various scenarios. Table 114 summarizes the characteristics of these two sets of variables. For some of the scales, internal consistency reliability is acceptable, but for others it is unacceptably low. Overall, this task requires further development and refinement as a measure of these tendencies.

Table 114. Conflict resolution scale characteristics for Experiment 2

	Cronbach's alpha	M	SD
Ratings			
Competition	.54	0.28	0.94
Accommodation	.28	0.32	0.72
Avoidance	.43	-1.17	0.81
Collaboration	.66	1.14	0.91
Compromise	.49	1.65	0.69
Rankings			
Competition	.36	3.45	0.61
Accommodation	-.23	3.09	0.45
Avoidance	.45	4.28	0.59
Collaboration	.50	2.28	0.63
Compromise	.37	1.87	0.52

Note. Each scale is the average score for that concept rated on 5 conflict scenarios. Ratings indicate the absolute likelihood that the individual will use that strategy to resolve conflict, and are on a scale from –3 through +3. Rankings indicate the relative likelihood of using that strategy in relation to the others, and are on a scale from 1 through 5.

The ratings and the rankings were analyzed in separate MANOVA models (e.g., either 5 ratings or 5 rankings were the dependent variables), in each case a 2 x 2 (Lighting Design by Reflectance) MANOVA. There were no statistically significant effects in these analyses.

6.4.3 Willingness to Volunteer

The scores on two questions related to the participants' willingness to volunteer were standardized and then averaged to form an overall score. One question concerned the number of unpaid hours (0 through 10) that the individual would be willing to put in at home to complete questionnaires as part of other research. The other question concerned their willingness to return for a second session if requested (a probability, 0 through 100). As one would expect for a standardized variable, the mean was 0.00 and the standard deviation was 0.88 (i.e., close to 1). This single score was analyzed in a 2 x 2 (Lighting Design by Reflectance) ANOVA. There was no effect of lighting design or partition reflectance on willingness to volunteer.

6.4.4 Summary: Social Behavior

Based on other environmental psychology research, we had predicted that social behaviors would show differences associated with positive affect (e.g., Baron, Rea, & Daniels, 1992). People in a state of positive affect should rate other people more positively, resolve conflicts with more collaborative or compromising strategies, and be more willing to volunteer. As in Experiment 1, these lighting and partition reflectance conditions did not exert sufficient influence on positive affect to lead to the expected effects on social behavior.

6.5 Linked Mechanisms Tests

6.5.1 Analytic Strategy

We repeated exactly the set of linked mechanisms tests that we had conducted for Experiment 1. This is an independent replication of the tests because the participants in Experiment 2 were different people from those who had taken part in Experiment 1. The best comparison of these results is the between-groups analyses in Experiment 1, with its larger, independent sample.

To recap the method, we tested each link using non-randomly formed groups in an analysis of variance approach. For each antecedent variable we formed two groups based on a split around the median (50^{th} percentile) value for this sample. We determined that the distribution of people into the above- and below-median groups was independent of lighting condition, using chi-squared tests of independence. The median split groups were equally distributed across lighting conditions. This is not surprising, given the results reported above. We then tested the effect of being in the high or low group, on the outcomes linked to the consequent variable at that link. The original linked mechanisms map, showing the links that we tested, is shown in Figures 1 and 37.

6.5.2 Appraisal Path

6.5.2.1 Median-split groups.

Table 115 shows the values used to create the median split groups for the tests along this path, and the resulting group sizes and descriptive statistics for the split groups. Many of the groups had almost the same median value as in Experiment 1, although some were different (compare to Table 65). Two variables in particular were difficult to split because of their skewed distributions. Motivation (Speed of Random Responding) and Environmental Competence both resulted in two groups with very different group sizes, that no adjustment to the split value could alter.

Table 115. Median split groups for appraisal path analyses.

Variable Set	Split	Low Group N	Low Group M	Low Group SD	High Group N	High Group M	High Group SD
Appraisal (Lighting Quality)							
EX2	3.00< or >= 3.00	48	1.99	0.63	59	3.37	0.38
Preference (Attractiveness)*							
EX2	46.33< or => 46.33	53	34.68	10.07	54	58.56	9.43
Mood (Pleasure)							
EX2	4.33< or => 4.33	45	3.16	0.80	53	5.46	0.97
Mood (Arousal)							
EX2	3.33< or => 3.33	46	2.38	0.72	53	4.29	0.78
Mood (Dominance)							
EX2	4.08<= or => 4.08	53	3.64	0.44	53	4.80	0.81
Motivation (Spd. Random Resp.) **							
EX2	<= 729 or = 729	35	337.38	207.08	79	729	0
Environmental Competence							
EX2	6.00<= or => 6.00	33	4.88	0.69	74	6.50	0.56

Note. * Attractiveness was chosen as the only Preference variable for this analysis because it was the strongest of the four components; in the principal components analysis it explained the largest percentage of variance, and its scale scores showed high internal consistency reliability. ** The maximum possible value for this variable was 729.

6.5.2.2 Appraisal → Preference link.

The first link to be tested was between Appraisal and Preference, with its four dependent variables. There was a statistically significant multivariate test and significant univariate test for Attractiveness and Visibility. People who rated the lighting quality as higher also rated the room as more attractive and more visible. These effects had also appeared in Experiment 1, although in that case the between-groups data showed statistically significant relationships between appraisal and all four measures of preference. The effect sizes for the two effects that replicated were different; for Experiment 2, the effect on Attractiveness was smaller than in Experiment 1, but the effect for Visibility was large (compare to Table 66).

Table 116. Test of Appraisal → Preference link.

Link / Dependent Variable	EX2 Wilks' Λ	F	df	$\eta^2_{partial}$
Appraisal → Preference	0.80	6.34***	4, 102	.06
Attractiveness		8.73**	1, 105	.08
Comfort		1.54	1, 105	.01
Visibility		18.01***	1, 105	.15
Spaciousness		0.91	1, 105	.01
	Low M	Low SD	High M	High SD
Attractiveness	42.01	14.93	50.57	14.86
Visibility	56.27	10.64	65.17	10.91

Note. ***p<=.001. **p<=.01. *p<=.05.

6.5.2.3 Preference → Mood link.

We focused, as before, on the Attractiveness component of Preference scores and examined the effect on three mood scales of being high or low on this variable. The results are summarized in Table 117. Only the Pleasure scale showed a statistically significant univariate relationship under the significant multivariate effect. People who rated the room as more attractive rated their own mood as more pleasant. This effect was smaller than the corresponding effect in Experiment 1 (see Table 67). In Experiment 1, all three mood subscales showed statistically significant relationships. The difference could be the result of the smaller sample size in Experiment 2, but there is also evidence that the Experiment 2 participants had an unusual mood response (see section 6.2.3).

Table 117. Test of Preference → Mood link.

Link / Dependent Variable	EX2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Preference→ Mood	0.91	3.21*	3. 93	.04
Pleasure		8.64**	1, 95	.08
Arousal		2.29	1, 95	.02
Dominance		0.41	1, 95	.00
	Low		High	
	M	*SD*	*M*	*SD*
Pleasure	3.96	1.38	4.79	1.39

Note. ***p<=.001. **p<=.01. *p<=.05.

6.5.2.4 Mood → Health and Well-being link.

As in Experiment 1, we tested the median splits on each mood subscale in turn, as antecendents of measures of health (physical and visual comfort) and well-being (satisfaction and task difficulty).

