This article was downloaded by: [Northeastern University]
On: 25 November 2014, At: 19:04
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Journal of the Illuminating Engineering Society
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uzie20

Lighting Quality and Energy-Efficiency Effects on Task Performance, Mood, Health, Satisfaction, and Comfort

J.A. Veitch Ph.D.[a] & G.R. Newsham Ph.D.[a]
[a] National Research Council of Canada, Institute fin-Research in Construction, Ottawa, ON, Canada
Published online: 19 Sep 2013.

To cite this article: J.A. Veitch Ph.D. & G.R. Newsham Ph.D. (1998) Lighting Quality and Energy-Efficiency Effects on Task Performance, Mood, Health, Satisfaction, and Comfort, Journal of the Illuminating Engineering Society, 27:1, 107-129, DOI: 10.1080/00994480.1998.10748216

To link to this article: http://dx.doi.org/10.1080/00994480.1998.10748216

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Lighting Quality and Energy-Efficiency Effects on Task Performance, Mood, Health, Satisfaction, and Comfort

J.A. Veitch, Ph.D. and G.R. Newsham, Ph.D.

THIS PAPER WAS PRESENTED AT THE 1997 IESNA ANNUAL CONFERENCE
PAPER OF DISTINCTION

Downloaded by [Northeastern University] at 19:04 25 November 2014

Introduction

Many jurisdictions have legislated energy codes that restrict building energy consumption for all uses, including lighting.[1,2] Quality lighting systems today must meet or better these energy-use levels so that both the environment and resources are conserved (meeting long-term human needs), while still meeting immediate task, social, behavioral, aesthetic, emotional, health, and safety needs. Maintaining this balance is important to building owners, employers, and occupants, for lighting is an important feature in office design and furnishings.[3,4]

Although many lighting systems exist that meet energy-efficiency requirements, concern persists that more energy-efficient lighting design will result in poorer quality lighting. The absence of a common definition of lighting quality has been one impediment to progress on this issue. Veitch and Newsham[5,6] have proposed a behaviorally based definition of lighting quality. According to this definition, lighting quality exists when the luminous conditions are suitable for the needs of the people who will use the space. They grouped these needs in six categories: visual performance; post-visual performance (e.g., reading, eating, sewing, walking); social interaction and communication; mood state (happiness, alertness, satisfaction, preference); health and safety; and aesthetic judgments (assessments of the appearance of the space or the lighting).

This approach is consistent with other environment-behavior research, particularly that aimed at understanding relationships between working conditions and individual and organizational-level outcomes. For example, Rubin[7] conducted an extensive literature review and concluded that there existed no measures of organizational effectiveness and/or productivity relevant to micro-environmental design. He proposed that the best approach to developing such a measure would be to use a variety of subjective and objective measures.

A review of the literature revealed only a weak understanding of the relationships between luminous conditions and these human needs.[5,6] Poor research design, small sample sizes, and inappropriate use of statistical tests all mar the quality of the literature base. This experiment was designed to expand our understanding of lighting quality by relating lighting quality, as understood by lighting designers (designers' lighting quality, or DLQ), to behavioral measures from each of the six categories listed above. The experiment also examined the relationship between DLQ and lighting energy-efficiency, as represented by the lighting power density (LPD). Our working hypotheses were, first, that increasing DLQ would cause better outcomes on all dependent measures; and, second, that this relationship would be independent of LPD. That is, we predicted that low-LPD lighting systems, when selected and designed by professionals, could provide equivalent lighting quality to high-LPD systems. Any effect in which high-LPD systems produced better lighting than low-LPD systems would cast doubt on the wisdom of current energy-efficiency strategies for lighting.

Authors' affiliation: National Research Council of Canada, Institute for Research in Construction, Ottawa, ON, Canada

Method

This experiment was a 3 by 3 (DLQ by LPD) factorial between-subjects experiment. There were three levels of DLQ (low, medium, and high, described below), and for each DLQ level there was a low LPD (8.0-9.3 W/m^2 [0.7-0.9 W/ft^2]), medium LPD (11.3-13.3 W/m^2 [1.1-1.2 W/ft^2]) and high LPD (19.8-25.1 W/m^2 [1.8-2.3 W/ft^2]) variant.

Participants

Participants (142 women and 150 men, aged 18-61 years) were recruited through a temporary employment agency, which pre-screened potential participants to ensure that they met the following criteria: minimum 18 years of age; normal or corrected-to-normal vision; normal hearing; no mobility impairments; experience with Windows™-based word processing and spreadsheets; and, fluency in English. The participants were paid for their participation at the standard rate for the type of clerical work characterized by the tasks. Participants were recruited, informed, and debriefed according to the exacting standards of the institution's Human Subjects Research Ethics Committee.

Each participant was randomly assigned to one of the nine lighting conditions and worked under that lighting system for one day on a variety of computer- and paper-based office tasks and questionnaires. Random assignment to experimental conditions distributes individual differences equally across the groups; therefore, if between-group differences in outcome measures are found, they may be attributed to the experimental manipulation.[8,9]

Downloaded by [Northeastern University] at 19:04 25 November 2014

Setting

The experiment took place in a mocked-up 12.2 by 7.3 m (40 by 24 ft) office designed for acoustics, lighting, ventilation, and indoor air quality research,[10] furnished as a typical mid-level clerical or administrative office. For this experiment, the space was windowless (a false partition wall concealed the windows). Modular furniture systems with 1.67 m partitions (66 inches) provided six open-plan workstations of approximately 6 m² (65 ft²), with space for shared file cabinets and printers at the ends of the room. Each workstation had a 486 computer with a 14 inch color monitor (0.28 inch dot pitch). The furnishings and contents of the workstations were identical.

The office space is acoustically shielded from building noise; therefore, simulated ventilation noise was added to keep the ambient noise at 48 dB(A). Temperature, humidity, ventilation, and noise conditions were monitored continuously to ensure that they remained within the accepted range of conditions for offices. A dedicated reception room outside this space was used for initial instructions and for coffee and lunch breaks. The experimental manipulation consisted of changing the lighting system used in the office space. All other environmental conditions were constant.

Lighting conditions

Previous experiments of a similar nature used arbitrarily chosen lighting designs.[11,12] This study adopted a different tactic; three local lighting designers were hired to form a panel to design nine different lighting systems for the office space. They were instructed to produce lighting designs that met current recommended illuminance levels for offices with VDTs, that met the three LPD levels, and that could currently exist in open-plan offices in North America. The designers were free to develop their own definitions of low, medium, and high DLQ.

Through discussion, the designers determined that all high-DLQ conditions would consist of indirect or direct/indirect lighting systems, all medium-DLQ conditions would consist of semi-specular parabolic-louvered recessed troffers, and all low-DLQ conditions would consist of recessed prismatic lensed troffers. These definitions of DLQ levels were validated for North American designers in a survey.[13] Photographs taken at standing height at the entrance to one workstation are shown in **Figure 1**.

Table 1 identifies the lighting equipment for all nine lighting conditions. Note that all medium- and low-LPD conditions used electronic ballasts; all high-LPD conditions used magnetic ballasts. All low-LPD conditions included undershelf and angle-arm task lighting, which also used electronic ballasts. All lamps, including those used for task lighting, were 3500K and had a CRI of 80 or

Table 1—*Specifications for experimental conditions*

	LPD		
DLQ	**1 (low)**	**2 (medium)**	**3 (high)**
1 (low) *Recessed troffer with K12 prismatic lens*	electronic ballasts 1x4 ft fixture with 1-lamp T8 (x20) •angle-arm task lamps (x6) •undershelf task lamps (x6)	electronic ballasts 1x4 ft fixture with 2-lamp T8 (x20)	magnetic ballasts 1x4 ft fixture with 2-lamp T12 (x20)
2 (medium) *Recessed troffer with parabolic louver*	electronic ballasts 0.66x4 ft fixture with 1-lamp T8 (x20) •angle-arm task lamps (x 6) •undershelf task lamps (x 6)	electronic (ballasts) 1x4 ft fixture with 2-lamp T8 (x20)	magnetic ballasts with 2-lamp T12 (x25)
3 (high) *Indirect or direct/indirect*	electronic ballasts 4 ft long fixture with 2-lamp T8 (x 8) (furniture-mounted indirect) •angle-arm task lamps (x 6) •undershelf task lamps (x 6)	electronic ballasts 36 ft long fixture with 16 4 ft lamps T8 (x 2) (18 inch suspended direct/indirect)	magnetic ballasts 36 ft long fixture with 16 4 ft T8 lamps (x 3) (12 inch suspended indirect)

Lamp types: All T8 lamps: 32 W, 3500K, CRI ≥ 80. All T12 lamps: 40 W, 3500K, CRI ≥ 80. Articulated task lamps with 3500K 13 W compact fluorescents. Undershelf task lights with T8 2 ft, 3500K lamps. All ballasts were non-dimming.

Figure 1—Photographs of experimental conditions, with spot photometric measurements.

better. **Table 2** summarizes spot photometric measurements and power measurements for the nine lighting conditions.

Procedure and dependent variables

Participants arrived at the institution at 8:30 am, in same-sex groups of 3 to 6 people. They had read a simplified explanation of the study prior to their agreement to attend, and were given a more detailed explanation by videotape on arrival. Thus began the day, which proceeded as shown in the schedule shown in **Table 3**. The session was conducted by an experimenter who was blind to the hypotheses under test, although she was aware that different lighting systems were in use in various experimental sessions. Participants remained in the reception room until all had completed the baseline visual performance test (described below). In the office space, they were assigned at random to one of the six workstations. Participants were free to adjust the chair and keyboard height and angles, but were asked not to alter the posi-

Table 2—Photometric summary

Lighting Design		Desk Illuminance (lux)		Desktop luminance (cd/m²)		Partition luminance (cd/m²)		LPD (W/m²)
DLQ	LPD	left	right	left	right	upper	lower	
1	1	745	365	93	50	29	22	9.25
1	2	430	395	57	64	46	34	12.91
1	3	525	490	77	72	62	42	19.80
2	1	748	400	110	57	21	22	9.25
2	2	510	465	65	60	30	27	13.34
2	3	540	450	70	61	33	27	25.07
3	1	540	320	74	46	33	24	7.96
3	2	655	560	98	88	47	43	11.30
3	3	705	618	100	90	72	54	20.87

Note: LPD values are measured, including task lighting for the LPD=1 conditions.

Downloaded by [Northeastern University] at 19:04 25 November 2014

tion or appearance of the computer screen.

When the participants were seated comfortably, they proceeded to follow instructions presented on-screen. Custom software automated most of the data collection and controlled the duration of each task. The experimenter left the room, but monitored their progress from an adjacent control room using a security video camera system and communicated with the participants by telephone. At the conclusion of the final questionnaire, participants were debriefed by videotape and provided with written information about the purpose of the experiment.

The following sections describe the tasks grouped according to the outcome categories to which they relate (refer to **Table 3** for the sequence of presentation).

Demographics and individual differences—The initial questions at the start of the day concerned demographics and necessary background information such as age, vision, hearing, and smoking status. Other questions concerned personal and family health, including disorders such as migraine, high blood pressure, and epilepsy, which may relate to sensitivity to flicker.[14]

Other individual difference (personality, beliefs, and attitudes) variables were included because of their potential to bias the results. Actual and desired influence over the physical environment in everyday settings were assessed using the Survey of Personal Influence in Common Environments.[15] The other such variables were environmental sensitivity,[16] beliefs about the effects of the physical environment on people,[17] and beliefs about the

Table 3—Schedule of experimental sessions

Approx. Time	Task	Duration (min)	
8:45 a.m.	Arrival, greeting, instructions, consent, visual performance (VALiD), assign to workstations	45	
9:30	Demographics, health history	5	
9:35	Mood questionnaire (baseline)	5	
9:40	Reaction time (instructions and task)	12	
9:52	Typing (instructions and task)	42	
10:35	*Break (Reception room)*	*15*	
10:50	Reading Comprehension (paper-based)	30	
11:20	Proofreading (instructions and task)	32	
11:52	Job Candidate Evaluation (paper-based)	20	
12:10 p.m.	Physical Sensations Questionnaire	5	
12:15	*Lunch (Reception room)*	*45*	
1:00 p.m.	Typing	40	
1:40	Proofreading	30	
2:10	Reaction time	10	
2:20	Mood Questionnaire	5	
2:25	Aesthetics and Satisfaction Questions	15	
2:40	Physical Sensations Questionnaire	5	
2:45	*Break (Reception room)*	*15*	
3:00	Creative Writing	30	VALiD 2nd test in parallel
3:30	Survey of Personal Influence in Common Environments	15	
	Environmental Sensitivity Scale	5	
	Person-Surroundings Scale	5	
	Lighting Beliefs Questionnaire	10	
4:05	Typing Probe (after all participants have done VALiD): gray text on gray background	10	
4:15	Workday Experiences Questionnaire	10	
4:30 p.m. (approx.)	*Debriefing and farewell*	*30*	

Note: The reading comprehension, job candidate evaluation, and creative writing tasks all required that participants leave their workstations to retrieve or return materials to shared file cabinets. This procedure ensured that participants viewed the entire space over the course of the day.

