ELSEVIER

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs





Cognitive abilities predict performance in everyday computer tasks

Erik Lintunen ^a, Viljami Salmela ^b, Petri Jarre ^b, Tuukka Heikkinen ^b, Markku Kilpeläinen ^b, Markus Jokela ^b, Antti Oulasvirta ^{a,*}

- ^a Aalto University, Finland
- ^b University of Helsinki, Finland

ARTICLE INFO

Dataset link: https://osf.io/7b6uv

Keywords: Cognitive abilities Task performance Computer use

ABSTRACT

Fluency with computer applications has assumed a crucial role in work-related and other day-to-day activities. While prior experience is known to predict performance in tasks involving computers, the effects of more stable factors like cognitive abilities remain unclear. Here, we report findings from a controlled study (N=88) covering a wide spectrum of commonplace applications, from spreadsheets to video conferencing. Our main result is that cognitive abilities exert a significant, independent, and broad-based effect on computer users' performance. In particular, users with high working memory, executive control, and perceptual reasoning ability complete tasks more quickly and with greater success while experiencing lower mental load. Remarkably, these effects are similar to or even larger in magnitude than the effects of prior experience in using computers and in completing tasks similar to those encountered in our study. However, the effects are varying and application-specific. We discuss the role that user interface design bears on decreasing ability-related differences, alongside benefits this could yield for functioning in society.

1. Introduction

This paper sheds new light on a question that is critical for the information society: does cognitive ability predict one's success in using computers, or do differences among individuals merely reflect prior experience of computer use or differences in sociodemographic factors such as age? The answer carries profound implications. Computer applications have become necessary for work, leisure, and even social relations today. Fluency in using them has been associated, among other factors, with earnings (Falck et al., 2021) and well-being (Nguyen et al., 2021). Furthermore, the inability to use computers has been shown to cause frustration (Hertzum and Hornbæk, 2023; Lazar et al., 2006). At a societal level, individual differences may drive a wedge between people who are able to fully engage in a technologically mediated society and those who cannot (Hargittai et al., 2019). It is therefore unsurprising that governments (e.g., UK, EU, and USA) have adopted the development of digital skills and equity on their agendas. However, if cognitive abilities do play a role, it is not enough to train people to use computers, or provide access; computer interfaces need to be redesigned to better match cognitive abilities (Wobbrock et al.,

Higher intelligence has been shown to predict a wide array of positive life outcomes, such as better job performance (Schmidt and Hunter,

1998), better mental and somatic health (Deary et al., 2010), and even higher creative achievement (Kim, 2008). However, while large individual-level differences have been reported in user performance across computer applications (Boot et al., 2015; Ownby et al., 2008; Quiroga et al., 2015; Sharit and Czaja, 1999; Wagner et al., 2010; Westerman et al., 1995), research still has no answer as to whether they are attributable to relatively stable cognitive abilities. Research in gaming (Bediou et al., 2018; Quiroga et al., 2015) offers some evidence of the role of cognitive abilities. In complex games such as *League of Legends*, cognitive abilities predict a gamer's ranking (Kokkinakis et al., 2017). However, such games are *designed* to be challenging, whereas standard computer applications are designed with usability in mind. User interface designers aim to reduce cognitive load, improve ease of use, and guarantee access to all (Johnson, 2020). Therefore, results from the gaming domain might not transfer to general computer use.

Research has given little attention to the relationship between cognitive ability and everyday computer use, despite its crucial role in contemporary life. While a questionnaire-based study found a link between self-reported success in computer use and cognitive abilities (Tun and Lachman, 2010), controlled studies in which the participants actually carry out tasks with a computer are the gold standard. Direct

E-mail addresses: erik.lintunen@aalto.fi (E. Lintunen), viljami.salmela@helsinki.fi (V. Salmela), petri.jarre@helsinki.fi (P. Jarre), tuukka.heikkinen@helsinki.fi (T. Heikkinen), markku.kilpelainen@helsinki.fi (M. Kilpeläinen), markus.jokela@helsinki.fi (M. Jokela), antti.oulasvirta@aalto.fi (A. Oulasvirta).

https://doi.org/10.1016/j.ijhcs.2024.103354

Received 6 February 2024; Received in revised form 29 July 2024; Accepted 9 August 2024 Available online 16 August 2024