Regarding health (Table 118), for Pleasure the results mirrored Experiment 1 and were approximately the same size (see Table 68 for comparison). Both visual and physical comfort scores were better (lower) for people scoring higher on Pleasure. For Arousal, there was no multivariate significance but we note that there was a statistically significant univariate test for physical comfort (as there had been in Experiment 1). For Dominance there were no statistically significant tests. Thus, we replicated nearly exactly the Pleasure link but did not see the same patterns for Arousal or Dominance as we had in Experiment 1.

Table 118. Tests of Mood → Health link.

Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Pleasure) → Health	0.81	10.88***	2, 95	.13
Visual Comfort		7.72**	1, 96	.07
Physical Comfort		21.98***	1, 96	.19
	Low		High	
	M	*SD*	*M*	*SD*
Visual Comfort	5.62	5.49	3.09	3.42
Physical Comfort	5.27	4.77	1.94	1.82
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Arousal) → Health	.95	2.68	2, 96	
Visual Comfort		1.00	1, 97	
Physical Comfort		5.07*	1, 97	
	Low		High	
	M	*SD*	*M*	*SD*
Physical Comfort	4.44	4.45	2.72	3.10
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Dominance) → Health	.99	0.62	2, 103	

Note. ***p<=.001. **p<=.01. *p<=.05.

Table 119 summarizes the tests of the link between mood and satisfaction. The multivariate link of Pleasure to Satisfaction was associated with a statistically significant test on Performance Satisfaction. People who were in the high-Pleasure group reported higher satisfaction with their performance during the day. This effect replicated the result in Experiment 1, but in Experiment 1 there were also significant tests for the other two satisfaction measures that did not replicate here (compare to Table 69). In neither Experiment 1 nor Experiment 2 were there Arousal or Dominance links to Satisfaction.

Table 119. Tests of Mood → Satisfaction link.

Link / Dependent Variable	EX2 Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Pleasure) → Satisfaction	.92	2.80*	3, 92	.02
Environ. Satis.		1.77	1, 94	.00
Perform. Satis.		7.20**	1, 94	.07
Self-Rated Prod.		0.03	1, 94	.00
	Low		High	
	M	*SD*	*M*	*SD*
Perform. Satis.	2.73	0.59	3.05	0.58
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Mood (Arousal) → Satisfaction	.98	0.58	3, 92	
Mood (Dominance) → Satisfaction	.97	1.01	3, 99	

Note. ***p<=.001. **p<=.01. *p<=.05.

The final set of tests for the health and well-being outcomes concerned ratings of task difficulty, shown in Table 120. Here, as in Experiment 1, there were no relationships between mood and task difficulty (compare to Table 70).

Table 120. Tests of Mood → Task Difficulty link.

Link / Dependent Variable	EX2 Wilks' Λ	*F*	*Df*	$\eta^2_{partial}$
Mood (Pleasure) → Task Difficulty	.91	1.33	7, 88	
Mood (Arousal) → Task Difficulty	.91	1.26	7, 89	
Mood (Dominance) → Task Difficulty	.90	1,59	7, 96	

6.5.2.5 Mood → Motivation link.

The tests of the link between mood and motivation are shown in Table 121. There was a significant link between Arousal and motivation, but in a counterintuitive direction: People in the high-arousal group scored lower on motivation. This was not what we had predicted, but might be explained as a function of the high task demands; more complex or difficult tasks might be expected to suffer under very high arousal (Landy, 1985). This outcome is different to the Experiment 1 result for this test, in which high-dominance individuals were lower in motivation (cf., Table 71), but arousal did not predict motivation.

Table 121. Tests of Mood → Motivation link.

Link / Dependent Variable	EX2			
	F		*df*	$\eta^2_{partial}$
Mood (Pleasure) → Motivation	0.12		1, 96	
Mood (Arousal) → Motivation	7.38**		1, 97	.07
	Low		High	
	M	*SD*	*M*	*SD*
Motivation	687.73	133.60	584.19	226.52
Mood (Dominance) → Motivation	.07		1, 104	

Note. *p<=.05.

6.5.2.6 Competence → Motivation link.

We had also predicted that environmental competence would relate to motivation, and tested this link, with results shown in Table 122. As in Experiment 1, the link did not hold (cf., Table 72). This is probably related to the high scores for environmental competence and the resulting imbalance in group sizes after the median split.

Table 122. Test of Environmental Competence → Motivation link.

Link / Dependent Variable	EX2		
	F	*df*	$\eta^2_{partial}$
Environmental Competence → Motivation	0.82	1, 105	

6.5.2.7 Motivation → Performance link.

We tested the effect of motivation on the various performance scores, as we had done in Experiment 1. There was one statistically significant multivariate test, for vigilance, and it was accompanied by a statistically significant test for hit rate (Table 123). This effect replicated the results of Experiment 1 (see Table 73): People higher in motivation scored higher hit rates (were more vigilant). Other relationships that had been seen in Experiment 1, weak links between motivation and cognitive performance and cognitive judgments, did not replicate.

Table 123. Tests of Motivation → Performance link.

Link Dependent Variable	Wilks' Λ	EX2 *F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Attention)	.72	16.95***	2, 85	.14
Hit Rate		32.47***	1, 86	.27
Envelope Speed		0.58	1, 86	.01
	Low		High	
	M	*SD*	*M*	*SD*
Hit Rate	12.30	21.81	34.28	13.90
Link Dependent Variable	Wilks' Λ	EX2 *F*	*df*	$\eta^2_{partial}$
Motivation → Performance (Typing)	.97	1.14	3, 97	
Motivation → Performance (Cognitive)	.98	0.56	4, 102	
Motivation → Performance (Cognitive Judgments)	.99	0.23	4, 102	
Motivation → Performance (Break Durations)	0.97	1.15	3, 103	

Note. ***p<=.001. **p<=.01. *p<=.05.

6.5.2.8 Summary: appraisal path.

Overall, the links in the appraisal path held up in the Experiment 2 sample. People who appraised the lighting as being of higher quality rated the room as more attractive and visibility higher (in terms of its distinctness, clarity, and brightness; see Section 4.3.2 and Table 20). Those who rated the room as more attractive reported higher pleasure. Those who reported higher pleasure reported greater satisfaction with their performance. Those tests that did not replicate in Experiment 2 were generally the tests with the smallest effect sizes in Experiment 1. Given the larger sample size in Experiment 1, which gave greater statistical power, failing to replicate the small effects with a smaller sample size in Experiment 2 is not surprising.

6.5.3 Vision Path

6.5.3.1 Median split groups.

As for the appraisal path variables, we formed median-split groups for the antecedent variables involved in this path. Their characteristics are found in Table 124.

Table 124. Median split groups for vision path analyses.

Variable Set	Split	Low Group N	*M*	*SD*	High Group N	*M*	*SD*
Visual Comfort							
EX2	3.00< or => 3.00	45	0.98	0.89	62	6.92	5.48
Visual Capability (Composite visual performance at contrast = 16)							
EX2	1.59< or => 1.59	51	1.31	0.24	51	1.87	0.18

6.5.3.2 Appraisal → Visual Comfort link.

We used the Appraisal split (see above) to test whether appraisals of the lighting were associated with visual comfort directly, in addition to the indirect linkage we tested above. Table 125 shows the result of this test. Indeed, those who rated the lighting as being of higher quality had better visual comfort. The effect size is medium-large. This replicates the finding in Experiment 1 (see Table 75).

Table 125. Test of Appraisal → Visual Comfort link.

Link Dependent Variable	EX2 *F*	*df*	$\eta^2_{partial}$	
Appraisal → visual comfort	19.20***	1, 105	.16	
	Low		High	
	M	*SD*	*M*	*SD*
Visual Comfort	6.65	6.55	2.61	2.43

Note. ****p*<=.001. **p*<=.05.