Downloaded by [Northeastern University] at 19:04 25 November 2014

Downloaded by [Northeastern University] at 19:04 25 November 2014

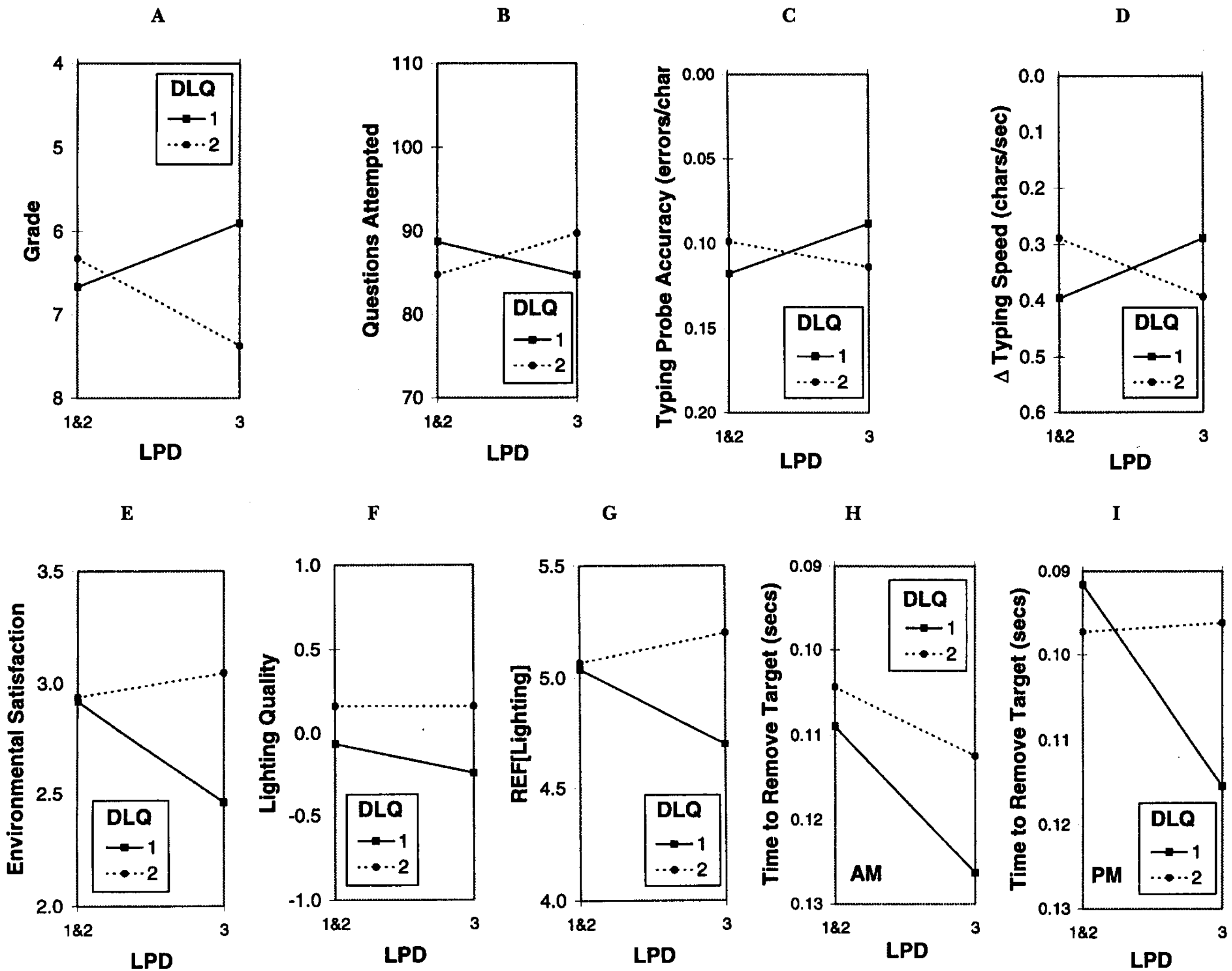

Figure 2—Graphic presentation of significant LPDA x DLQB interactions. (A) Writing, (B) Reading, (C) Typing probe accuracy (D) Typing probe speed, (E) Environmental satisfaction, (F) Lighting quality, (G) REF, and (H) & (I) Psychophysical performance (LPDA x DLQB x Time). Note that y-axis scales are oriented with preferable outcomes towards the top. All scale values are raw (untransformed scores).

effects of lighting on people.[18] These four questionnaires were presented on the computer in the late afternoon, so that they would not influence the more important performance, mood, and satisfaction responses.

Visual performance—Visual performance was measured using the Vision and Lighting Diagnostic kit (VALiD),[19] both at the start and end of the day. The testing was performed individually in a private room off the reception room. Incandescent lamps provided 200 lx on the task for this test. The VALiD test was conducted under constant conditions, rather than as a test of visibility under the test conditions, to provide an examination of aftereffects of working under the various lighting conditions.

Performance measures—A variety of computer-based and paper-based tasks were used to assess verbal/intellectual and clerical work performance. Verbal/intellectual tasks were represented by a standardized reading comprehension test[20] and a creative writing task. The creative writing task used the images from the Thematic Apperception Test[21] as stimuli; participants wrote stories to explain the action in the images. This task was used by Nelson et al.[22] but was modified here in that the stories were typed on the computer instead of written in longhand. The intellectual tasks were presented once each during the day.

Clerical work was measured using computerized typing and proofreading tasks. Both tasks were presented in both morning and afternoon sessions. Each 40-min typing session consisted of two 20-min sessions with either white characters on a blue background (negative polarity), or black characters on a white background (positive polarity). Proofreading sessions were 30 min long, with 15 min in each polarity condition. Order of presentation of the two polarity conditions was counterbalanced. At the end of the day there was an extra 10-min period of typing using dark gray text on a light gray background, which we labeled the "Typing Probe." This low-contrast task was chosen to provide a severe test of the lighting effects on task performance.

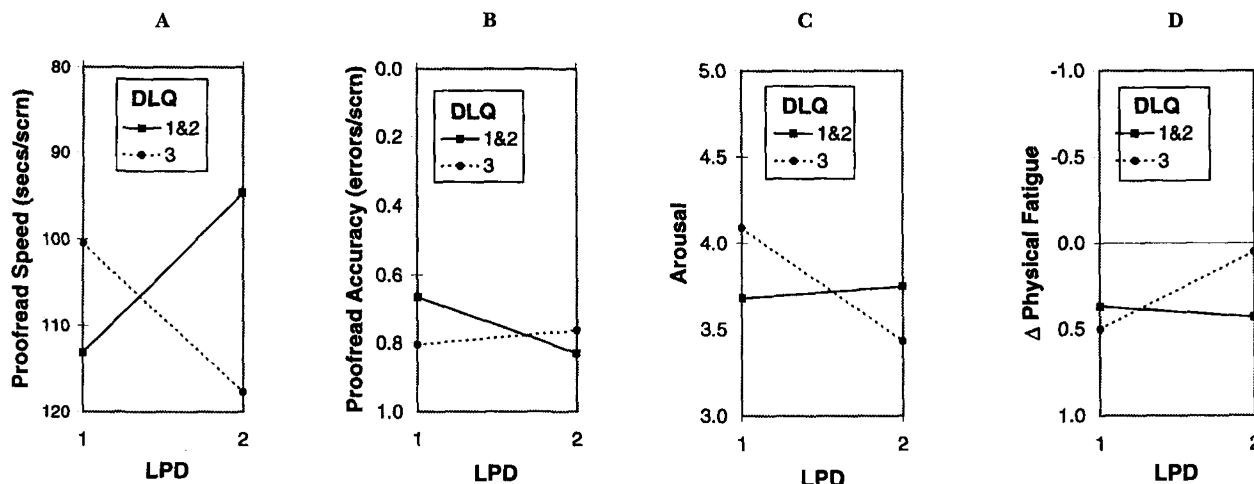

Figure 3—Graphic presentation of significant LPDB x DLQA interactions. (A) Proofreading speed, (B) Proofreading accuracy (C) Mood (arousal), and, (D) Physical fatigue.

Downloaded by [Northeastern University] at 19:04 25 November 2014

Regardless of the screen conditions, the typing task required participants to exactly retype model text, presented in a window at the top of the screen, into a window at the bottom of the screen. Model text was presented a paragraph at a time, and on-screen instructions prompted participants to press a function key to advance to the next paragraph when ready. When the typist made a mistake, the computer would beep and would not accept further new keystrokes until the error had been corrected.

Model text for the typing task was taken from a report of a government department and concerned the implementation of a computer system. Following consultation with peers, the text was modified slightly by the introduction of random errors at a mean rate of 1 error per 200 characters. Because the software was configured to require exact replication of the model text, the errors in the model text required that participants pay close attention to the text; even expert typists were expected to have some errors.

The proofreading task was a modified form of the numerical verification task developed by Rea.[23] Each screen presented 20 lines of text. Each line displayed a 10-character string, a checkbox, and another 10-character string. The second string was either identical to the first, or differed by one character. The strings contained upper- and lower-case letters and numbers (pilot testing suggested that the five-digit strings used by Rea resulted in a task that was too easy for this sample). Participants marked the lines in which the two strings differed, using the spacebar to mark an X in the checkbox. Each screen contained from none to six differences. At the completion of each screen, the participants pressed a function key to advance to the next screen.

Self-ratings of task performance were included in the Workday Experiences Questionnaire at the end of the day. Ratings of the difficulty of the tasks and questionnaires were provided on a five-point scale. Participants also estimated the effect that the physical environment in this setting might have had on their performance in comparison to other places they were accustomed to, using a nine-point scale from "-40 percent or more," up in 10 percent increments to "+40 percent or more" (for which the midpoint was "0 percent").

Social behavior—Although the tasks were individual in nature, two indirect indices of social behavior were

Table 4—*Summary of planned comparisons for MANOVA*

Comparison	Label	Levels tested	Comments
Main effects	**DLQA**	**DLQ3 vs (DLQ1 and 2)**	**direct/indirect vs recessed troffers**
	DLQB	**DLQ1 vs DLQ2**	**prismatic lens vs parabolic louver**
	LPDA	**LPD3 vs (LPD1 and 2)**	**magnetic ballast vs electronic ballast**
	LPDB	**LPD1 vs LPD2**	**task lighting vs no task lighting**
Interactions	**DLQA x LPDA**		**interaction of main effects**
	DLQA x LPDB		
	DLQB x LPDA		
	DLQB x LPDB		

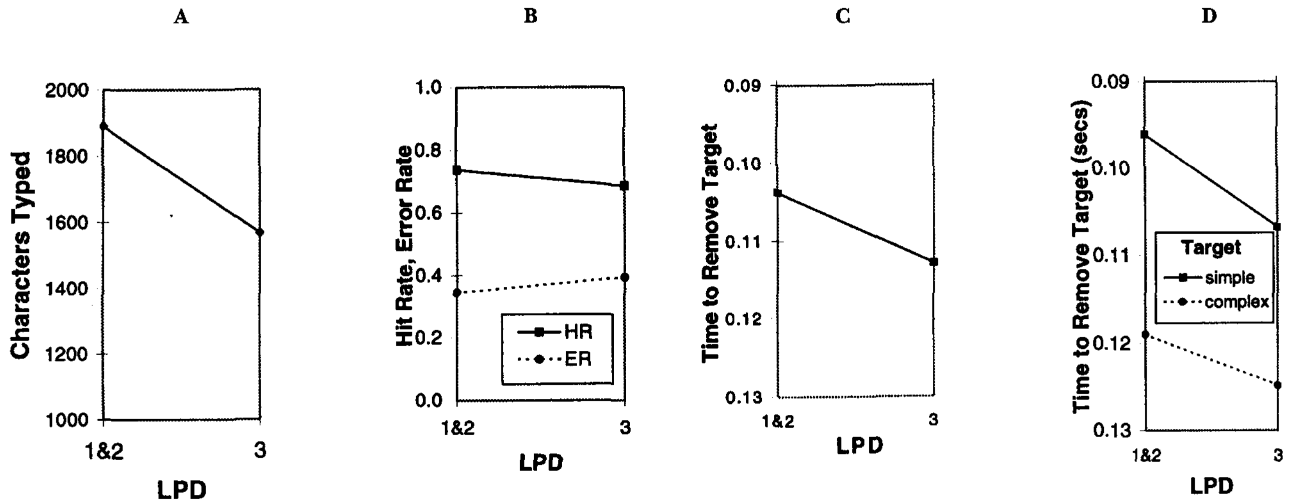

Figure 4—Graphic presentation of significant LPDA main effects. (A) Writing, (B&C) Psychophysical performance, and (D) Psychophysical performance: LPDA x Target.

included. Participants evaluated fictional job candidates and assigned them a starting salary based on their resumes and job descriptions; this task is thought to be an analogue of interpersonal evaluations that occur in offices.[24] This task was entirely paper-based. The second index of social behavior was willingness to help,[25] measured as the number of hours the participants were willing to donate to voluntarily participate in another experiment. This measure has been previously shown to be sensitive to changes in lighting conditions.[26]

Mood, satisfaction, and psychophysical performance—This category includes all manner of emotional (affective) reactions to the work and to the workplace. These include satisfaction with the lighting and the physical features of the workplace, mood, and psychophysical performance.