^{*} Corresponding author.

measurement of performance is less reliant on self-perception and emotional factors and may better predict real-world success. However, only a few such studies exist, all of which have been narrow and focused on either special applications or specific demographic groups. In a study carried out in the 1990s, younger age and higher visuo-motor abilities were correlated with better performance in a banking task (Sharit and Czaja, 1999). Also, research has identified a link between higher cognitive abilities and better performance in information-seeking tasks (Pak et al., 2006; Sharit et al., 2008; Westerman et al., 1995), and higher fluid intelligence has been shown to be associated with faster learning in an e-mail task among aging adults (Nair et al., 2007). Computer use is like other skills in that it depends on - and improves with practice and familiarity (Nair et al., 2007; Choi et al., 2021; Mitzner et al., 2019). For gauging the relationship between cognitive abilities and actual success in using computers, substantive effort is needed in collecting comprehensive measurements. An ideal study should probe both measures in controlled conditions. For generalizability, an ideal study should cover a wide range of everyday tasks, each of which may draw on particular components of cognitive ability (Kovacs and Conway, 2016), as studies of video games attest (Dobrowolski et al., 2015; Quiroga et al., 2019). Moreover, data on covariates like experience, fatigue, and demographic factors should be collected.

Here, we study which cognitive factors predict task success, task completion time, and mental load, across a wide spectrum of popular computer applications. In our preregistered study, we employed a set of 18 common tasks involving computers. The set was designed to represent not only personal computing tasks but the demands of functioning in digital society (OECD, 2012; Wilson et al., 2015). It represents the most comprehensive cross-section of work and leisure domains investigated thus far, including applications for word processing, banking, information search, maps, e-mail, and other operations. Examples are given in Fig. 1. The difficulty levels of the tasks were pre-calibrated in a pilot study targeting average success rates between 25% and 75%. Our prespecified, stratified sample (N=88) covers wide age range and key demographic factors.

We hypothesized that general cognitive ability predicts performance in computer-use tasks, even when prior experience and sociodemographic variables are accounted for. At its most fundamental level, cognitive ability can be measured in terms of the general factor of intelligence (Spearman, 1904), but it is critical to understand the contribution of more specific cognitive domains (Sachdev et al., 2014). Specifically, executive functions may have an effect independent from that of general intelligence (Friedman et al., 2006; Draheim et al., 2021). These are needed for controlling and inhibiting actions, and for shifting attention in a task-relevant way. Clues about the effect of executive functioning can be derived from eye-tracking data, which can show how an individual explores and homes in on the information at hand (Liversedge and Findlay, 2000). Secondly, earlier computational modeling suggests that challenging engagement with computers demands working memory (Card et al., 1980; Kieras and Polson, 1985; Schraagen et al., 2000). The user must bear in mind intermediate results and goals, such as what has been typed or what the next subgoal is. Thirdly, perceptual reasoning and linguistic abilities are important in graphical user interfaces that rely on visuo-spatial presentation and label-based information, respectively. Since all the factors depend on the task at hand, we expected to see task-specific effects of these different cognitive domains (Kovacs and Conway, 2016). In addition, we captured both task-specific levels of familiarity and participants' self-reported efficacy in computer use. This rich dataset allows us to quantify the contributions both of general cognitive ability and of its constitutive components across multiple application types, all while controlling for demographic factors and various aspects of prior experience, including familiarity with user interfaces, applications, operating systems, and computers in general, as well as the beliefs that one has about their capabilities to successfully use computers.

2. Methods

The two main goals behind our study design were (i) designing a set of realistic but challenging everyday tasks performed on a computer and (ii) obtaining a sample that is large, representative, and diversity-rich — especially with regard to age, expressed gender, cognitive ability, and computer skills.

2.1. Ethics

The study protocol adhered to the ethics principles of the Declaration of Helsinki (WMA, 2013) and the guidelines of the Finnish National Board on Research Integrity, TENK (Kohonen et al., 2019). Informed consent was obtained in writing from each participant.

2.2. Preregistration

Prior to collection of the data, a preregistration document was published in an Open Science Foundation (OSF) project repository². Only one aspect of the work deviated from the prescribed protocol: the final sample size was slightly smaller than planned due to difficulties in recruiting older males.