6.5.3.3 Visual Comfort → Visual Capability link.

We tested this link using MANOVA, in which composite visual performance scores at each contrast level on the timed vision test were the six dependent variables (Table 126). As in Experiment 1, there was no relationship between visual comfort and visual capability (compare to Table 76). The likely explanation is that the lighting conditions in both experiments did not create any extreme visual discomfort, and were not bad enough to adversely affect visual capability.

Table 126. Test of Visual Comfort → Visual Capability link.

Link Dependent Variable	Wilks' Λ	EX2 *F*	*df*	$\eta^2_{partial}$
Visual comfort → visual capability	.94	1.01	6,95	

6.5.3.4 Visual Capability → Competence link.

We tested the effects of visual capability on competence using MANOVA, for there were measurements of task competence and environmental competence. Once again, the link between visual capability and competence was not statistically significant (Table 127, compare to Table 77).

Table 127. Test of Visual Capability → Competence link.

Link Dependent Variable	Wilks' Λ	EX2 *F*	*df*	$\eta^2_{partial}$
Visual capability → competence	.99	0.23	2, 98	

6.5.3.5 Visual Capability → Task Performance link.

We tested the effects of visual capability on each type of task performance in turn, as we had done above. In general the results were the same as in Experiment 1 (see Table 128, compare to Table 78). There was a significant relationship between visual capability and vigilance; in Experiment 2 this was associated with a significant effect on envelope speed only, although in Experiment 1 both dependent variables were involved. In both experiments, visual capability was associated with typing performance on all three print sizes. People high on visual capability did better on the typing tasks, with roughly similar effect sizes as in Experiment 1. The tests that did not reach multivariate significance in Experiment 2 had also failed in Experiment 1.

Table 128. Tests of Visual Capability → Performance link.

Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Attention)	.89	5.35**	2, 85	.06
Hit Rate		1.16	1, 86	.01
Envelope Speed		9.71**	1, 86	.10
	Low		High	
	M	*SD*	*M*	*SD*
Envelope Speed	0.28	0.15	0.37	0.12
Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*Df*	$\eta^2_{partial}$
Visual capability → Performance (Typing)	.88	4.56**	3, 97	.09
8 point		8.37**	1, 99	.08
12 point		12.66***	1. 99	.11
16 point		7.22**	1, 99	.07
	Low		High	
	M	*SD*	*M*	*SD*
8 point	2.62	1.21	3.31	1.21
12 point	2.72	1.10	3.54	1.23
16 point	2.87	1.24	3.51	1.18
Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Cognitive)	.96	1.08	4, 97	
Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Cognitive Judgments)	0.98	0.45	4, 97	
Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Visual capability → Performance (Break Durations)	.93	2.50	3, 98	

Note. ***p<=.001. **p<=.01. *p<=.05.

6.5.3.6 Summary: vision path.

In Experiment 2 as in Experiment 1, this path received partial support. Lighting appraisals predicted visual comfort, but visual comfort (at least at the levels achieved here) did not related to visual capability, as assessed with composite visual performance on the timed vision test. Visual capability was not related to competence. Visual capability did relate to task performance on the tasks where such relationships were to be expected, the vigilance and typing tasks. This clarifies questions that remained following the Experiment 1 tests with the repeated measures sample.

6.5.4 Exploratory Tests

In addition to the planned tests of the linked mechanisms map, we considered several additional questions; again we repeated exactly the tests that had been conducted for Experiment 1. Only tests with statistically significant results are reported here.

6.5.4.1 Motivation → Health and Well-being?

In Experiment 1, none of these tests were statistically significant, but in Experiment 2 a link was found between motivation and satisfaction. People who were higher in motivation, oddly, were lower in environmental satisfaction (Table 129). This was a small-medium effect. It is in the opposite direction to the prediction, and is a puzzle. There were no statistically significant links between motivation and health, or motivation and task difficulty ratings.

Table 129. Tests of Motivation→ Satisfaction link.

Link	EX2			
Dependent Variable	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Motivation → Satisfaction	.84	6.35***	3, 100	.04
Environ. Satis.		7.23**	1, 102	.07
Perform. Satis.		3.63	1, 102	.03
Self-Rated Prod.		0.82	1, 102	.01
	Low		High	
	M	*SD*	*M*	*SD*
Environ. Satis.	3.03	0.44	2.65	0.72

6.5.4.2 Satisfaction → Task Performance?

We formed new median split groups on the satisfaction variables to test whether or not satisfaction was related to performance (Table 130). Again, we caution that the design of this experiment means that these measurements break the temporal chain of causality, and emphasize that the tests are exploratory, for the purposes of stimulating new research.

Table 130. Median split groups for satisfaction analyses.

Variable		Low Group			High Group		
Set	Split	N	*M*	*SD*	N	*M*	*SD*
Environmental Satisfaction							
EX2	3.00<= or > 3.00	50	2.25	0.55	57	3.23	0.36
Performance Satisfaction							
EX2	3.00< or => 3.00	52	2.38	0.48	53	3.33	0.37
Self-Rated Productivity							
EX2	0.00<= or > 0.00	61	-0.66	0.91	45	1.98	1.03

Only the statistically significant tests are reported in Table 131. In general the results differ from Experiment 1. Environmental satisfaction showed a statistically significant relationship to cognitive judgments, with ratings of the summaries both for "Grammar" and "Well-written" being higher for people high in environmental satisfaction. The corresponding effect in Experiment 1 had a smaller size and only operated on "Well-written" (compare to Table 80). Environmental Satisfaction was related to cognitive performance in Experiment 2, although it had not been in Experiment 1 (see Table 80). In Experiment 2, people high in environmental satisfaction took longer to categorize summaries and longer to create summaries from whole articles. There were no significant tests involving the median-split groups on self-rated productivity, although there had been in Experiment 1.

Table 131. Tests of Satisfaction → Task Performance link.

Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Environmental Satisfaction → Performance (Cognitive Judgments)	.89	3.02	4, 102	.02
Interest		0.05	1, 105	.00
Factual		0.43	1, 105	.00
Grammar		6.61**	1. 105	.06
Well-written		10.89***	1, 105	.09
	Low		High	
	M	*SD*	*M*	*SD*
Grammar	2.40	0.54	2.65	0.45
Well-written	2.30	0.41	2.59	0.49
Link / Dependent Variable	EX2			
	Wilks' Λ	*F*	*df*	$\eta^2_{partial}$
Environmental Satisfaction → Performance (Cognitive)	.88	3,37**	4, 102	.03
Cat. Time		4.24*	1, 105	.04
Cat. Acc.		0.65	1, 105	.01
Judge Time		0.45	1, 105	.00
Summ. Time		6.52**	1, 105	.06
	Low		High	
	M	*SD*	*M*	*SD*
Cat. Time	14.06	5.47	17.13	9.21
Summ. Time	32.81	17.81	41.49	17.30

Note. ****p*<=.001. ***p*<=.01. **p*<=.05.

6.5.4.3 Performance → Satisfaction?

Conversely, we examined whether people in high-performance groups were more satisfied than others. The median split groups for these analyses are reported in Table 132. We selected the same variables for these analyses as we had used in Experiment 1.

Table 132. Median split groups for performance analyses.

Variable / Set	Split	Low Group N	Low Group *M*	Low Group *SD*	High Group N	High Group *M*	High Group *SD*
Typing (8 pt)							
EX2	2.97<= or > 2.97	50	1.93	0.63	51	3.99	0.78
Summary Breaks							
EX2	1.70<= or > 1.70	53	1.18	0.52	54	1.85	0.11
Interest in Article							
EX2	50.00<= or > 50.00	56	39.81	10.19	51	62.27	9.61

There was only one statistically significant relationship in these tests. People who were better typists rated the typing task as easier (Table 133). This effect replicated the one seen in Experiment 1 (cf., Table 82).