Satisfaction was measured in the afternoon, when participants had had time to form judgments about the space and its suitability for computer- and paper-based work. A four-item scale assessed lighting quality in terms of the light available for reading, overall satisfaction with the lighting at the workspace, the adequacy of the quantity of light and the effects of the light on the participant's ability to work; this was taken from previous research.[27] Environmental satisfaction (satisfaction with the fit between the physical workplace and the requirements of the job) was the second such measure; this four-item scale also has been used previously.[28] In addition, participants completed a 23-item scale rating the effects of physical environmental features on their ability to work.[29]

Affective and psychophysical changes over the work-

Table 5—*Data Means and standard deviations for significant verbal-intellectual effects*

Effect	Writing-Grade		Writing-Words/sen		Writing-Char/word		Writing-characters		Reading Questions Attempted	
					F(1,281)	*R^2*	*F(1,281)*	*R^2*	*F(1,281)*	*R^2*
*LPDA**					*3.96*	*0.014*	*6.70*	*0.023*	*4.07*	*0.014*
Group					**M (SD)M (SD)**		**M (SD)**			
LPD1 and 2					**4.14 (0.29)**		**1892(915)**		**87.4 (15.5)**	
LPD3					**4.06 (0.41)**		**1568(634)**		**84.4 (15.2)**	
	F(1,281)	*R^2*	*F(1,281)*	*R^2*					*F(1,281)*	*R^2*
*LPDA x DLQB***	*5.14*	*0.018*	*8.43*	*0.029*					*4.87 0.017*	
Group	**M (SD)**		**M (SD)**						**M(SD)**	
LPD1 and 2/DLQ1	**6.7(2.6)**		**16.3(6.3**						**88.7(15.2)**	
LPD1 and 2/DLQ2	**6.3(1.9)**		**14.9(4.2)**						**84.8(15.5)**	
LPD3/DLQ1	**5.9(1.9)**		**14.1(4.2)**						**84.7(13.9)**	
LPD3/DLQ2	**7.4(3.2)**		**17.5(7.1)**						**89.7(12.3)**	

Note. Only the statistically significant effects are shown. R^2 is the proportion of variance accounted for. For multivariate effects, it is the arithmetic mean of all univariate effect sizes. *Wilks Λ l =0.941; F(7,275)=2.45; R^2_{ave} =0.008. **Wilks' Λ l =0.938; F(7,275)=2.62; R^2_{ave} =0.011. Cells display raw scores. Analyses were performed on transformed scores, as follows: Writing - Grade = √(mean grade across para.). Writing - Words/sentence = Ln(mean words/sentence across para.). Writing - Characters/word = $(5.1 - \text{mean characters/word across para.})_{0.25}$. Writing - Total Characters = √(mean total characters across para.). Reading - Attempts' = √(107-reading questions attempted).

Downloaded by [Northeastern University] at 19:04 25 November 2014

day in this category of outcomes were assessed in two ways: with a mood questionnaire and with a computerized psychophysical task combining target detection and reaction time (thus, there were both self-report and behavioral measures of activation or fatigue). The mood questionnaire consisted of 18 semantic differential pairs representing three components of mood: pleasure, arousal, and dominance.[30] Both measures were presented in both the morning and afternoon.

The psychophysical task was developed specifically for this experiment. It was a simplified version of tasks common to video games. Colored symbols (blue or red circles, squares, and clubs) appeared at a random height at the left edge of the computer screen and traveled horizontally from left to right on a black background, simulating objects travelling on a conveyor belt. The top and bottom of the screen were white. Grey dotted lines defined a rectangular removal zone. The participants' task was to remove target symbols as quickly as possible after they had entered the removal zone. Pressing the spacebar after a target symbol had entered the zone removed that symbol, but had no effect on non-target symbols. Within each 10-min block of this task were three 200-sec sections. The first two sections featured simple targets: targets were either all symbols of one color, or all of one shape. The third section of each block was always a complex target, identified by both its color and its shape (e.g., red clubs). Within each section, targets traveled for 100 sec at 0.39 m/sec and for 100 sec at 0.46 m/sec.

Health—The one-day exposure did not permit examination of long-term health consequences of the lighting. However, participants did complete a 17-item scale to report their physical comfort at the end of both the morning and afternoon sessions, using a five-point scale to report the severity of symptoms such as headache, eyestrain, sore wrists or arms, and general fatigue. This questionnaire was adapted for this experiment from others in the literature.[31,32] Thermal comfort was also measured, using the ASHRAE and McIntyre scales,[33] at mid-morning and mid-afternoon.

Aesthetic impressions—Judgments about the aesthetic appearance of the room were assessed using 27 semantic differential pairs (**Tables 1–6**) adapted from previous research to make them appropriate to this study.[12,34] Participants were instructed to base their responses on their opinions of the appearance of the room. We judged that an instruction to focus on the lighting alone would be meaningless to these naive participants; moreover, we wished to avoid creating response biases.

Manipulation checks—At the end of the session, participants were asked to respond to three open-ended questions about their beliefs about the purpose of the study, the expectations of the researchers, and factors that might have influenced their responses. These deliberately vague questions provide an indication of the extent to which the participants' experiences during the day might have biased their responses. The information thus provided can also be useful for the development of new hypotheses for future research.

Results

This experiment used a straightforward analysis of variance design. There were two independent variables,

Table 6—*Data means and standard deviations for significant typing and proofreading effects*

Effect	Typing Speed		Proofreading Speed		Proofreading Errors	
			F (1,276)	*R^2*	*F (1,276)*	*R^2*
*Time**			*71.97*	*0.207*	*30.61*	*0.100*
Group			***M (SD)***		***M (SD)***	
Morning			**111 (38)**		**0.680 (0.540)**	
Afternoon			**104 (43)**		**0.795 (0.580)**	
	F (1,276)	*R^2*	*F (1,276)*	*R^2*		
*Screen***	*19.71*	*0.066*	*28.04*	*0.092*		
Group	***M (SD)***		***M (SD)***			
White on Blue (-ve)	**2.65 (1.03)**		**110 (41)**			
Black on White (+ve)	**2.70 (1.05)**		**106 (40)**			
			F (1,276)	*R^2*	*F (1,276)*	*R^2*
LPDB x DLQA•			*7.52*	*0.027*	*3.81*	*0.014*
Group			***M (SD)***		***M (SD)***	
LPD1/DLQ1 and 2			**113 (43)**		**0.665 (0.563)**	
LPD1/DLQ3			**100 (33)**		**0.805 (0.524)**	
LPD2/DLQ1 and 2			**95 (37)**		**0.829 (0.525)**	
LPD2/DLQ3			**117 (54)**		**0.762 (0.583)**	

*Note. *Wilks' Λ =0.760; F(4,273)=21.55; R^2_{ave} =0.078. **Wilks' Λ =0.853; F(4,273)=11.74; R^2_{ave}=0.042. •Wilks' Λ 0.964; F(4,273)=2.58; R^2_{ave}=0.012. Cells display raw scores. Analyses were performed on transformed scores, using the following transformations: Typing scores: Speed=$\sqrt{}$(characters/sec); Errors=$\sqrt{}$errors/character); Proofreading scores: Speed=Ln(secs/screen); Errors=Ln(1+errors/screen).*

Downloaded by [Northeastern University] at 19:04 25 November 2014

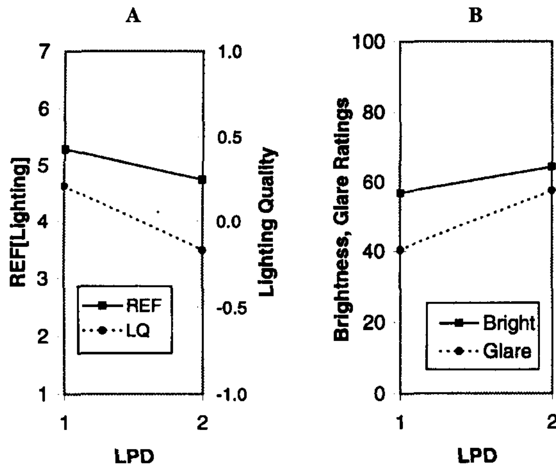

Figure 5—Graphic presentation of significant LPDB main effects. (A) Satisfaction ratings and (B) room appearance ratings.

DLQ and LPD, each having three levels. This allowed the eight orthogonal planned comparisons summarized in **Table 4**. This set of comparisons was applied to conceptually related groups of dependent variables (e.g., three mood subscales; or, measures of typing speed and accuracy) using multivariate analyses of variance (MANOVA) to minimize Type I error. This technique gives an overall test of significance for all dependent measures and individual, univariate significance tests. We interpreted significant univariate tests only if the multivariate test had also reached significance (p<0.05).

Prior to analysis, we examined the individual differences and demographic data. There were no between-groups differences on the personality variables (environmental sensitivity, desire for environmental control, actual environmental control, person-surroundings scale, lighting beliefs), nor on measures of individual or family health. Random assignment effectively equated the groups on these variables. However, we did note

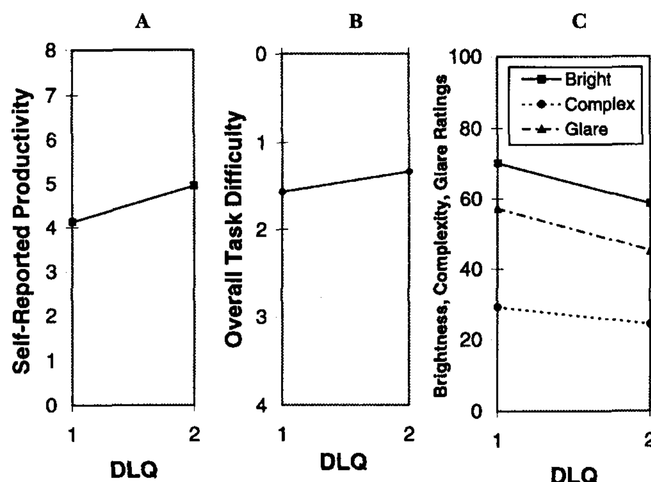

Figure 6—Graphic presentation of significant DLQB main effects. (A) Performance rating: self-reported productivity, (B) Performance rating: overall task difficulty, and (C) Room apperance ratings.

between-groups differences in age. Therefore, it was decided to use age as a covariate in analyses of variables for which age might bias the outcome (e.g., visual performance data).

Many of the dependent variables were not normally distributed; 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, analyses proceeded with the original data. The number of participants included varies slightly from one analysis to another because of missing data.

The following sections summarize the statistical analyses for each outcome category, with **Tables 5-10** providing the details for these analyses. Important statistically significant effects are shown graphically in **Figures 2-6**, grouped by interactions and main effects; this summary of outcomes will be used for the final discussion and interpretation of results. Detailed results are available in the project report.[35]

Visual performance

The morning and afternoon visual performance scores each consisted of the number of correctly identified Landolt rings (maximum 13) for each of six contrast rows. Morning scores (as a measure of pre-existing visual abilities) and age were used as covariates, a statistical technique for controlling the effects of these variables on the dependent measures, which were the afternoon scores.