2.3. Participants

With our sampling process, we aimed at stratification based on expressed *Gender* and *Age*: 58% females (N=51) and 42% males (N=37), with *Age* distributed uniformly across three ranges: [20, 35), [35, 50), and [50, 65] (see Appendix D). Among the criteria for recruitment were being a legally competent adult aged 20–65 with normal or corrected-to-normal vision, proficiency in the Finnish language, having no diagnosed specific learning difficulties, and being in an employment relationship at the time of testing. The final sample comprised 88 participants, with an even spread across all expressed genders and age bands (N=88, with $\chi^2=2.54$, p=.281). Data collection was conducted between September 2022 and February 2023. The preregistration materials (link above) contain sample-size estimates and power analyses.

2.4. Tasks and materials

We aimed to create a diverse and representative set of tasks, factoring in task difficulty and work-load variance. The set of tasks was designed iteratively to cover everyday computer tasks needed to function in the society (Wilson et al., 2015). Appendices E and F present a description of each task along with a gallery of screenshots.

Each task comprised several subtasks, divided into the categories *core* (critical for completing the task) and *additional* (success-related but not crucial). This allowed us to capture more variance among high-performing participants. Each task was to be completed within a three-minute limit. The limit was imposed to be able to test the participants on a wide range of tasks while keeping the session length reasonable. The level of challenge was calibrated through a pilot study, the details of which are reported in our preregistration document (?). The goal of the calibration was to avoid floor and ceiling effects while keeping the tasks realistic. Appendices G and H show individual differences and task differences, respectively. From these figures, we see that the calibration was successful in ensuring the tasks were realistic given the time limit

The experiment used the Windows 10 operating system, with the exception of the command-line processing task which utilized Windows Subsystem for Linux running Ubuntu. Before each trial, the operating system was restored to a pre-trial snapshot using the "System Restore" feature of Windows. A keyboard and mouse were supplied as input devices.

¹ See https://osf.io/7b6uv for details.

² https://osf.io/7b6uv

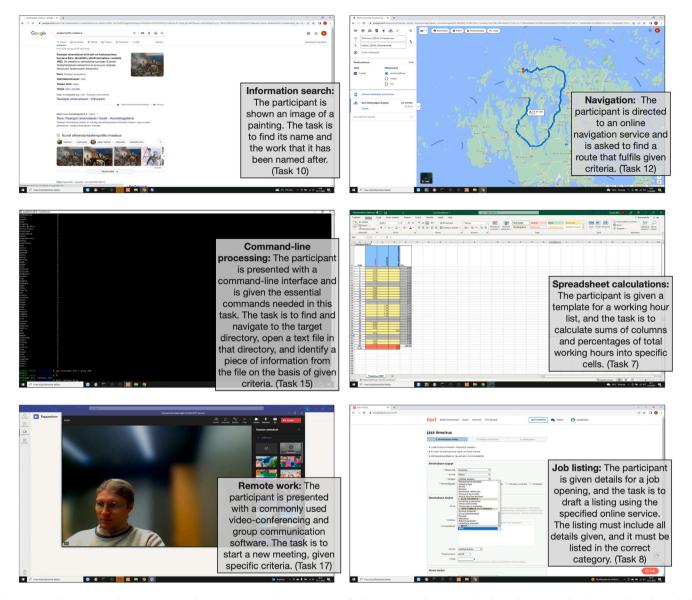


Fig. 1. Our tasks cover everyday activities carried out on computers that are important for functioning in the society. This figure shows a sample of six tasks from the total of 18 in the study.

2.5. Experimental design

The experiment followed a within-subjects design, in which participants were tested one at a time and in two sessions each. During the first (Session 1), the participant completed computerized tasks and provided responses to questionnaire items. Task order was randomized for each participant. In the second session (Session 2), we carried out the WAIS-IV cognitive ability assessment (see below).