Table 133. Tests of Task Performance → Satisfaction link.

Link Dependent Variable	EX2 Wilks' Λ	F	df	$\eta^2_{partial}$
Performance (Typing) → Task Difficulty	.86	2.21*	7, 91	.02
Typing		5.62*	1, 97	.05
Categorizing		0.90	1, 97	.01
Evaluating summ.		1.62	1, 97	.02
Extracting summ.		3.63	1, 97	.04
Questionnaires		0.29	1, 97	.0
Vision tests		0.79	1, 97	.01
All tasks		0.23	1, 97	.00
	Low		High	
	M	*SD*	*M*	*SD*
Typing	1.49	1.23	0.86	1.00

6.5.4.4 Mood affects Social Behavior?

We approached this question with two analyses, the first examining pleasure as a predictor and the second involving environmental satisfaction. The dependent variables were: liking for other group members, willingness to volunteer, and the two types of responses to the conflict resolution task (ratings and rankings).

Pleasure did not relate to any of the social behavior outcomes, nor did environmental satisfaction. Small effects that had been observed in the between-groups sample in Experiment 1 did not replicate. This might be because of the lower statistical power with the smaller sample in Experiment 2.

6.5.5 Summary: Linked Mechanisms Tests

In general the results of the Experiment 2 linked mechanisms tests are encouraging. Despite a much smaller sample, and different experimental conditions, most of the links in the linked mechanisms map replicated the results for the between-groups sample in Experiment 1. This gives added confidence that the appraisal path and the vision path are meaningful explanations for the effects of lighting on people in offices. The ambiguous results for the exploratory tests are unsurprising, given that the experiment was not designed for these tests and the sample was much smaller. These are matters for future research to consider, and are discussed in the General Discussion.

7. General Discussion

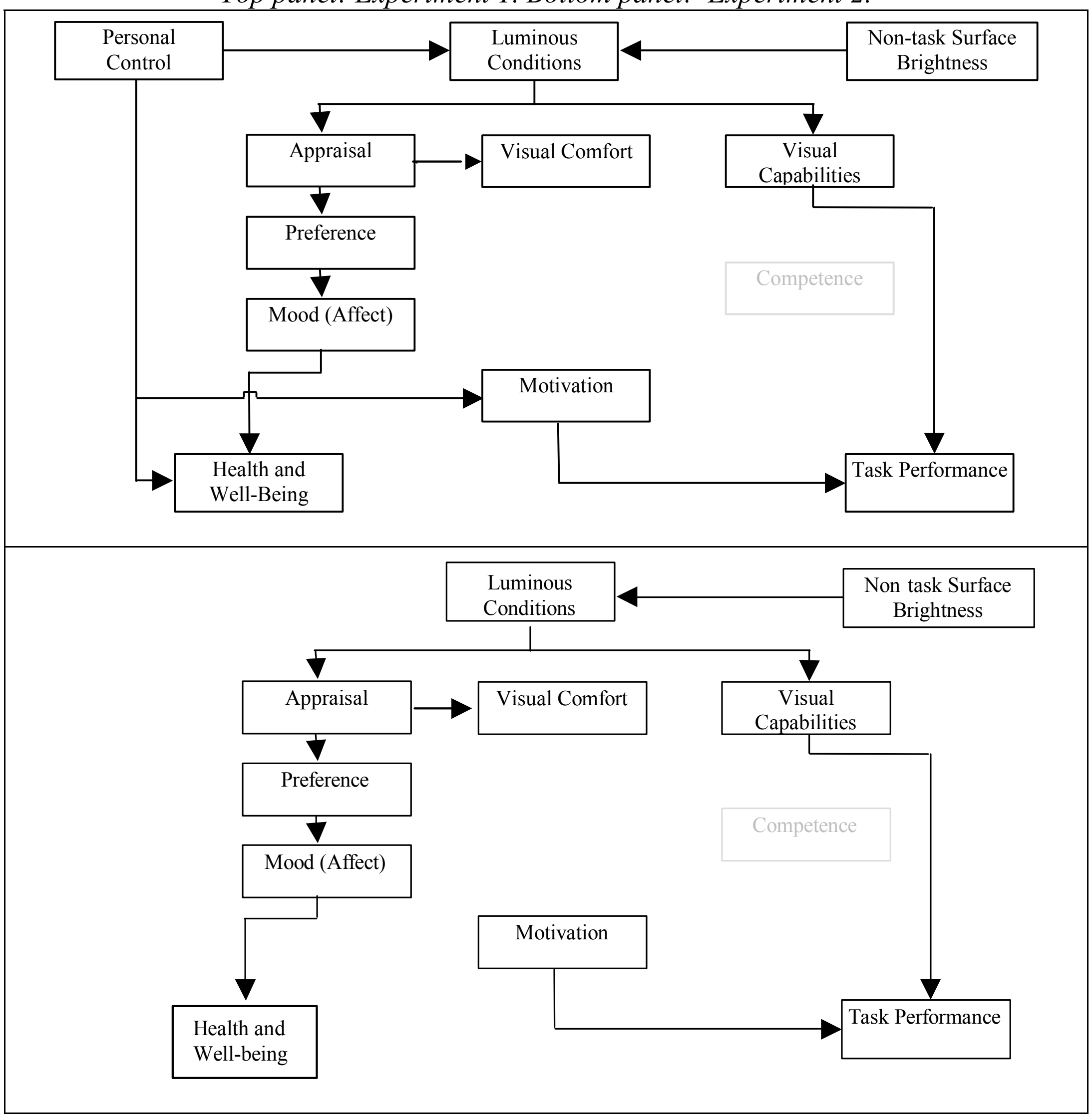
7.1 The Modified Linked Mechanisms Map

Having discussed the plethora of results obtained in the two experiments, it is now desirable to undertake a synthesis of the data so as to reveal the "big picture." A convenient place to start is with the linked mechanisms map shown in Figure 1 (and 37). This map was based on the mechanisms by which changes in surface brightness and having personal control of lighting were expected to improve task performance and enhance feelings of health and well-being. The procedures used and the dependent variables measured in Experiment 1 were selected so as to test all the links in the map.

The results of Experiment 1 and the statistical tests of each link in the linked mechanisms map (described in Section 4.7 above) allow the map to be modified to reflect what has been found. This is done by eliminating all the links where no statistically significant relationship was found. Figure 51 shows the modified linked mechanisms map resulting from Experiment 1 (top panel). Experiment 2 was designed to further test the effects of surface brightness and lighting conditions on task performance and feelings of health and well-being. Figure 51 (bottom panel) shows the modified linked mechanisms map derived from the results of Experiment 2 (based on the results reported in section 6.5).

The similarity between the linked mechanisms maps obtained from two experiments using different lighting systems and different individuals is notable and supports the validity of the relationships contained in the maps. The obvious difference between the linked mechanisms maps from the two experiments is the presence of the effects of having individual control in Experiment 1 and their absence in Experiment 2. This is to be expected because two of the lighting systems used in Experiment 1 provided the participants with individual control of their lighting while neither lighting condition used in Experiment 2 provided individual control. Competence appears in gray in both maps, with no links to other concepts, because although it appeared in the original model no statistically significant links to it were observed in either experiment.

Figure 51. The modified linked mechanisms maps.
Top panel: Experiment 1. Bottom panel: Experiment 2.