There was one statistically significant multivariate effect, LPDA (Wilks' Λ=0.955, F(6,269)=2.14, R^2_{ave} =0.007.).* This contrast compares lighting conditions using magnetic ballasts (LPD3) to those using electronic ballasts (LPD1 and 2), regardless of luminaire type. One univariate test reached statistical significance, performance on rows E/F (contrast 0.31) (F (1,274)=5.97, R^2=0.02). The mean performance under magnetic ballasts (M=11.20, SD=0.74) was 2.2 percent lower than under electronic ballasts (M=11.46, SD=0.75).

Task performance

Verbal-intellectual tasks—There were three reading scores: the number of questions attempted (maximum 107), the number of questions correctly answered (maximum 107), and the percentage of attempted questions

R^2 is the proportion of variance accounted for, and is a measure of effect size. For multivariate effects, it is the arithmetic mean of all univariate effect sizes. In behavioral research, small effect sizes account for 1-5 percent of the variance; medium effect sizes account for approximately 9-11 percent, and large effect sizes account for 25 percent or more.[45]

Downloaded by [Northeastern University] at 19:04 25 November 2014

correctly answered (maximum 100). For the writing task, the stories written in response to the pictures were analyzed using Grammatik™ IV, a grammar-checking software package,[36] to generate scores for grade level,* number of words per sentence, number of characters/word, and total number of characters. The scores for each story were averaged to give scores for the individual's writing performance.

The results of this MANOVA are displayed in **Table 5** (we report raw scores for easier interpretation, although transformed data were analyzed). There was a significant LPDA by DLQB interaction and a significant LPDA main effect. The effects were small, explaining 1-2 percent of the variance in scores.

Typing and proofreading—Typing speed was measured in net characters per second (absolute speed corrected for error frequency), and typing accuracy was the number of errors per correct keystroke. Proofreading speed was the length of time (seconds) spent checking each screen of 20 lines; accuracy was the number of errors per screen. The analysis was a 3 by 3 by 2 by 2 (DLQ by LPD by Screen Polarity by Time) MANOVA. Screen and time were within-subjects variables; DLQ and LPD were between-subjects variables.

The within-subject variables showed expected main effects. Participants were faster, but made more errors, on the proofreading task in the afternoon (**Table 6**). These changes likely reflect a combination of practice and fatigue. For both typing and proofreading, the speed was higher with positive polarity (black on white), a finding that is consistent with previous research.[37] One lighting interaction effect was significant, LPDB by DLQA (**Figure 3**). This effect on proofreading performance showed that for LPD1 systems, the lowest speed and lowest error rate occurred with lensed or louvered (direct-only) systems (DLQ1 and 2); for LPD2, the lowest speed and lowest error rate occurred with the direct/indirect system (DLQ3). These interaction effects were small in terms of explained variance, but reflect mean differences on the same order as the time-of-day effect.

Typing probe—At the end of the day, participants spent 10 min doing very difficult typing: dark gray text on a light gray background. These data were analyzed separately in a 3 by 3 (LPD by DLQ) MANOVA, using four dependent variables: speed and accuracy as before, and the difference between performance on the typing probe and their typing speed and accuracy for the first 10-min block in the afternoon typing session. Performance on the low-contrast task was always lower than for the high-contrast task, as one might expect. Both the LPDA by DLQB and LPDB by DLQB interactions had statistically significant multivariate effects (**Table 7**).

Task difficulty and productivity ratings—The eight end-of-day estimates of the difficulty of the tasks and questionnaires were scored on five-point scales with minimum 0 and maximum 4. The productivity estimate was one question, and used a nine-point scale (0-8). These nine dependent variables were analyzed using the 3 by 3 MANOVA model and the planned comparisons. This

Software calculates the grade level of written work using the following formula: (0.39 - average number of words/sentence) + (11.8 - average number of syllables per word). Writing in the range of grade 6-10 is considered most readable, the range in which these participants wrote. More complex writing may be less effective; therefore, we assumed that for this task, lower grade scores would be preferable to higher ones.

Table 7—*Typing probe significant effects, data means and standard deviations*

Effect	Typing speed - difference		Typing errors -probe	
	F (1,283)	R^2	*F (1,283)*	R^2
*LPDA x DLQB**	*3.92*	*0.014*	*7.95*	*0.027*
Group	***M (SD)***		***M (SD)***	
LPD1 and 2/DLQ1	**0.40 (0.44)**		**0.118 (0.075)**	
LPD1 and 2/DLQ2	**0.29 (0.41)**		**0.099 (0.056)**	
LPD3/DLQ1	**0.29 (0.40)**		**0.088 (0.051)**	
LPD3/DLQ2	**0.39 (0.35)**		**0.114 (0.050)**	
	F (1,283)	R^2		
*LPDB x DLQB***	*5.43*	*0.019*		
Group	***M (SD)***			
LPD1/DLQ1	**0.29 (0.43)**			
LPD1/DLQ2	**0.31 (0.37)**			
LPD2/DLQ1	**0.52 (0.43)**			
LPD2/DLQ2	**0.27 (0.45)**			

*Note. * Wilks' Λ =0.948; F(4,280)=3.87; R^2_{ave}=0.010. ** Wilks' Λ = 0.957; F(4, 280)=3.18; R^2_{ave}=0.007. Cells display raw scores. Analyses were performed on transformed scores, as follows: Speed Difference = (char/sec[BW] - char/sec[GG] +0.55)$^{0.30}$, Typing Errors-Probe = (errors/char[GG])$^{0.33}$. BW=high contrast (first 10 min, afternoon session). GG=low contrast (typing probe session).*

Downloaded by [Northeastern University] at 19:04 25 November 2014

analysis revealed main effects for DLQB and LPDA

The multivariate DLQB main effect (Wilks' Λ =0.920; F(9,269)=2.59, R^2_{ave}=0.009) was associated with statistically significant univariate effects on two variables: overall task difficulty (F (1,277)=4.36, R^2=0.015), and productivity effect estimate (F (1,277)=10.15, R^2=.035). DLQ2 (parabolic louvered systems, M=1.33, SD=0.64) groups rated the tasks as easier than DLQ1 groups (prismatic lenses, M=1.57, SD=0.78), and also estimated that the physical environment had a more beneficial effect on their productivity that day (M=4.96 [SD=1.68] vs M=4.13 [SD=1.81]). The mean difference between the groups in productivity effect estimates corresponds to 0.83 scale units. Each unit represents a 10 percent change in productivity, suggesting that the participants who experienced the DLQ2 conditions believed that the physical environment had improved their productivity by 8.3 percent more than participants who experienced the DLQ1 conditions.

The multivariate LPDA main effect (Wilks' Λ =0.929; F(9,269)=2.29, R^2_{ave}=0.004) was associated with a significant univariate effect on one variable, vision tests' difficulty (F (1,277)=7.80, R^2=0.027). Groups whose lighting had used electronic ballasts (LPD1 and 2, M=1.10 [SD=1.00]) rated the vision tests as slightly more difficult than those whose lighting had used magnetic ballasts (LPD3, M=0.79 [SD=0.89]). This finding is unexpected, given the visual performance results reported above.

Social behavior

Job candidate evaluation—Participants rated four fictional job candidates for skill, competence, and intelligence, and assigned a starting salary to each. There were no statistically significant effects on the ratings of job candidates willingness to volunteer. The overall mean number of hours that participants were willing to volunteer to serve as a participant in future investigations was 3.4 hours (SD=3.9), and responses covered the range from 0 to 10 hours. These data were analyzed using a 3 by 3 ANOVA design. There were no statistically significant effects on this variable.*

Mood, satisfaction and psychophysical performance

Mood—Arousal, pleasure, and dominance were the three mood scores, each calculated as the mean response on six semantic differential pairs. The range of possible scores is from 0 to 8. Internal consistency reliability values for the scales were consistent with previous research.[30,38]

The scale scores were analyzed in a 3 by 3 by 2 model (DLQ by LPD by Time). Multivariate effects for time and for the LPDB by DLQA interaction were found. The time effect (Wilks' Λ =0.338; F(3,281)=183.39; R^2_{ave}=0.337) revealed that all the mood measures were lower in the afternoon than in the morning. The drop in pleasure scores was large (F (1,283)=543.04, R^2=0.657); participants were moderately lower in feelings of dominance (F (1,283=)91.14, R^2=0.244), and the smallest drop occurred in the arousal scale (F (1,283)=34.97, R^2=0.110).†

The significant LPDB by DLQA interaction (Wilks' Λ = 0.968; F(3, 281)=2.13; R^2_{ave}=0.011) was associated with changes in Arousal (F (1,283)=6.74, R^2=0.023). Systems with an indirect component (DLQ3) had higher arousal when combined with task lighting (LPD1) (M=4.09 [SD=1.27] for LPD1/DLQ3, vs M=3.43 [SD=1.31] for LPD2/DLQ3), but direct systems (DLQ1 and 2) had higher arousal without task lighting, (M=3.75 [SD=1.20] for LPD2/DLQ1 and 2, vs M= 3.68 [SD=1.23] for LPD1/DLQ1 and 2) (**Figure 4**).

Satisfaction—Three sets of ratings assessed aspects of satisfaction: Ratings of Environmental Features, Environmental Satisfaction, and Lighting Quality. For Environmental Satisfaction, possible scores ranged from 0 to 4. Its internal consistency reliability was good (Cronbach's alpha=0.83). The Lighting Quality scale items[27] included two different response scales (two items on 1-5, and two 1-4); therefore, the individual item scores were converted to standardized scores prior to averaging to obtain the scale score. This gave scale scores with a theoretical range of from negative to positive infinity (in practice, they ranged from -2.21 to 1.22); higher values reflect better lighting quality. The zero point does not imply neutral lighting quality; rather, it is the mean point for this set of data.

The 23-item Ratings of Environmental Features (REF) covered all aspects of the physical environment. Principal components analysis (PCA) with Varimax rotation reduced this to five components explaining 59.9 percent of the variance. The five components were: noise (four questions), ventilation (four questions), furniture (five questions), washrooms (two questions), and lighting (two questions). Six questions did not load on any of the components. Subscale scores to correspond to the components were calculated by taking the mean response for the questions that loaded on the component; thus, the subscale scores have the range 1-7, with higher values indicating a better effect of the environment on the ability to work. All but the Washroom scale had acceptable or good internal consistency reliability.

**The absence of lighting effects on these indirect measures of social behavior might reflect the relative isolation in which the participants worked, with social contacts only during the coffee breaks and lunch period. To fully assess lighting influences on social behavior will require research designs that target these processes.*

†For Pleasure, a.m. M=5.68 (SD=1.26); p.m. M=3.67 (SD=1.54). For Dominance, a.m. M=4.08 (SD=1.08), p.m. M=3.49 (SD=1.08). For Arousal, a.m. M=3.96 (SD=0.99), p.m. M=3.44 (SD=1.34).

Downloaded by [Northeastern University] at 19:04 25 November 2014

Thus, MANOVA of the satisfaction data included seven variables: REF [Noise]; REF [Ventilation]; REF [Furniture]; REF[Washrooms]; REF [Lighting], Environmental Satisfaction, and Lighting Quality. **Table 8**, the summary of these results, shows an LPDA by DLQB interaction effect and an LPDB main effect. These effects are small, explaining 1-2 percent of the variance. The mean differences are also not large; however, they are similar in size to those of Collins et al.[27] for comparably lit workstations.

Psychophysical performance—The psychophysical performance measures were derived from performance on the video-game task. Three measures were used: hit rate (proportion of correctly hit targets), error rate (proportion of spacebar presses that failed to remove a target), and time-to-hit (elapsed time from the entry of the target into the removal zone until its removal). Recall that the task occurred in both morning and afternoon, at slow and fast speeds, and with three target types. Thus, the analysis was a 2 by 2 by 3 by 3 by 3 (Time by Speed by Target by DLQ by LPD) MANOVA. Time, Speed, and Target were within-subjects variables; DLQ and LPD were, of course, between-subjects variables. Interactions involving more than three independent variables were not examined, as these were likely to be uninterpretable.

The within-subjects variables showed expected relationships and large effect sizes[35]: performance grew more accurate (higher hit rate) and faster with experience (Multivariate Time: Wilks' Λ=0.736, F(3, 281)=33.60, R^2_{ave}=0.09). The speed at which targets moved had a large effect on both hit and error rates, and a moderate effect on the time-to-hit (Multivariate Speed: Wilks' Λ =0.471, F(3, 281)=105.311, R^2_{ave}=0.31). Time and speed interacted: the effect of speed was larger in the morning than in the afternoon for both hit rate and error rate, as would be expected for a practice effect. However, the effect of speed on time-to-hit was larger in the afternoon than in the morning, which suggests that fatigue was also occurring (Multivariate Speed by Time: Wilks' Λ =0.869, F(3, 281)=14.14, R^2_{ave}=0.053).