2.6. Procedure

At the beginning of Session 1, the participant was given a concise verbal overview of the procedure, to guarantee them a clear understanding of the study, its objectives, and the tasks involved. Informed consent was then requested, after which the participant was given the *CUSE* questionnaire to complete (described further down). Next, the participant was guided through a written list of instructions to ensure standardized task administration. The participant was directed to begin each task as soon as the instructions appeared on the screen, which is

when the timer was started. The instructions remained visible throughout the task. No timer was displayed; instead, the participant received a simple audible alert when 30 s of task time remained. Were the time to expire mid-task, the participant would be interrupted and asked to stop. After calibrating the eye-tracker and completing a practice task for familiarization with the conditions, the participant verbally confirmed readiness to begin. At this point, the first half of tasks was presented, with task-specific questionnaires administered after each via a browser-based interface. The second half of tasks followed after a two-minute break. Once the 18 computerized tasks were completed, the participant was asked to respond to the background questionnaire. Session 1 ended with two computerized cognitive tests (Antisaccade and Selective Visual Arrays).

Session 2 was held on a separate occasion to minimize the potential effects of fatigue. The timing, order, and other administrative aspects of the WAIS subtests were handled in the manner specified in the WAIS-IV manual (Wechsler, 2008a). There were no scheduled breaks, but the participant was allowed a short break between the tasks on request.

2.7. Experimental setup

Both sessions were completed in a soundproofed and light-controlled testing room. In Session 1, there was one computer monitor for task instructions (on the left) and one monitor for completing the computerized tasks (at the center of the work space). An eye-tracker was positioned below the latter screen. The instructor was stationed unobtrusively outside the participant's field of view, monitoring the test screen via a remote connection. In Session 2, the participant and test administrator sat face-to-face, across a table from each other.

2.7.1. Performance-related metrics

We measured the following task-specific variables: *Task Success, Elapsed Time, Mental Load*, and *Familiarity*. The first of these captures the proportion of the work successfully completed, and the time variable refers to the time taken to complete the core subtasks of each task. Task outcomes and questionnaire responses were collected using Python SQLite (version 3.33.0).

2.7.2. Measurement of cognitive abilities

We estimated general cognitive ability (Full-Scale IQ) via the Wechsler Adult Intelligence Scale (WAIS), 4th edition (Wechsler, 2008a). As part of the core battery, Verbal Comprehension is tested with subtests Similarities, Vocabulary, Information; Perceptual Reasoning with Block Design, Matrix Reasoning, Visual Puzzles; Working Memory with Digit Span, Arithmetic; and Processing Speed with Symbol Search, Coding. The assessment was administered in person by a master's student in psychology, under the supervision of a licensed psychologist. While most of the subtests were administered verbally, some involved the use of pictures, pencil markings, and manual manipulation of objects.

To assess executive functions, we employed the accuracy-based Antisaccade instrument for capturing the *Response Inhibition* variable and Selective Visual Arrays for *Attention Control*, on the basis of best practices (Draheim et al., 2021). In the former, participants are tasked to inhibit distractors by making saccades in the opposite direction (left or right) of cued stimuli and identify target stimuli before they are masked; and in the latter, participants must attend to either blue or red arrays and provide a response indicating whether any change occurred between two consecutive array configurations (first cue shown briefly). These tests were completed using E-Prime (version 3.0).

2.7.3. Measurement of eye movements

Eye movements were measured using the GP3 HD Eye Tracker (150 Hz, accuracy 0.5-1°) and Gazepoint Analysis (version 6.8.0). From the gaze fixations, we calculated the following statistics: Fixation Count, Fixation Duration, Off-Screen Fixations, and Explorative Behavior. These were calculated for all tasks and participants from fixations occurring during the first 60 s of a given task. Fixation Count represents the absolute number of fixations, Fixation Duration their mean duration, Off-Screen Fixations the percentage of them outside of the screen area, and Explorative Behavior the dispersion of within-screen fixations. After applying a convolution on each fixation point using a two-dimensional Gaussian kernel - modeling 1° of the gaze - dispersion was computed as the count of nonzero pixels in the fixation heatmap divided by the total number of pixels in the heatmap. Our measure of dispersion is consistent with standard measures of fixation spread from the eye tracking literature, e.g., "spatial density", (see Goldberg and Kotval, 1999; Moacdieh and Sarter, 2015) where the display area is divided into a grid and the number of cells with at least one fixation is divided by the total number of cells. However, our method of computing the proportion after the application of convolutions, and using pixel-sized cells, allowed us to avoid extreme effects of discretization and achieve more "smoothness" in the measurements.