There are five features of the modified linked mechanisms maps that deserve discussion. They are:

1. The path from lighting conditions through visual capability to task performance is present in both experiments. This means that lighting conditions and task conditions that improve the visibility of the task lead to better task performance. This is hardly news but its presence is essential to the credibility of the modified linked mechanisms map. It is also useful in that it extends a path largely derived from laboratory studies using abstract stimuli (Rea and Ouellette, 1991; Weston, 1945) to a much more realistic setting and tasks.

2. The path from luminous conditions through appraisal, preference, and mood to health and well-being is present in both experiments. This means that people who are more satisfied with their lighting consider the space to be more attractive, are happier, and are more comfortable and more satisfied with their environment and their work. Various pieces of this path have been demonstrated before but this is the first time the complete path from lighting conditions to feelings of health and well-being has been demonstrated.

3. The linked mechanisms map for Experiment 1 shows that having personal control of the lighting in the form of dimming has a direct link to feelings of health and well-being, in that people with dimming control reported higher ratings of lighting quality, overall environmental satisfaction, and self-rated productivity. Presumably this direct link occurs because having personal control in the form of dimming allows the individual the potential to adjust the lighting to match his/her preferences. That these preferences can vary widely between individuals for the same tasks is evident from Figure 36, which shows the illuminances chosen by the people with dimming control. That these preferences are not closely linked to specific tasks is shown by the fact that the individuals with dimming control tended to adjust the illuminance once at the beginning of the day and not every time they change tasks.

4. The linked mechanisms map for Experiment 1 also shows that there is a link between having personal control and task performance operating through motivation, in that people with dimming control showed more sustained motivation and improved performance on a measure of attention. Again , this presumably occurs because having personal control in the form of dimming gives the individual the potential to adjust the lighting to match his/her preferences, which acts as a motivator.

5. What is absent from both modified linked mechanisms maps is also important. The most notable absence is any cross-links between the vision path that connects luminous conditions through visual capability to task performance, and the appraisal path that connects luminous conditions to feelings of health and well-being through appraisal, preference, and mood. Such cross-links are widely believed to exist, operating either through visual discomfort, which affects both the vision path and the appraisal path, or between mood and task performance via motivation (Boyce and Rea, 2001; Cropanzano and Wright, 2001).

This study offers no support for these links. The failure to find a link through visual discomfort may be easily explained by the fact that the levels of discomfort experienced by the participants in both experiments were slight, certainly not enough to cause a change in visual capability. If lighting conditions that caused severe visual discomfort had been used, it is likely that such a link would have been found. However, such lighting installations are unrealistic because, in practice, such lighting installations would lead to occupant complaints and therefore would not be common lighting practice.

The failure to find a link between mood and task performance is more difficult to explain. There is no doubt that mood is associated with behavior relevant to work (Isen and Baron,1991) and that aspects of the physical environment can affect mood (e.g., Baron, 1990b; McCloughan et al., 1999) but whether the generally comfortable lighting conditions examined in this study would be enough to produce changes in mood sufficient to change motivation over one day's work is open to question. Again, if much more uncomfortable conditions had been used, a link between mood and task performance through motivation might have been found. An alternative explanation is that the measure of motivation used was insufficiently sensitive to detect small changes. A more sensitive measure of motivation might have revealed such a link.

Although the repeatability of the linked mechanisms maps across two experiments suggests that the basic structure is sound, it should not be thought that the maps are incapable of refinement. Some exploratory tests of relationships between satisfaction with the environment and task performance show statistically significant links, but the relationships did not appear stable from Experiment 1 to Experiment 2, and some were in the opposite direction to what one would have predicted. It is interesting that the people who are more satisfied with the environment are also those who appear to be more thoughtful in making cognitive judgments (taking longer), and rate the reading material more highly. These relationships deserve further investigation; these two experiments were not designed to resolve questions about the cross-linkages between performance and satisfaction.

7.2 The Importance of the Lighting Conditions for Office Work

This study was undertaken to determine to what extent different levels of lighting quality, as judged by the Light Right consortium, affect task performance of office workers and their health and well-being. That lighting conditions can have such effects is shown by the existence of the vision path and the appraisal path in both revised linked mechanisms maps. The question that will be considered here is whether the lighting conditions compared in these experiments influence lighting appraisal or vision sufficiently to change the final outcomes at the end of the linked mechanisms for people working in real offices.

In Experiment 1, three "Best Practice" office lighting design conditions were compared with a base-case lighting condition representative of modern office lighting practice, a regular array of parabolic-louvered luminaires. All of the "Best Practice conditions used direct / indirect luminaires, and two of them provided a degree of individual control. In Experiment 2 two fixed lighting installations were compared, one being another common form of office lighting (Base Case 2), i.e., recessed prismatic luminaires in a regular array; the other being the same Best Practice installation used in Experiment 1 but set to a lower illuminance (Best Practice 2). In both experiments, the comparison was done in terms of office workers' task performance and feelings of health and well-being, over a complete working day. The office workers who experienced the lighting conditions were temporary office workers, with no expert knowledge of lighting. Different participants provided data in each experiment. The tasks examined were representative of those done in a modern office, but with an emphasis on knowledge-based tasks.

The statistical analyses of the data collected revealed both good news and bad. The good news is that the lighting designs are discriminated in terms of comfort and in comparison to other office lighting installations with which the participants were familiar. For example, in Experiment 1, the Base Case parabolic lighting was considered comfortable by 71 percent of the participants who experienced it but the Dimming Control lighting was considered comfortable by 91 percent of those who experienced it. In Experiment 2, the Base Case 2 prismatic lighting was considered comfortable by 69 percent of those who experienced it. Further, the Best Practice installations were considered comfortable by 85 percent of subjects in Experiment 1 and 81 percent in Experiment 2, the slight decline probably being associated with the lower illuminance used in Experiment 2.

The bad news is that there were no statistically significant simple main effects of lighting conditions on the performance of any of the tasks in either experiment. This implies that there is no consistent and clear advantage to one or other of the lighting conditions as regards task performance. There were some statistically significant effects of lighting conditions but these were all interactions involving some combination of time, cubicle reflectance, and task visual difficulty, as well as lighting conditions. The interaction effects suggested that there might be small benefits to a direct / indirect lighting system, depending on the illuminance level, the partition reflectance, and the task. At the illuminance levels said to be typical of current practice (Best Practice 2), the best outcomes for the Best Practice 2 lighting mostly (but not always) occurred for people in cubicles with light-colored partitions. At the higher illuminance levels of the Best Practice in Experiment 1, whether the interaction showed better performance for Base Case or Best Practice depended on the partition reflectance and the task.

Interactions in Experiment 1 also suggested some advantage of having personal control of the lighting for task performance over time. More importantly, these interaction effects suggest that the availability of control can offset declining performance over the working day; that is, it might be a means of inoculation against fatigue.

The lack of a simple, clear, and consistent effect of lighting conditions on task performance cannot be explained by a lack of sensitivity in the design of the tasks. The expected effects of changing visibility on task performance were found in both experiments, e.g., decreasing typing speed with smaller print sizes in the typing task. Similarly, the expected changes in task performance over time due to practice and fatigue were found, e.g., the change over time of the cognitive tasks.

What all this implies is that although the participants did perceive and appreciate the differences between the lighting conditions, this perception made very little difference to their performance of the tasks. One plausible explanation for this pattern of results is that the performance of the tasks is essentially governed by the stimuli the tasks provide to the visual system and the luminous conditions in the immediate surrounding area, while the appraisal of the lighting is influenced by a wider range of considerations, among them certainly being the ability to see the details necessary to do the task, but also the appearance of the office as a whole, and the appearance of the fixtures.