Target complexity had a very large effect on all three measures: performance was less accurate and slower when the targets were distinguished by both their shape and color, than by either characteristic alone (Multivariate Target Complexity: Wilks' Λ =0.105, F(3, 281)=799.47, R^2_{ave}=0.78). Also, colored targets were more accurately and more quickly identified than shape targets, particularly when the target speed was fast (Multivariate Shape/Color contrast: Wilks' Λ =0.433, F(3, 281)=122.67, R^2_{ave}=0.27; Multivariate Shape/Color By Speed contrast: Wilks' λ =0.306, F(3, 281)=212.43, R^2_{ave}=0.48).

Table 9 summarizes the significant lighting effects. The lighting effects on these measures are complex, and all the effects are small in terms of explained variance. There was one significant multivariate three-way interaction, LPDA by DLQB by Time, which was associated with a significant univariate effect for the time-to-hit measure only (**Figure 2**). Although all groups improved their afternoon performance relative to the morning (shorter times to remove target), the LPD3/DLQ1 group showed the smallest improvement, and the combined LPD1 and 2/DLQ1 groups showed the greatest improvement.

The significant multivariate LPDA contrast was mirrored by small effects for all three measures. Performance was more accurate both in terms of hit rate and error rate, and faster, for participants in LPD1 and LPD2 than those in LPD3. For the time-to-hit measure, this main effect was moderated by target complexity. The effect of complex targets was much greater for the LPD1 and LPD2 groups than for the LPD3 groups.

Finally, there was a statistically significant interaction

Table 8—Summary of significant lighting quality and environmental satisfaction effects

Effect	*REF noise*	*REF ventilation*	*REF furniture*	*REF lighting*	*Lighting quality*	*Environmental satisfaction*
		F(1,280) R^2	*F(1,280)* R^2	*F(1,280)* R^2	*F(1,280)* R^2	*F(1,280)* R^2
*LPDB**		*4.00* *0.014*	*4.77* *0.018*	*7.13* *0.025*	*9.26* *0.032*	*4.98* *0.017*
Group		***M (SD)***	***M (SD)***	***M (SD)***	***M (SD)***	***M (SD)***
LPD1		**5.04(0.93)**	**5.69(0.78)**	**5.29 (1.24)**	**0.21 (0.76)**	**3.00(0.67)**
LPD2		**4.75(0.96)**	**5.43(0.93)**	**4.74 (1.48)**	**-0.17(0.92)**	**2.76(0.82)**
	F(1,280) R^2			*F(1,280)* R^2	*F(1,280)* R^2	*F(1,280)* R^2
*LPDA x DLQB***	*4.04* *0.014*			*5.37* *0.019*	*4.13* *0.015*	*6.17* *0.022*
Group	***M (SD)***			***M (SD)***	***M (SD)***	***M (SD)***
LPD1 and 2/DLQ1	**4.53 (1.24)**			**5.18 (1.44)**	**0.05 (0.86)**	**2.92(0.81)**
LPD1 and 2/DLQ2	**4.46 (1.15)**			**5.05 (1.38)**	**0.12 (0.87)**	**2.94(0.69)**
LPD3/DLQ1	**4.17 (1.10)**			**4.45 (1.47)**	**-0.42(0.89)**	**2.46(0.78)**
LPD3/DLQ2	**4.75 (0.94)**			**5.22 (1.12)**	**0.14(0.73)**	**3.05(0.65)**

Note. REF scales from 1-7. Lighting Quality is on a continuous scale centered at zero. Environmental Satisfaction is on a scale from 0 to 4. In all cases, higher positive values indicate more favorable ratings. * Wilks' Λ =0.949, F(7,274)=2.10, R^2_{ave}=0.016. **Wilks' Λ =0.942, F(7,274)=2.40, R^2_{ave}=0.013.

Downloaded by [Northeastern University] at 19:04 25 November 2014

of DLQA by target complexity. Complex targets, as compared to simple targets, were hit more slowly by participants in the DLQ1 and DLQ2 conditions (direct-only systems) than by those in the DLQ3 conditions.

Health

Change scores (afternoon-morning) were calculated for the 13 items for which there were symptoms reported, and these data were reduced to two subscales using PCA: General Fatigue (tired eyes, irritated eyes, mental fatigue, dry eyes, tired), and Physical Fatigue (sore back, sore wrists/arms). Scores on these variables represented the average change in symptoms from morning to afternoon on their component items; possible scores range from -4 to +4, with positive scores indicating more severe symptoms in the afternoon.

There was one significant multivariate effect: LPDB by DLQA (Wilks' Λ =0.978, F(2, 281)=3.15, R^2_{ave}=0.01), associated with the Physical Fatigue subscale (F (1,282)=5.95, R^2=0.021). On average, all groups increased in fatigue in the afternoon over the morning (all means >0). Within the LPD1 conditions (which included task lighting), there was a very slightly greater increase in physical fatigue for DLQ3 (indirect, M= 0.50 [SD=0.58]) than for DLQ1 and DLQ2 combined (M=0.37 [SD=0.68]). Within the LPD2 conditions, there was almost no change in physical fatigue for DLQ3 (M=0.05 [SD=0.61]) from afternoon over morning, but a small increase for the

Table 9—*Summary of psychophysical performance effects*

Effect	Hit Rate		Error Rate		Time-to-Hit	
					F (1,283)	*R^2*
DLQA x Simple/ Complex Target•					*6.19*	*0.021*
Group					*M (SD)*	
DLQ1 and 2 / Simple					**0.097 (0.037)**	
DLQ1 and 2 / Complex					**0.120 (0.030)**	
DLQ3 / Simple					**0.105 (0.037)**	
DLQ3 / Complex					**0.123 (0.035)**	
					F (1,283)	*R^2*
LPDA x Simple / Complex Target ••					*7.13*	*0.025*
Group					*M (SD)*	
LPD1 and 2 /Simple					**0.096 (0.036)**	
LPD1 and 2 / Complex					**0.119 (0.028)**	
LPD3 / Simple					**0.107 (0.044)**	
LPD3 / Complex					**0.125 (0.037)**	
	F (1,283)	*R^2*	*F (1,283)*	*R^2*	*F (1,283)*	*R^2*
LPDA•••	*6.36*	*0.022*	*5.49*	*0.019*	*8.27*	*0.028*
Group	*M (SD)*		*M (SD)*		*M (SD)*	
LPD1 and 2	**0.74 (0.26)**		**0.35 (0.27)**		**0.104(0.035)**	
LPD3	**0.69 (0.31)**		**0.39 (0.30)**		**0.113(0.043)**	
					F (1,283)	*R^2*
LPDA x DLQB x Time					*5.03*	*0.018*
Group					*M (SD)*	
AM:						
LPD1 and 2/DLQ1					**0.109 (0.038)**	
LPD1 and 2/DLQ2					**0.104 (0.032)**	
LPD3/DLQ1					**0.126 (0.050)**	
LPD3/DLQ2					**0.112 (0.041)**	
PM:						
LPD1 and 2/DLQ1					**0.092 (0.026)**	
LPD1 and 2/DLQ2					**0.097 (0.027)**	
LPD3/DLQ1					**0.116 (0.048)**	
LPD3/DLQ2					**0.096 0.028)**	

Note. Hit Rate is proportion of targets correctly hit. Error rate is proportion of spacebar presses that were errors (false positives and outside the removal zone). Time-to-hit is the number of seconds required to remove a target in the removal zone. •DLQA x Simple/Complex Target: Wilks' Λ =0.967, F(3, 281)=3.15, R^2_{ave}=0.012. ••LPDA x Simple/Complex Target: Wilks' Λ =0.969, F(3, 281)=2.96, R^2_{ave}=0.013. •••LPDA: Wilks' Λ =0.967, F(3, 281)=3.15, R^2_{ave}=0.023. LPDA x DLQB x Time: Wilks' Λ =0.972, F(3, 281)=2.74, R^2_{ave}=0.006. Cells display raw scores. Analyses were performed on transformed scores, using the following transformations: Time-to-Hit=Log_{10}(time-to-hit).

Downloaded by [Northeastern University] at 19:04 25 November 2014

DLQ1 and DLQ2 conditions (M=0.43 [SD=0.76]).

Aesthetic impressions

The 27 semantic differential items were scored continuously from 0 to 100. The most interpretable solution obtained from PCA was an unrotated three-factor solution, presented in **Table 10.** Nine scales, including some of Flynn's[34] originals, did not load highly on any component. The largest component was labeled "Visual Attraction," the second "Complexity," and the third "Brightness." Only one variable ("bright - dim") loaded higher than the 0.500 cutoff on the third component, which in conventional PCA would result in that component not being interpreted. It was retained here because of the intrinsic interest of this concept to lighting research. Likewise, Glare was retained as a separate variable, although it did not load higher than the cutoff on any component.

Analysis proceeded using the 3 by 3 MANOVA design with four dependent measures, these being subscale scores (averages of the scores on the highly loading items, reverse-coded for negatively loading items) for Visual Attraction, Complexity, Brightness, and Glare. Interestingly, there were no statistically significant effects involving the Visual Attraction scale.

Significant multivariate effects for DLQA (Wilks' Λ=0.947, F(4, 280)=3.94, R2(ave)=0.013.), DLQB (Wilks'Λ =0.879, F(4, 280)=9.63, R^2_{ave}=0.029), and LPDB (Wilks' Λ =0.901, F(4, 280)=7.72, R^2_{ave}=0.026) were obtained. Overall, the room, when lighted with direct systems, was judged to be brighter than when indirect systems were in use (F (1,283)=11.05 R^2=0.038, M=64.38 [SD=22.00] for DLQ1 and 2, M=55.33 [SD=23.88] for DLQ3).

The judgments of room appearance under parabolic louvered systems (DLQ2) were less complex (F

Table 10—*Principal components analysis of room aesthetics judgments*

	Unrotated Component Loadings		
	Visual attraction	**Complexity**	**Brightness**
like - dislike	0.837		
pleasant - unpleasant	0.836		
beautiful - ugly	0.825		
attractive - unattractive	0.820		
interesting - monotonous	0.767		
colorful - colorless	0.724		
comfortable - uncomfortable	0.718		
subdued - stimulating	-0.673		
gloomy - radiant	-0.611		
spacious - cramped	0.610		
somber - cheerful	-0.608		
tense - relaxing	-0.540		
distinct - vague	0.505		
cluttered - uncluttered		-0.610	
nonuniform - uniform		-0.589	
constant - flickering		0.569	
complex - simple		-0.515	
bright - dim			0.655
glaring - not glaring			-0.462
warm - cool			
overhead - peripheral			
large - small			
dramatic - diffuse			
faces clear - faces obscure			
public - private			
formal - casual			
clear - hazy			
Percent Variance Explained	28.77	9.27	6.18
Cronbach's α	0.92	0.51	N/A

Note. The instructions for this task were: "We would like you to tell us your opinions of the appearance of this room and of people and objects in it. There are no right or wrong answers. Different people will judge the room in different ways. Please mark the place on the scale close to the adjective that you feel best describes your opinion. The more strongly you feel about the adjective, the closer you should place your mark to it." Responses were scored on a continuous scale from 0-100.

Downloaded by [Northeastern University] at 19:04 25 November 2014

(1,283)=7.27, R^2=0.025, M=29.24 [SD=13.86] for DLQ1, M=24.37 [SD=11.61] for DLQ2), less bright (F (1,283)=13.32, R^2=0.045, M=70.16 [SD=20.88] for DLQ1, M=58.66 [SD=21.70] for DLQ2), and less glaring (F (1,283)=11.80, R^2=0.040, M=57.30 [SD=25.41] for DLQ1, M=45.17 [SD=25.17] for DLQ2) than under lensed systems.

Lighting conditions that included task lighting with lower overall ambient illumination (LPD1) led to room appearance judgments that were less bright and less glaring than under higher-LPD systems without task lighting (LPD2). For Brightness, F (1,283)=6.71, R^2=0.023; M=56.78 (SD=23.68) for LPD1, and M=64.33 (SD=22.64) for LPD2. For Glare, F (1,283)=23.27, R^2=0.076; M=40.40 (SD=24.36) for LPD1, and M=57.46 (SD=26.20) for LPD2.

Downloaded by [Northeastern University] at 19:04 25 November 2014

Manipulation checks

Participants ended the day by answering three open-ended questions concerning the impressions they had gathered about the purpose of the experiment, the experimenters' expectations concerning its outcomes, and factors that might have influenced their responses during the day. Examination of these data revealed that less than one-quarter of respondents believed that lighting was the primary focus of the experiment (66 responses) or felt that the lighting had influenced their work or other responses (71 responses). Participants believed that the experimenters were concerned with identifying working conditions that would maximize productivity or performance, while being comfortable, healthy, and not stressful. Lighting levels were the only specific lighting dimension identified in response to this question, although most responses said only that "lighting" was of interest to the researchers.

Glare was the most frequently identified influence on performance. Seven participants said that the light was too bright; 64 reported glare problems, either alone or with other variables. Glare complaints occurred in all experimental conditions. (Detailed frequency analyses were not conducted because of the overall low frequency of specific complaints.)

Discussion

We had hypothesized that the behavioral outcomes would improve in a linear fashion with Designers' Lighting Quality. We also tested the hypothesis that the behavioral outcomes would be affected by changes in luminous conditions associated with reductions in lighting power density. The results were not as clear as we had hoped, but contain good news for energy conservation and for people who want to improve the quality of the luminous environment.

The good news for energy conservation is this: We found no support for the fear that as LPD declines, so does lighting quality. There was no overall main effect of LPD in which performance, satisfaction, or mood, was worse for the lowest LPD condition. Moreover, a clear pattern of evidence supports the adoption of energy-efficient electronic ballasts because of their effects on people.

The news for people interested in lighting quality is more complex. There was no simple main effect of DLQ favoring the indirect systems; rather, there were interactions of LPD and DLQ, in which the effect of the lighting system depended also on system characteristics associated with energy efficiency. There were also significant main effects for both LPD and DLQ that modify these interaction effects.

The independent variables are valid instances of the constructs Designers' Lighting Quality and Lighting Power Density, but the luminous conditions change in several dimensions at once from one level of each variable to another. For example, between LPD2 and LPD3, the ballast type changes from electronic to magnetic. The differences in luminous conditions from LPD2 to LPD3 include a change in illuminance and in luminances in the field of view, and a change in flicker rate. This confounding proved to be an unavoidable consequence of the decision to use lighting installations that might occur in real offices, rather than unrealistic models built for laboratory use only. Future analyses will probe the photometric data for a useful predictor of the performance and satisfaction outcomes; for the moment, we will deal with differences between lighting system characteristics.

Guide to applications

In keeping with our model of lighting quality based on behavioral outcomes, we interpret these effects in terms of the trend across outcome categories. We have summarized all of these effects in **Figure 7**, which we present as a simplified guide to the application of these findings to lighting decisions. As with all scientific results, of course, the relationships in the figure can be expanded and qualified; this, we do in the text. Moreover, we can only consider lighting conditions that occurred in the experiment.

Ballast type—The figure begins with the issue of ballast choice: magnetic or electronic? First costs and applicable codes and standards will influence this selection (LPDA). Our findings suggest that the best choice is an electronic ballast; therefore, we present that path in bold. The italics beside the path identify the outcomes that favor that particular choice.

The main effect for LPDA occurred on visual performance, verbal-intellectual performance (writing and reading), task difficulty ratings, and psychophysical per-

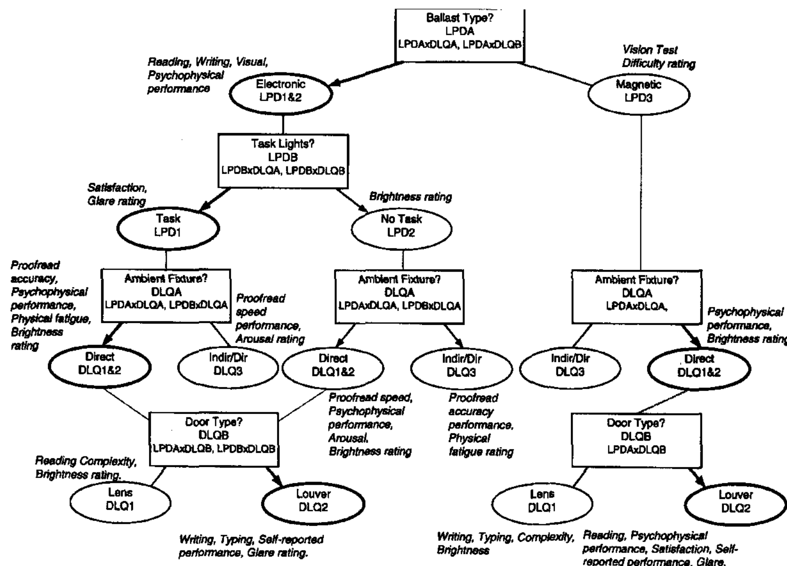

Figure 7—Schematic guide to choices between lighting systems used in this experiment, summarizing experimental results.

formance. **Figure 4** shows the results graphically for selected dependent measures. With one exception (ratings of the difficulty of the vision tests), the results were consistent, with better performance for people working under electronically ballasted systems than under magnetically ballasted systems.

The findings for this effect are also consistent with previous research. Participants wrote more and attempted more reading test questions in LPD1 and 2 conditions than LPD3 conditions. These effects might be explained by the disruption in saccadic eye movements that Wilkins[39] observed under the low-frequency flicker typical of magnetic ballasts in comparison to high-frequency electronic ballasts. This type of effect might also account for the psychophysical performance task results: the hit rate was higher, error rate lower, and time-to-hit (reaction time) shorter for LPD1 and LPD2 than for LPD3. Likewise, participants in LPD1 and LPD2 conditions had better visual performance for low-contrast (0.31) material at the end of the day than those who had worked in LPD3 conditions. The existence and the direction of the effect are consistent with previous research.[40] Overall, the LPDA effects support the use of electronic ballasts over magnetic ballasts; in addition to their energy-saving potential, they appear to be beneficial to people. (There was one paradoxical result: LPD1 and LPD2 participants reported that the vision tests were more difficult than the LPD3 participants.)

Task/ambient combinations—Given the choice to use energy-efficient electronic ballasts, one then faces a choice between a task/ambient combination (in our case, the most energy-efficient systems), and ambient-only (typical of current energy-efficiency strategies). This is the LPDB main effect, although our results are modified by interactions of LPDB by DLQA and LPDB by DLQB. Considering the main effects, the balance favors the task/ambient combination. Small but significant main effects on satisfaction and room appearance ratings were found (**Figure 5**). Overall, the task/ambient combinations improved environmental satisfaction and lighting quality ratings and were rated as lower in glare. LPD1 conditions were also (accurately) rated as less bright than LPD2 conditions overall. Based on these results, we present the task/ambient pathway in bold; however, results discussed in the following section can also justify choosing an ambient-only system.

Direct vs indirect systems—The results for this comparison are ambiguous. There were few overall main effects, suggesting that the behavioral effects of the DLQ3 systems were not different from those of the DLQ1 and DLQ2 systems. Brightness ratings and psychophysical task performance were the only small effects: The room was rated as brighter under DLQ1 and DLQ2 than DLQ3. In addition, for simple targets, although not for complex ones, time-to-hit (reaction time) on the psychophysical task was lower under direct-only than under indirect systems.

However, for the comparison of direct and indirect systems at LPD1 and LPD2, we look also to the significant LPDB by DLQA interaction (**Figure 3**). This shows that best choice varies depending on whether one has chosen the task/ambient or ambient-only route.

For proofreading, the interaction results show a trade-off between speed and accuracy. The faster groups were also less accurate: At LPD1, the slower, more accurate performance was for DLQ1 and DLQ2; at LPD2, the slower, more accurate performance was for DLQ3. Determining which of these groups had the best performance will depend on the relative importance of speed and accuracy. For some applications, speed is of the essence; however, the failure to detect errors is likely to be more costly in the long run. For this reason, we emphasize accuracy over speed in our interpretation. (The proofreading instructions were neutral, favoring neither accuracy nor speed.)

The results for arousal and physical fatigue follow an interesting pattern. For the direct systems (DLQ1 and DLQ2), arousal and physical fatigue did not differ with or without task lighting. The LPD1/DLQ3 group, how-

Downloaded by [Northeastern University] at 19:04 25 November 2014

ever, had higher arousal and a larger increase in physical fatigue than the LPD2/DLQ3 group. Under indirect systems, higher arousal appears to have led to greater physical fatigue.

Considering the LPDB by DLQA results, the pattern is difficult to interpret. Proofreading performance was slower but more accurate for LPD2/DLQ3 than LPD2/DLQ1 and DLQ2; and these participants had the lowest overall fatigue. However, speed and errors for LPD1/DLQ1 and DLQ2 were very similar to the LPD2/DLQ3 scores, making it difficult to conclude that any group had superior performance on this task. This outcome is comparable to other work: Most researchers comparing direct and indirect systems have failed to observe significant effects on any outcomes.[41-43]

Faster but less accurate performance, and more fatigue, occurred in the LPD1/DLQ3 group (indirect furniture-mounted luminaires + task lighting). This group was also more aroused than the others. These findings might relate to those of Collins et al.,[27] who reported that furniture-mounted indirect lighting with task lighting was rated most unsatisfactory among the systems examined in the field. They surmised that the pattern of luminances in the space was at fault, a hypothesis that will be examined further in these data (and reported elsewhere).

Overall, taking into account both the main DLQA effect and LPDB by DLQA interaction, it may be said that either direct or direct/indirect lighting systems can provide good lighting quality at low LPDs. However, this experiment cannot provide conclusive results about the behavioral effects of indirect systems because, as we discovered, there are many different indirect systems; the higher variability between conditions across the DLQ3 level might be one reason for the lack of effects involving this set of lighting systems.

Lensed vs louvered energy-efficient systems—If one has chosen a direct lighting system, the next choice in our scheme (**Figure 7**) is between flat prismatic lenses and parabolic louvers. In this section, we present the interpretation of this interaction for the LPD1 and LPD2 conditions; below, we discuss the LPD3 conditions.

For computer-based tasks (typing probe, writing), the shape of this effect is consistent: at LPD1 and 2, performance is better under parabolic louvered luminaires. For the one paper-based task, the reading test, the effect is reversed: participants attempted more items under prismatic lenses.

Interestingly, satisfaction ratings do not directly follow the performance data. The ratings for DLQ1 and DLQ2 did not differ at LPD1 and LPD2. Overall, it appears that for low-LPD lighting designs that are typical of current and cutting edge installations (LPD1 and LPD2), parabolic louvered systems are judged to be equal in quality to lensed systems, but permit better performance on computer-based tasks. Lensed systems might be preferable for paper-based tasks, but with only one such task, this experiment cannot provide conclusive evidence in this regard.

Main effects over all the data also influence this choice. Significant main effects emerged for performance ratings and room appearance ratings (**Figure 6**). Parabolic louvered systems (DLQ2) were judged as causing better self-reported productivity and lower overall task difficulty than lensed systems. They were judged as creating a less complex and less bright room appearance with lower glare.

Past practice choices—Here, we turn to the lighting quality implications of lighting system choices typical of past practice (LPD3). We note that many of these systems are still in use. The lighting systems in this category include one of the most common early retrofits for computer offices: shallow paracube louvers to replace prismatic lenses in identical luminaires.

There is little evidence relevant to the choice of direct or indirect systems in LPD3. There was no significant LPDA by DLQA main effect to isolate this choice. Only the small significant DLQA main effect applies, which tends to favor direct systems.

Turning to the direct systems, we return to the significant LPDB by DLQA interaction for clerical and verbal-intellectual performance and the satisfaction ratings. Typing probe and writing performance were better for lensed than louvered luminaires. For the one paper-based task, the reading test, the effect is reversed: participants attempted more items under louvered than lensed luminaires. Satisfaction ratings clearly favor the louvered luminaires over the lensed, as do the psychophysical performance data. On balance, the louvered luminaires appear to produce superior outcomes to the flat prismatic lenses, even for the LPD3 level, but over the short term, computer-based work is not necessarily helped by the simple change from lenses to louvers.

Conclusions

Overall, these results suggest that for computer-based office work under present-day energy-efficiency targets, parabolic-louvered systems provide superior lighting quality over lensed systems. Suspended direct/indirect lighting also provided high-quality lighting by certain measures. Interestingly, this direct/indirect system also incorporated parabolic louvers. This finding is, in a sense, a validation of recommended practices such as RP-1.[44] Moreover, the change to electronic ballasts for energy efficiency has beneficial effects on task and visual performance.

At energy consumption levels typical of past practice, the superiority of parabolic louvers over lenses is less

Downloaded by [Northeastern University] at 19:04 25 November 2014

obvious. Retrofits with the particular louvers used in this study are not sufficient to clearly provide good quality lighting. In any case, one would hope that such retrofits for computer office lighting include energy efficiency measures such as reflectors or new luminaires and electronic ballasts, not only a replacement door for an existing luminaire.