7.3 Patterns of Use of Lighting Control

Another feature of the data collected in Experiment 1 is the possibility of looking at the ways in which the desk light in the Best Practice with Switching Control condition and the dimming system in the Dimming Control condition were used. Before considering the results it is worth pointing out that the switching and dimming systems allow the participants to achieve rather different ranges of illuminances. The switched desk lamp could only add light to the illuminance provided by the ambient Best Practice lighting, giving a minimum illuminance on the desk of 400 lx and a maximum of 1131 lx. By adjusting the light output of two of the lamps in the luminaire centered over each cubicle, the dimming system allowed the ambient illuminance to be reduced as well as increased thereby giving a minimum illuminance on the desk of 251 lx and a maximum of 1176 lx. Although these two systems allowed somewhat different ranges of illuminance they were both used in a similar way.

For both systems, a clear majority of participants switched or dimmed the lighting once at the beginning of the day and then left the control alone. This is a similar pattern of behavior to that found in other studies of lighting control use (Crisp, 1978; Maniccia et al. 1999; Moore et al., 2002). Another interesting observation for lighting practice is the illuminance chosen. For the dimming control, Figure 36 shows the distribution of horizontal illuminances on the desk selected by those who had the dimming control. Two features of this figure are notable. The first is the wide range of illuminances selected, ranging from the 200 lx to 1100 lx. Such wide ranges have been found previously in laboratory experiments in which people were able to adjust the illuminance to their preference (Boyce et al., 2000; Veitch & Newsham, 2000b). The second is the shape of the distribution, which is strongly skewed with a peak close to the lowest illuminance possible. On average these participants preferred illuminances considerably below those currently recommended and used for much office lighting.

Figure 36 also shows the distribution of horizontal illuminances selected by the participants with switching control. Again there is a wide range of illuminances selected but the most frequently selected horizontal illuminances were in the range 660 lx to 1000 lx. These illuminances are higher than those currently recommended and used for much office lighting. Provided that the illuminance is sufficient to allow easy task performance, then the illuminance preference is partially determined by the characteristics of the device used to adjust the illuminance.

7.4 Implications for Lighting Practice

The findings of this study carry a number of implications for lighting practice. The first is quite simply that current office lighting practice is acceptable to many people. The Base Case 1 installation, consisting of a regular array of parabolic luminaires, and the Base Case 2 installation, consisting of the same regular array of recessed prismatic luminaires, were assessed as comfortable by 70 percent and 69 percent of the people who experienced them, respectively. Further, the performance of all the visual and cognitive tasks showed no difference between Base Case lighting and the fixed Best Practice lighting (fixed direct / indirect lighting) in either experiment. However, ventilation engineers have set 80% comfortable as the target for choosing thermal conditions (ASHRAE, 1992); should we accept 70% as the target for comfort with lighting?

If it is desirable to increase the percentage of people who assess the lighting as comfortable, these experiments provide evidence that it should be possible to do so, by going to the direct / indirect lighting without personal control, or even better, with individual dimming control. In Experiment 1, 85 percent of those who experienced the Best Practice lighting considered it comfortable and 53 percent said it was better than the lighting in similar workplaces in other buildings. Similarly, 91 percent who experienced the Dimming Control condition considered it comfortable and 50 percent considered it better than the lighting in similar workplaces in other buildings. The advantage for direct / indirect lighting was not so marked in Experiment 2, 81 percent of those who experienced it considered it comfortable but only 27 percent considered it better than the lighting of similar workplaces.

It is important to appreciate that there were features of the Best Practice 1 and 2 and Dimming Control lighting that might limit the implications of these results. In both experiments, the Best Practice lighting consisted of regular lines of suspended direct / indirect luminaires but in Experiment 1 the mean illuminance on the desk was similar to that provided by the Base Case 1 recessed parabolic lighting. whereas, in Experiment 2 the mean illuminance was about 27 % lower. It is common practice when using direct / indirect systems to provide about a 20 percent lower illuminance than when using direct lighting. The slight reduction in the percentage of subjects considering the Best Practice 2 lighting comfortable in Experiment 2 is probably associated with this reduction in illuminance, causing discomfort for people whose preference is for higher illuminances.

As for the Dimming Control lighting used in Experiment 1, the obvious feature of this direct / indirect lighting that differentiates it from other forms of office lighting is indeed the facility for the occupant of a cubicle to adjust the illuminance on the whole desk over a large range. The important words here are "cubicle", "whole desk", and "large range." The change in illuminance is provided by altering the light output of the downward component of a direct / indirect luminaire suspended over the center of the cubicle. This means that adjusting the light output of the luminaire alters the illuminance at all points in the cubicle, with very little effect on the illuminances in adjacent cubicles. Similar perceptions of comfort and better quality are unlikely to have occurred had the changes in illuminance been confined to a small area of the desk or had those changes affected adjacent cubicles in a disturbing way. As for the "large range", the point is that the control can be used to both increase and decrease the illuminance on the desk from the illuminance that is usually provided in a fixed office lighting installation. One doubts that the same high level of comfort and perception of better quality would have been found had the control been limited to reducing the illuminance to less than the illuminance provided for a fixed office lighting installation, particularly if there had also been no attempt to keep the perimeter walls bright using separate wall-washing (cf., Newsham, Marchand, & Veitch, in press).

Another way to consider the implications of these results to lighting practice is to consider the extent to which they agree with or require modification of the published guidelines for office lighting. The most widely used guidelines for office lighting in North America are given in the IESNA/ANSI *American National Standard for Office Lighting* RP-1-1993 (IESNA, 1993). This publication states that both direct lighting, as used in the Base Case lighting, or direct / indirect lighting, as used in the Best Practice lighting, can provide good office lighting, but suggests that the latter is more likely to produce excellent lighting than the former. This supports the finding that the Base Case lighting is perceived as reasonable and that the Best Practice lighting is perceived as better. There is no discussion of the role of individual dimming in the Recommended Practice, probably because the provision of this facility is still rare. One aspect of office lighting that is given little attention in the Recommended Practice is illuminance uniformity. This will need to change if installations similar to that used in the Dimming Control lighting become more widely used. Such installations will inevitably produce very non-uniform patterns of illuminance across the office, although within the cubicle the distribution is relatively uniform. This is to be expected, but some guidance setting limits on the range of illuminances that can be achieved would be useful, as would some limits on the interaction between adjacent cubicles.

Another aspect of these results with rather more wide ranging implications than the choice of lighting system is the finding that giving people personal dimming control leads to the selection by the majority of lower illuminances than those currently recommended for offices. In fact, the illuminances selected represent about a 20 to 40 percent energy saving. This might be taken to imply that the recommended illuminances for offices should be reduced to a much lower value. A word of warning is necessary about such an approach. The problem with this approach is evident in the distribution of illuminances selected (Figure 36). Although the peak of the distribution is in the range 200 to 400 lx, there is a long tail extending up to much higher illuminances than those currently recommended. People who like such high illuminances are unlikely to be appreciative of illuminances in the 200 to 400 lx range. A better approach might be to provide the Dimming Control through some form of task lighting, while lowering the level of ambient lighting. However, as stated earlier, task lighting that does not illuminate the whole desk might not be acceptable.

7.5 The Way Ahead

Finally, it is necessary to consider if, where, and how the study of the effects of lighting quality on the performance of office work should continue. Previous experiments have attempted to uncover effects of lighting design on performance and satisfaction in laboratory settings with varying degrees of realism (e.g., Boyce et al., 2000; Veitch & Newsham, 1998a). This experiment has extended this work to a very realistic, field simulation setting. Having done so, and found few simple effects of lighting design, the first thing to say is that there is little point in trying to repeat this experiment unless a deliberate effort is made to improve the ability to detect small effect sizes by dramatically increasing the number of participants. Of course, whether spending more and more resources to detect smaller and smaller effects is worthwhile is a question that needs to be carefully considered by all involved, particularly when there are alternative approaches. Specifically, there are two diverging paths for study that can be recommended.

One path would be to conduct targeted, rigorous laboratory experiments to examined elements in the linked mechanism map. Laboratory studies should be used for this purpose because these can be undertaken with much more control. One important theoretical question that could be answered by such laboratory studies would be to test whether the missing links between the vision path and the appraisal path would appear if clearly uncomfortable conditions were to be included for comparison. There would also be the possibility of using populations of subjects chosen for their sensitivity to lighting conditions and for developing a better measure of motivation. These actions would tend to increase the sensitivity of the experiment and hence would provide a rigorous test of the revised linked mechanisms map.

To continue the study of the effects of lighting quality on task performance for offices, it is recommended that effort be moved into the field. This would have three advantages. The first is that the results would have a high level of realism. Lighting is only part of the work context, and any simulation experiment can only simulate part of that context. Effects involving visibility occur regardless of context, but mood and motivation effects are context-dependent. Studying these latter effects requires real people in real organizations.

The second advantage of field investigation is that it would allow the accumulation of results over an extended period of time. This is of value because it may be that lighting conditions which can be ignored for one day become important when one is exposed to them for many days and months. Conversely, it may be that lighting conditions that are seen as better on first acquaintance become the norm over many months and so reduce in impact. Given the results of this experiment, it would be interesting to determine whether working under well-designed lighting (particularly with some degree of individual control) could offset the ill effects of high job demands. It would be next to impossible to simulate such conditions in a laboratory setting, but they occur naturally in organizations.

The third advantage of field research into lighting effects on people at work is that it allows for the possibility of measuring the effects of lighting conditions on aspects of performance and behavior at an organizational level rather than an individual level. This study has focused on the performance of individuals and has made no attempt to measure such variables as absenteeism, recruitment, and staff retention that are important at an organizational level. However, the results of the exploratory tests of the relationships between environmental satisfaction, motivation, and performance suggest that measurements at an organizational level would be valuable.

Although we know of no targeted research into lighting's effect on organizational productivity, there are other investigations that suggest that physical environmental conditions influence important organizational outcomes. Several investigations have found that satisfaction with the physical environment predicts job satisfaction and/or organizational commitment (Carlopio, 1996; Veitch, Farley, & Newsham, 2002; Wells, 2000). It is logical to expect that any design choice, including lighting, that can improve environmental satisfaction should also lead to beneficial outcomes on these variables. Carefully-designed field investigations, over long periods and with large samples, could provide the evidence that many have long sought, that lighting quality contributes to organizational productivity.

8. References

American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE). (1992). *Thermal environmental conditions for human occupancy* (ANSI/ASHRAE Standard 55-1992). Atlanta, GA: ASHRAE.

Amick, B., Robertson, M., Derango, K., Palacios, N., Allie, P., Rooney, T., & Bazzani, L. (2002). The health consequences of an office ergonomics training coupled with an ergonomically designed chair: Preliminary results. In H. Luczak, A. E. Cakir and G. Cakir (Eds.), *WWDU 2002 - World Wide Work. Proceedings of the 6th International Scientific Conference on Work with Display Units* (pp. 371-373). Berlin: ERGONOMIC Institut fur Arbeits- und Sozialforschung, Forschungsgesellschaft mbH.

Bandura, A. (1982). Self-efficacy mechanisms in human agency. *American Psychologist, 37,*122-147.

Barnes, R. D. (1981). Perceived freedom and control in the built environment. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 409-422). Hillsdale, NJ: Erlbaum.

Baron, R. A. (1990a). Conflict in organizations. In K. R. Murphy & F. E. Saal (Eds.), *Psychology in Organizations: Integrating Science and Practice* (pp. 197-216). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baron, R. A. (1990b). Environmentally induced positive affect: Its impact on self-efficacy, task performance, negotiation, and conflict. *Journal of Applied Social Psychology*, *20*(5), 368-384.

Baron, R. A., Rea, M. S., & Daniels, S. G. (1992). Effects of indoor lighting (illuminance and spectral distribution) on the performance of cognitive tasks and interpersonal behaviors: The potential mediating role of positive affect. *Motivation and Emotion*, *16*, 1-33.

Becker, F. D. (1986). Quality of work environment (QWE): Effects on office workers. *Prevention in Human Services*, 4, 35-57.

Borenstein, M., Rothstein, H., & Cohen, J. (1997). *Power and precision* (Release 1). Teaneck, NJ: Biostat.

Boyce, P .R.. (2003). *Human factors in lighting* (2nd ed.). London: Taylor and Francis.

Boyce, P. R., & Eklund, N. H. (1996, August). *An evaluation of office lighting options*. Paper presented at the 1996 Annual Conference of the Illuminating Engineering Society of North America, Cleveland, OH, August 5-7, 1996.

Boyce, P. R., Eklund, N. H., & Simpson, S. N. (2000). Individual lighting control: Task performance, mood, and illuminance. *Journal of the Illuminating Engineering Society*, *29*(1), 131-142.

Boyce, P. R., and Rea, M.S., (2001). *Lighting and Human Performance II: Beyond Visibility Models toward a Unified Human Factors Approach to Performance*. Palo Alto, CA: EPRI; Arlington, VA: National Electrical Manufacturers Association, VA; and Washington, D.C.: U.S. Environmental Protection Agency Office of Air and Radiation.

Carlopio, J. R. (1996). Construct validity of a physical work environment satisfaction questionnaire. *Journal of Occupational Health Psychology, 1*(3), 330-344.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Committee on Recommendations for Quality and Quantity of Illumination. (1966). Outline of a standard procedure for computing visual comfort ratings for interior lighting - Report No. 2. *Illuminating Engineering, 61,* 643-666.

Commission Internationale de l'Eclairage [CIE]. (1995). *Discomfort glare in interior lighting* (Publication No. 117-1995). Vienna, Austria: CIE.

Crisp, V. H. C. (1978). The light switch in buildings. *Lighting Research and Technology*, *10*, 69-82.

Cropanzano, R. and Wright, T. A. (2001). When a "happy" worker is really a "productive" worker: A review and further refinement of the happy-productive worker thesis. *Consulting Psychology Journal, 53*(3), 154-168.

Eklund N. H. and Boyce P. R. (1996). The development of a reliable, valid and simple office lighting survey. *Journal of the Illuminating Engineering Society*, *25*(2), 25-40.

Eklund, N. H., Boyce, P. R., & Simpson, S. N. (2000). Lighting and sustained performance. *Journal of the Illuminating Engineering Society*, *29*(1), 116-130.

Eklund, N. H., Boyce, P. R., & Simpson, S. N. (2001). Lighting and sustained performance: Modeling data-entry task performance. *Journal of the Illuminating Engineering Society*, *30*(2), 126-140.

Feather, N. T. (1962). The study of persistence. *Psychological Bulletin, 59*, 94-115.

Flynn, J. E., Hendrick, C., Spencer, T., & Martyniuk, O. (1979). A guide to methodology procedures for measuring subjective impressions in lighting. *Journal of the Illuminating Engineering Society*, *8*, 95-110.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco, CA: W. H. Freeman.

Gifford, R., Hine, D. W., & Veitch, J. A. (1997). Meta-analysis for environment-behavior research, illuminated with a study of lighting level effects on office task performance. In G. T. Moore & R. W. Marans (Eds.), *Advances in Environment, Behavior, and Design* (Vol. 4 , pp. 223-253). New York: Plenum.