As energy efficiency goals tighten, it appears that the combination of task and ambient lighting will be a reasonable strategy to maintain satisfaction while reducing overall energy used for lighting. Moreover, there were no significant performance, mood, or health effects (for good or ill) involving the contrast between very energy-efficient practice and current energy code levels (LPD1 versus LPD2). Energy-efficient lighting need not cause ill effects, provided that the luminaire selection and layout take into account the furniture layout, finishes, tasks, and ages of occupants.

These effects are small to medium (1-7 percent explained variance[45]) in size, but are typical of effect sizes in the lighting literature.[37,46,47] Moreover, the effects cover important behaviors for office workers and their employees. The fact that lighting design had any effect on these complex behaviors and evaluations confirms the importance of efforts to understand lighting quality. The costs and benefits of lighting choices are not only those of equipment, installation, maintenance and energy; they include the effects of the lighting on the people in the space. Fisk and Rosenfeld[48] have estimated that the potential for financial benefits by improving indoor environments are considerably greater than the costs of such improvements. The findings we have reported are but one step towards a better understanding of the luminous conditions required for a good quality indoor environment.

The results also raise serious measurement issues that future research must address. Further refinement is needed for the health measurements, room appearance judgments, and social behaviors in particular. For example, the room appearance judgments did not resolve into a component pattern typical of research in this area,[12] despite a more than adequate subjects:items ratio and a selection of semantic differential scales chosen from the literature. The weakness of these measures made it more difficult to detect significant effects. Having attempted to use only measures that had proved useful in previous research, we are led to the conclusion that on the measurement side, our best is not yet good enough.

These results describe the effects in terms of the nine lighting systems. Further analyses will attempt to identify the dimensions in the luminous environment that explain these differences in performance and satisfaction. Partition luminance, computer screen glare, and average luminance in the field of view will be the first targets of these regression analyses, as we attempt to test relationships reported in prior research.[12,49]

Downloaded by [Northeastern University] at 19:04 25 November 2014

References

1. American Society of Heating, Refrigeration and Air Conditioning Engineers/Illuminating Engineering Society of North America. 1989. *Energy Efficient design of New Buildings Except New Low-Rise Residential Buildings,* (ASHRAE/IES Standard 90.1). Atlanta: ASHRAE.

2. Canadian Codes Centre. 1995. *National Energy Code for Buildings* (Public Review No. 1.0). Ottawa, ON: National Research Council Canada.

3. Harris, L. 1980. Office lighting, comfort, and productivity—How the workers feel. *LD+A* 10(no.7):35-40.

4. Spreckelmeyer, K. 1993. Office relocation and environmental change: A case study. *Environment and Behavior* 25:181-204.

5. Veitch, J.A. and Newsham, G.R. 1998. Determinants of lighting quality I: State of the science. *J of the IES*, this issue.

6. Veitch, J.A. and Newsham, G.R. submitted. Determinants of lighting quality II: Research and recommendations. Manuscript submitted for publication.

7. Rubin, A. 1987. *Office design measurements for productivity—a research overview,* (National Bureau of Standards Report NBSIR 87-3688). Washington, DC: U.S. Department of Commerce.

8. Cook, T.D. and Campbell, D.T. 1979. *Quasi-experimentation: Design and analysis for field settings.* Boston: Houghton Mifflin.

9. Kerlinger, F.N. 1986. *Foundations of Behavioral Research* (3rd ed). New York: Holt, Rinehart, and Winston.

10. Shaw, C.Y., Barakat, S., Newsham, G., Veitch, J.A., and Bradley, J. 1995. NRC Indoor Environment Research Facility. *Proceedings of the 16th Annual AIVC Annual Conference on Implementing the Results of Ventilation Research, Palm Springs, CA. September 19-22, 1995* (Vol. 1):210-219). Coventry, UK: Air Infiltration and Ventilation Centre.

11. Katzev, R. 1992. The impact of energy-efficient office lighting strategies on employee satisfaction and productivity. *Environment and Behavior* 24:759-778.

12. Loe, D.L., Mansfield, K.P., and Rowlands, E. 1994. Appearance of lit environment and its relevance in lighting design: Experimental study. *Lighting Research and Technology,* 26:119-133.

13. Veitch, J.A. and Newsham, G.R. 1996. Experts' quantitative and qualitative assessments of lighting quality. Presented at the IESNA 1996 Annual Conference, Cleaveland, OH.

14. Wilkins, A., Nimmo-Smith, I., Tait, A., et al. 1984. A neurological basis for visual discomfort. *Brain* 107:989-1017.

15. Gifford, R. and Eso, S.J. 1988. *The survey of person-*

al influence in common environments. Victoria, BC: University of Victoria, Department of Psychology.

16. Focused Research Program on Work Environments. 1988. *Environmental Sensitivity Scale.* Irvine, CA: University of California, Irvine, Program in Social Ecology.

17. Gifford, R. 1992. *The person-surroundings scale.* Victoria, BC: University of Victoria, Department of Psychology.

18. Veitch, J.A. and Gifford, R. 1996. Assessing beliefs about lighting effects on health, performance, mood and social behavior. *Environment and Behavior* 28:446-470.

19. Rea, M.S. 1988. *Population data on near field visual acuity for use with the Vision and Lighting Diagnostic (VALiD) Kit* (Report no. CR5544.3). Ottawa, ON: National Research Council of Canada.

20. Karlsen, B., Madden, R., and Gardner, E.F. 1977. *Stanford Diagnostic Reading Test* (Blue level). New York: Harcourt Brace Jovanovich.

21. Murray, H.A., Barrettt, W.G. Homburger, E., et al. 1938. *Explorations in personality.* New York: Oxford University Press.

22. Nelson, T.M., Nilsson, T.H., and Johnson, M. 1984. Interaction of temperature, illuminance and apparent time on sedentary work fatigue. *Ergonomics* 27:89-101.

23. Rea, M.S. 1986. Toward a model of visual performance: Foundations and data. *J of the IES* 15(no. 2):41-57.

24. Sauser, W.I., Arauz, C.G., and Chambers, R.M. 1978. Exploring the relationship between level of office noise and salary recommendations: A preliminary research note. *Journal of Management,* 4:57-63.

25. Kahn, A.S. (Ed.). 1984. *Social psychology.* Dubuque, IA: Wm. C. Brown.

26. Baron, R.A., Rea, M.S., and 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.

27. Collins, B.L., Fisher, W.S., Gillette, G.L., and Marans, R.W. 1990. Second level post-occupancy evaluation analysis. *J of the IES,* 19(No. 2):21-44.

28. Sundstrom, E., Town, J.P., Rice, R.W., Osborn, D.P., and Brill, M. 1994. Office noise, satisfaction, and performance. *Environment and Behavior* 26:195-222

29. Stokols, D. and Scharf, F. 1990. Developing standardized tools for assessing employees' ratings of facility performance. In: G. Davis and F.T. Ventre (Eds.) *Performance of Buildings and Serviceability of Facilities* (STP 1019). Philadelphia: American Society for Testing and Materials; (55-79).

30. Russell, J.A. and Mehrabian, A. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality* 11:273-294.

31. Hedge, A., Erickson, W., and Rubin, G. 1991. VDT use, job stress, job satisfaction and the sick building syndrome in offices. In Y. Queinnec and F. Daniellou (Eds.). *Designing for everyone. Proceedings of the 11 Congress of the International Ergonomics Association, Paris. (Vol. 1:)* 713-715. London: Taylor and Francis.

32. Hedge, A., Erickson, W., and Rubin, G. 1992. Effects of personal and occupational factors on sick building syndrome reports in air conditioned offices. In: Quirk, J.C., Murphy, L.R., and Hurrell, J.J. (Eds). *Stress and Well-being at Work.* 286-298. Washington, DC: American Psychological Association;

34. Flynn, J.E., Hendrick, C., Spencer, T., and Martyniuk, O. 1979. A guide to methodology procedures for measuring subjective impressions in lighting. *J of the IES* 8(No. 2): 95-110.

33. Schiller, G.E., Arens, E.A., Bauman, F.S., Benton, C., Fountain, M., and Doherty, T. 1988. A field study of thermal environment and comfort in office buildings (No. 3164 [RP-462]). *ASHRAE Transactions* 94(No. 2):280-308.

35. Veitch, J. A. and Newsham, G.R. 1997. *Experimental investigations of lighting quality, preferences, and control effects on task performance and energy-efficiency: Experiment 1 primary analyses* (IRC Internal Reprt No. IRC-IR-744). Ottawa, ON: National Research Council of Canada, Institute for Research in Construction.

36. Reference Software International. 1990. *Grammatik IV User's Guide* (2nd ed). San Francisco, CA: RSI.

37. Sanders, P.A. and Bernecker, C.A. 1990. Uniform veiling luminance and display polarity affect VDU user performance. *J of the IES* 19(No. 2):113-123.

38. Veitch, J.A., Gifford, R., and Hine, D.W. 1991. Demand characteristics and full spectrum lighting effects on performance and mood. *Journal of Environmental Psychology,* 11:87-95.

39. Wilkins, A.J. 1986. Intermittent illumination from visual display units and fluorescent lighting affects movements of the eyes across text. *Human Factors* 28:75-81.

40. Veitch, J.A. and McColl, S.L. 1995. On the modulation of fluorescent light: Flicker rate and spectral distribution effects on visual performance and visual comfort. *Lighting Research and Technology* 27:243-256.

41. Bennett, C.A. 1986. Lighting, comfort, and performance in the workplace. *LD+A* 16(No. 8):40-44.

42. Harvey, L.O., DiLaura, D.L., and Mistrick, R.G. 1984. Quantifying reactions of visual display operators to indirect lighting. *J of the IES,* 14(No. 1):515-546.

43. Yearout, R. and Konz, S. 1989. Visual display unit workstation lighting. *International Journal of Industrial Ergonomics* 3:265-273.

44. Illuminating Engineering Society of North America (IESNA). 1993. *American National Standard Practice for Office Lighting.* ANSI/IESNA RP-1 1993. New

Downloaded by [Northeastern University] at 19:04 25 November 2014

York: IESNA

45. Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

46. Tiller, D.K. and Veitch, J.A. 1995. Perceived room brightness: Pilot study on the effect of luminance distribution. *Lighting Research and Technology* 27:93-101

47. Veitch, J.A. and Gifford, R. 1996. Choice, perceived control, and performance decrements in the physical environment. *Journal of Environmental Psychology* 16:269-276.

48. Fisk, W.J. and Rosenfeld, A.H. 1997. Estimates of improved productivity and health from better indoor environment. *Indoor Air* 7 (3):158–172.

49. Miller, N., McKay, H., and Boyce, P.R. 1995. An approach to a measurement of lighting quality. Proceedings of the 1995 IESNA Annual Conference New York, NY.

50. Veitch, J.A., Miller, N., McKay, H., and Jones, C.C. 1996. Lighting system effects on judged lighting quality and facial appearance. Presented at the 1996 IESNA Annual Conference, Cleveland, OH.

51. Newsham, G.R. and Veitch, J.A. 1997. Energy-efficient lighting options: Predicted savings and occupant impressions of lighting quality. Presented at CLIMA 2000 Conference, Brussels, Belgium.

Downloaded by [Northeastern University] at 19:04 25 November 2014

Acknowledgments

The preparation of this paper was supported by the Canadian Electrical Association (Agreement No. 9433 U 1059), Natural Resources Canada, the Panel on Energy Research and Development, and the National Research Council of Canada (NRC), as part of the NRC project "Experimental Investigations of Lighting Quality, Preferences, and Control Effects on Task Performance and Energy Efficiency." Lighting equipment used in this experiment was donated by CANLYTE Inc., General Electric Co., Ledalite Architectural Products Inc., Litecontrol Corp., Luxo Lamp Ltd., Osram-Sylvania Inc., Peerless Lighting Ltd., and Philips Lighting. We thank Dahlia Barakat, Sherif Barakat, Cathy Brown, Peter Boyce, Martin Conboy, Philip Gabriel, Ralston Jaekel, Steffan Jones, Stuart Kaye, Pierre Landry, Michael Ouellette, Roger Marchand, Robert Magee, Terry McGowan, Michael Perry, Francis Rubinstein, Carol Scovil, Karen Shaver, Anthony Slater, Jana Svec, Dale Tiller, Martyn Timmings, Jacqueline Vischer, Greg Ward, Jean Wineman, Jim Wise, Arnold Wilkins, and Chuck Woods for their many and various contributions to this project. Other data collected in connection with this experiment is reported elsewhere.[51]

Discussions

This paper has numerous merits. It describes a major study done over a full working day, by a large number of individuals, under a number of different lighting installations, classified according to power density and perceived quality. The experimental design is carefully thought out and the statistical analysis is rigorous. In these respects it sets a high standard for future researchers to match.