Hedge, A., Sims, W. R., & Becker, F. D. (1995). Effects of lensed-indirect and parabolic lighting on the satisfaction, visual health, and productivity of office workers. *Ergonomics, 38*, 260-280.Jones, S. J., & Veitch, J. A. (2000, August). *The effects of workplace control on employee satisfaction and performance*. Paper presented at the 108th Annual Convention of the American Psychological Association, Washington, D.C., August 4 - 8, 2000.

Illuminating Engineering Society of North America [IESNA]. (1993). *American National Standard Practice for Office Lighting* (ANSI/IESNA RP-1-1993). New York. IESNA.

Isen, A. M., & Baron, R. A. (1991). Positive affect as a factor in organizational behavior. In B. M. Staw and L. L. Cummings (Eds), *Research in Organizational Behavior* (Vol. 13, pp. 1-53). Greenwich, CT: JAI Press.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth, TX: Harcourt College.

Kilmann, R. H., & Thomas, K. W. (1977). Developing a forced-choice measure of conflict-handling behavior: The "MODE" instrument. *Educational and Psychological Measurement*, *37*(2), 309-325.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guildford.

Landy, F. J. (1985). *Psychology of work behavior.* Homewood, IL: Dorsey.

Loe, D. L., Mansfield, K. P., & Rowlands, E. (1994). Appearance of lit environment and its relevance in lighting design: Experimental study. *Lighting Research and Technology*, *26*, 119-133.

Maniccia, D., Rutledge, B., Rea, M. S., & Morrow, W. (1999). Occupant use of manual lighting controls in private offices. *Journal of the Illuminating Engineering Society, 28*(2), 42-56.

McCloughan, C. L. B., Aspinall, P. A., & Webb, R. S. (1999). The impact of lighting on mood. *Lighting Research and Technology, 31,* 81-88.

Moore, T., Carter, D. J., & Slater, A. I. (2002). A field study of occupant controlled lighting in offices. *Lighting Research and Technology*, *34*(3), 191-206.

Newsham, G. R., & Veitch, J. A. (2001). Lighting quality recommendations for VDT offices: A new method of derivation. *Lighting Research and Technology, 33*, 97-116.

Newsham, G. R., Veitch, J. A., & Scovil, C. Y. (1995a). *Conveyor belt task: Software to test user reaction time using targets on a simulated conveyor belt* (IRC-IR-702). Ottawa, ON: National Research Council Canada, Institute for Research in Construction.

Newsham, G. R., Veitch, J. A., & Scovil, C. Y. (1995b). *Typing task: Software to measure the speed and accuracy with which presented text is typed* (IRC-IR-700). Ottawa, ON: National Research Council Canada, Institute for Research in Construction.

Newsham, G. R., Marchand, R. G., & Veitch, J. A. (2003). Preferred surface luminances in offices, by evolution. *Journal of the Illuminating Engineering Society*, in press.

Rea, M. S., & Ouellette, M. J. (1991). Relative visual performance: A basis for application. *Lighting Research and Technology, 23,* 135-144.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Revised ed.). Newbury Park, CA: Sage.

Russell, J. A., & Mehrabian, A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.

Sundstrom, E., Town, J. P., Rice, R. W., Osborn, D. P., & Brill, M. (1994). Office noise, satisfaction, and performance. *Environment and Behavior*, *26*, 195-222.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Thomas, K. (1976). Conflict and conflict management. In M. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.

Veitch, J. A. (Ed). (1998). *Proceedings of the First CIE Symposium on Lighting Quality* (Publication x015-1998). Vienna: CIE.

Veitch, J. A., Farley, K. M. J., & Newsham, G. R. (2002). *Environmental satisfaction in open-plan environments: 1. Scale validation and method* (IRC Internal Report IR-844). Ottawa, ON: National Research Council Canada, Institute for Research in Construction. Veitch, J. A., Gifford, R., & Hine, D. W. (1991). Demand characteristics and full spectrum lighting effects on performance and mood. *Journal of Environmental Psychology*, *11*, 87-95.

Veitch, J. A., & Newsham, G. R. (1996, August). *Determinants of lighting quality II: Research and recommendations*. Paper presented at the 104th Annual Convention of the American Psychological Association, Toronto, Ontario, Canada. (ERIC Document Reproduction Service No. ED408543).

Veitch, J. A., & Newsham, G. R. (1998a). Lighting quality and energy-efficiency effects on task performance, mood, health, satisfaction and comfort. *Journal of the Illuminating Engineering Society, 27*(1), 107-129.

Veitch, J. A., & Newsham, G. R. (1998b). Determinants of lighting quality I: State of the science. *Journal of the Illuminating Engineering Society*, *27*(1), 92-106.

Veitch, J. A., & Newsham, G. R. (2000a). Exercised control, lighting choices, and energy use: An office simulation experiment. *Journal of Environmental Psychology, 20,* 219-237.

Veitch, J. A., & Newsham, G. R. (2000b). Preferred luminous conditions in open-plan offices: Research and practice recommendations. *Lighting Research and Technology*, *32*, 199-212.

Wagner, F. R., & Morse, J. J. (1975). A measure of individual sense of competence. *Psychological Reports, 36*, 451-459.

Wells, M. M. (2000). Office clutter or meaningful personal displays The role of office personalization in employee and organizational well-being. *Journal of Environmental Psychology, 20*(3), 239-255.

Weston, H.C. (1945). *The relation between illumination and visual efficiency: the effect of brightness contrast* (Industrial Health Research Board, Report No. 87). London: His Majesty's Stationary Office.

Wibom, R. I., & Carlsson, L. W. (1987). Work at video display terminals among office workers. In B. Knave & P. G. Wideback (Eds.), *Work with Video Display Units 86* (pp. 357-367). Amsterdam, the Netherlands: Elsevier Science.

Wyon, D. P. (1996). Indoor environmental effects on productivity. In K. Y. Teichman (Ed.), *IAQ '96: Paths to Better Building Environments* (pp. 5-15). Atlanta, GA: American Society of Heating, Refrigerating, and Air Conditioning Engineers.

Acknowledgements

This investigation was supported by the Light Right Consortium, managed by Pacific Northwest National Laboratory under RFP # 404141, through contracts to Rensselaer Polytechnic Institute's Lighting Research Center (RPI # A11106) and the National Research Council of Canada's Institute for Research in Construction (NRC # B3214). This phase of the Light Right Consortium was supported by the Alliance to Save Energy, the Illuminating Engineering Society of North America, the International Association of Lighting Designers, the International Facility Management Association, Johnson Controls, the National Electrical Manufacturers Association, the New York State Energy Research and Development Authority, Steelcase, Inc., the US Department of Energy, and the US Environmental Protection Agency. The Light Right Consortium and the researchers gratefully acknowledge the contributions of lighting equipment provided by (in alphabetical order) Birchwood Lighting, Cooper Lighting, Day-Brite Lighting, Engineered Lighting Products, General Electric Lighting, Ledalite Architectural Products, Lightolier, Lutron, Osram Sylvania, Peerless Lighting, and Philips Lighting. Furnishings were provided by Armstrong World Industries Inc. and Steelcase, Inc.

The researchers would like to thank Carol Jones, Judi Heerwagen, and Leslie Nicholls at PNNL for their many contributions to the project. Other staff at the National Research Council of Canada who contributed to this project were Chantal Arsenault, Roger Marchand, and Clinton Marquardt. At the Lighting Research Center, Martin Overington and Owen Howlett also contributed.