Unfortunately, it also sets a high standard in complexity, partly because of the amount of data reported and partly because of the decision to present only the statistically significant effects, without any overall view of the data. However, even if all the data had been presented, I believe the pattern of results would still be complex because complexity is inherent in the choice of independent variables. There is no reason why either lighting power density or lighting quality, as such, should have any effect on the performance of visual tasks or visual comfort. They may or may not, depending on the stimulus they provide to the visual system. Until the ways in which the various lighting installations change the stimuli to the visual system are understood, the interpretation of the data collected will remain difficult and the ability to generalize the results uncertain. As evidence for the value of having an understanding as to why an effect occurs, I would simply point out that the effect which explains the most variance in any of the tasks is the choice of screen polarity. For typing speed and proofreading speed, screen polarity explains 7 and 9 percent of the variance respectively, with the faster speeds being achieved with the positive polarity screen. This effect can be understood from the fact that the relatively higher luminance of positive polarity screens makes the display less subject to interference from reflected images of the rest of the room. I am glad to learn that the authors are attempting to relate the results obtained to the photometric conditions. If they succeed, it will be much easier to understand the effects found and to know the extent to which these results can be generalized to other lighting conditions.

As for the current conclusions, I have no quarrel with the argument that low lighting power densities can provide quality lighting, as defined by outcomes such as task performance, worker satisfaction, or worker mood. It is consistent with previous work.[a] But low power density lighting is not a guarantee of desirable outcomes. Low or high power density lighting installations can be done well or badly. Similarly, the advocacy of the use of electronic ballasts seems reasonable and is consistent with previous work,[b] although the possibility that the magnetic ballasts had an adverse effect because of an audible hum cannot be ruled out. However, I am more doubtful of the value of the guide to applications. It would be a brave specifier who determined the choice of equipment

for a lighting installation based on a flow chart in which the effect determining the choice explained about 2 percent of the variance, leaving the other 98 percent unexplained, and which showed different effects on different tasks for different types of lighting equipment, when the same lighting installation might have to be used for many different types of task.

Overall, this paper is a distinct step forward in attempting to understand the influence of lighting on peoples' everyday working lives. The fact that it raises more questions than it answers is to be expected at this stage of the study of lighting quality.

References

a. Boyce, P.R. and Eklund, N.H. 1996. An evaluation of office lighting options. Proceedings of the 1996 *IESNA Annual Conference* 38-68.

b. Wilkins, A.J., Nimmo-Smith, I., Slater, A.I., and Bedocs L. 1989. *Fluorescent lighting, headaches and eyestrain. Lighting Research and Technology* 21: 11-18.

P. Boyce
Lighting Research Center,
Rensselaer Polytechnic Institute

As we have been hearing now for a number of years, accurate data on the performance effects of lighting quality will be difficult to obtain, hard to analyze, complicated to report and limited in what can be concluded. This is a "dense" paper filled with such information. It also represents another step in the welcome trend to analyze the human response to lighting systems using human factor analysis techniques. The appearance of the results, at least to this lighting person, are unfamiliar but when carefully and thoroughly read, provide important information. I welcome the change.

What has been reported here are the results of a relatively large-scale and extensive experiment designed to see if there is a link between lighting quality and lighting power density. The concern about the relationship (if any) and the impact on lighting design, should recommended or legislated lighting power densities be further reduced has resulted in heated discussions, numerous anecdotes, studies and reports; but, unfortunately, few conclusions. This paper, however, follows all of the rules. It provides extensive data, carefully documented and expertly analyzed to reach the conclusion that for these lighting situations, lower lighting quality as measured by the response of people to the lighted environment does not necessarily result when lighting power densities are decreased.

That is welcome and encouraging news. But it is also sobering to realize that the situations examined in this study were limited. My question to the authors is, therefore, how should your results be used? How far can what you have learned be extrapolated to the general case of enclosed spaces and visual tasks? And, what can be said about the results in the case of long-term working environments, not one-day studies, which are the norm for most people involved in the tasks which you have described?

T. McGowan
GE Lighting

This paper describes a comprehensive investigation into the lighting of offices. It combines human performance and subjective assessment together with physical light measurement for a range of conditions which embrace lighting quality and energy efficiency. It is not common, in these times of reduced research activity, for studies of this magnitude to be carried out and I compliment the research team and the institute for which they work.

My first consideration is the relevance of the experiment. Many of us who have undertaken projects of this nature have asked ourselves, or been asked by others, whether the tasks undertaken and the performance measurements used, represent real life situations and reliable measures of performance. Of course I understand the difficulty and in no way do I criticize, but nonetheless it is important that the criteria that develop from these studies be truly meaningful. In the past many of the performance experiments have provided good working standards but they do not, in my view, represent minimum standards, and hence I suspect we have recommended illuminances and glare criteria which are higher or more stringent that is absolutely necessary.

The same applies to this study, particularly as the office worker today is a very different person to the office worker of a few years ago. Yes they do read, write, type, use a computer and interface with colleagues, but it is a very much more multi-disciplinary job than earlier times and the lighting must respond to this change. Therefore is a one day period of assessment sufficient? It is because of this that I feel that the subjective assessment results are likely to be more useful, but not by naive observers who often have little understanding of visual quality and lighting, but by a team of trained observers. Authors comments please.

The modern office relies heavily on the face-to-face communication between people, yet this study seems to have omitted this consideration. Is it possible that an installation with some indirect light would provide softer modeling of the occupant's face and would be preferred over the direct light only as provided by the parabolic louver installations (DLQ2/LPD2 and LPD3). My experience is that people seem to prefer installations that provide both direct and indirect light and which have areas of light and shade together with room surface bright-

Downloaded by [Northeastern University] at 19:04 25 November 2014

ness. This study seems to support this—do the authors agree?

If this is the case, then office lighting might change dramatically in the future and relate to, or be part of, the workstation. This study has shown that this can be achieved with very low energy consumption (7.96 W/m^2). This would allow the workstation to be moved easily if required and if fitted with "user-friendly" controls the lighting could be altered by the occupant or for energy efficiency. Some of the room surfaces would also require illumination as was shown by the DLQ3/LPD1 installation. However to prove this approach to lighting design, I suggest, we need a trial installation that could be monitored over a reasonable length of time (one year at least) to assess user reactions and performance together with energy consumption. If it was successful it could be replicated. Do the authors agree that this should be the style of the next generation of lighting design investigations rather than the laboratory experiments that many of us here have conducted, and which have not changed lighting design very much—if at all.

Finally the photometric results presented are limited. It would be very useful to have detailed illuminance measurements across the task area and detailed luminance measurements for the task area and the surround field. It would also be useful to have values of the average luminance of the main surfaces within the field of view. My experience is that spot luminance values are of limited value.

D. Loe
Lighting Consultant and
Building Research Establishment Conservation Support Unit (BRECSU)

Authors' response

We thank our colleagues for their thoughtful comments on this paper and, not less importantly, for their helpful comments and contributions throughout this project. Our communications with them and with other lighting researchers and designers at previous conferences and in committee work has contributed greatly to the development of our ideas and to this experiment in particular.

To D. Loe, P. Boyce, and T. McGowan

David Loe requested additional photometric information, which we are pleased to provide in **Table A.**

Peter Boyce speculated that differences in ballast noise, rather than in flicker rate, might have confounded the ballast type effect (LPDA). However, any differences in ballast noise were masked by the addition of simulated ventilation noise. Noise levels in the space were kept at 48 dB(A) for all experimental conditions.Our schematic guide (**Figure 7**) is in no way intended as a scheme for lighting recommendations. It is a means to organize our complex set of results, in terms that might be useful to those for whom the arcane language of behavioral research is unfamiliar.

Moving to the more substantive comments, we wish to reply to all three discussants together. In various ways, each has queried the relevance of this experiment, and indeed experimental work in general for lighting practice.

Relevance, or external validity, is an important issue for all applied psychology research, and particularly so for attempts to study the effects of physical settings on behavior. We have attempted to create a setting that is

Table A—*Luminance values from digital image analysis (CapCalc).*

DLQ		LPD 1 (low)	2 (medium)	3 (high)	
1 (low)	*Scene mean*	**51**	**68**	**74**	
Recessed troffer	*40-deg mean*	**42**	**51**	**54**	
with K12	*40-deg SD*	**36**	**28**	**30**	
prismatic lens	*VDT Max/*	**69**	**110**	**137**	
	Ceiling max		**1371**	**2130**	**2170**
2 (medium)	*Scene mean*	**40**	**37**	**38**	
Recessed troffer	*40-deg mean*	**48**	**40**	**43**	
with parabolic	*40-deg SD*	**62**	**25**	**28**	
louver	*VDT Max/*	**31**	**61**	**53**	
	Ceiling max		**462**	**499**	**507**
3 (high)	*Scene mean*	**55**	**82**	**146**	
Indirect or	*40-deg mean*	**42**	**58**	**65**	
direct/indirect	*40-deg SD*	**39**	**34**	**34**	
	VDT Max	**29**	**167**	**78**	
	Ceiling max	**601**	**627**	**1604**	

Note. All values in cd/m^2. The image was taken to approximate view of seated occupant facing the VDT; this resulted in the camera standing on the tripod at head height, in the doorway of the workstation. The 40-degree band was +/- 20-deg around the horizontal axis, and subtended 45 degrees from side to side. VDT luminance values were taken from the image of a black screen and are a measure of screen glare. Ceiling values include luminaire luminance.

Downloaded by [Northeastern University] at 19:04 25 November 2014

typical of many open-plan offices. The interior design and furnishings are representative of existing Government of Canada installations; the lighting installations may be found in contemporary North American offices. The participants were drawn from the pool of temporary employees from which other local offices draw staff and were paid in accordance with the work performed.

However, we did conduct a controlled experiment in a laboratory. We have to agree that this places some limits on the results of this or any similar study. The participants had special rights, of which they were informed, and were always aware of the unusual nature of the work. Most particularly, the tasks dissected the components of office work in order to permit more accurate measurement. The components are reasonable ones familiar to anyone who does office work, but no one knows the relative importance of each type of performance in the life of an office worker. Such studies do not exist.

In this study we were not able to incorporate direct measures of social interaction. Whether or not it is true that face-to-face communication is more important than formerly thought (we are not aware of any comparative historical data on this point), this is undoubtedly a feature in most offices. We have no data to say what type of lighting would best support face-to-face communication, but of course we would like to see research that would address the question.

The one-day period is longer than most other investigators have been able to use, but still is a comparatively short time in which to form a firm opinion. Short-term laboratory studies of this type should of course be supplemented and extended using longer-term exposures. The strength of laboratory investigation is that it allows us to carefully test specific hypotheses so that we can eliminate some possibilities before moving to field studies. Although laboratory work is expensive, it is neither as costly nor as difficult as a well designed field study.

We do not recommend using expert observers only. There is ample evidence in psychology that experts and laypeople do not hold the same opinions about risks, benefits, or preferences. To know what our end-users think of lighting, we should ask them. For studies of a longer-term nature, this will probably require field studies or field experiments, possibly similar to that described by David Loe. The risk in using expert observers is that their special knowledge about lighting will lead to results that merely confirm what they already believe to be true.

Our results concerning the directionality of illumination were disappointing. We had hoped to see a clear preference for lighting systems that incorporated an indirect component. However, the three lighting systems selected for the DLQ3 conditions used very diverse distributions, more so than was true for the recessed troffer designs in DLQ1 and DLQ2. This experimental design problem makes it difficult to draw conclusions concerning indirect lighting; however, we can say that the performance, satisfaction, and other effects for those systems were not different from those for the direct-only systems. Moreover, interreflections from room surfaces would add some indirect illumination even to the designs using direct-only systems. A combination of direct and indirect illumination might indeed be desirable, but at what proportions? To date, research cannot answer that question.

In conclusion, we would like to add that this experiment, and indeed all lighting research, cannot itself change lighting practice. The problem is not chiefly the limited nature of each individual investigation or report. The problem lies at the interface between research and practice—and is by no means unique to lighting research! Lighting quality improvements will not occur until better-quality lighting is sold to clients as being a worthwhile investment. In addition to updating lighting education materials, textbooks, the Handbook, and recommended practice documents for our own circle, this especially will require disseminating this knowledge beyond the lighting industry to architects, interior designers, building engineers, facility managers and building owners. This is a job for the IESNA and IALD, and for every individual member of those organizations.

Downloaded by [Northeastern University] at 19:04 25 November 2014