2.7.4. Survey instruments

Questionnaires were administered either repeatedly, after each task, or only once. The two task-specific variables assessed were (i) subjective mental work load (Mental Load), measured by means of the NASA Task Load Index (Hart and Staveland, 1988), pared down to exclude physical demand, and (ii) task familiarity (Familiarity), measured via a linear combination of self-evaluated interface and task familiarity, each rated on a scale of 0 to 100. Task familiarity was measured to account for task-specific prior knowledge, and the participants were instructed to include analogues tasks (e.g., an online banking system with a different bank) in their subjective estimates. The written question read: "How often do you perform a comparable task?" Interface familiarity on the other hand was measured to account for prior experience in using exactly those systems encountered during the tasks. The written questions read: "How familiar were you with the software in Finnish/English [one question for each language]?" Familiarity with the operating system was measured separately from familiarity with tasks and interfaces as part of the background questionnaire (described next).

We measured self-efficacy (CUSE) via the Computer User Self-efficacy scale (Cassidy and Eachus, 2002); collected demographic details such as Age, Gender, and Education; and gathered background information on prior computer use (Exposure), such as the average number of hours' use per week, as well as data related to mobile-device and operating system use, collected to control for prior experience both in using task-relevant operating systems and in using computing systems more generally. We measured baseline alertness using the Toronto Hospital Alertness Test (THAT) (Shapiro et al., 2006), but the variable was dropped from our analyses after confirming that it was a non-significant predictor of the dependent variables of interest.

2.8. Data analysis

Before carrying out statistical tests of the effect of cognitive abilities, we had to ascertain also whether the sample exhibits sufficient variance for further analyses. Large individual-level differences were indeed evident for task performance and cognitive abilities. Mean values for *Task Success* fell within the range 16.5–91.4% (mean: 64.5, sd: 16.1), and the range for mean *Elapsed Time* was 53–176 s (mean: 108.9, sd 24.7). See Appendix G for details. Values for *Full-Scale IQ* varied from 78 to 140 (mean: 109.5, sd: 13.3). Statistical tests, described below, delved into the latter variable's positive correlation with Task Success and its negative correlation with *Elapsed Time* and with *Mental Load*. They revealed that the 18 tasks varied in their levels of difficulty (see Appendix H).

Several statistical analyses were made. First, partial correlations (controlling for age, gender and education) were calculated between continuous predictor and outcome variables averaged across tasks (Fig. 2). Second, mean data (over computer tasks) were analyzed with hierarchical linear regression analysis containing four blocks/models (Fig. 3). Third, hierarchical linear regression analyses were repeated separately for each task. Demographic factors were in the first block and either cognitive variables or experience related variables in the second block (Fig. 4). Fourth, linear regression analyses were repeated separately for each task with Full-Scale IQ replaced with the four WAIS subscales (Fig. 5). Fifth, a linear mixed effect analysis was conducted with task as a random factor. Different cognitive predictors were used to find the best model for the data (Table 2). Finally, a linear mixed effect analysis was conducted adding eye-movement variables to the model. In the regression analyses, the assumption of no multicollinearity was checked from VIF values (all were below two). The residuals were normally distributed in all analyses, except in a few cases of taskspecific regression analysis. The data analyses were conducted using Matlab, Rstudio and jamovi.

2.9. Data availability

A fully anonymized version of the dataset will be available on the OSF page of the project. All analysis scripts will be released before publication.

3. Results

In the following, we report our findings, starting with the general effect of cognitive abilities and continuing with comparisons against other factors. We then analyze the effects of cognitive components and look at how task-specific they are. Finally, we report on differences among individuals in eye-movements. Throughout, we denote dependent and predictor variables with *italics* and blocked factors with SMALL CAPITALS.

3.1. Cognitive abilities predict general performance

Our main finding is that general cognitive ability is a significant predictor of performance in computer tasks. Fig. 2 crystallizes this: after adjusting for *Age*, expressed *Gender*, and level of *Education*, *Full-Scale IQ* (WAIS 4th Ed.; Wechsler, 2008b) had a significant association with all three outcomes: *Task Success* (fraction of the task finished), *Elapsed Time*, and *Mental Load* (NASA-TLX; Hart and Staveland, 1988).

Task Success correlated most strongly with Full-Scale IQ, Perceptual Reasoning, Response Inhibition, Working Memory, and Computer User Self-Efficacy (CUSE). Elapsed Time, in turn, correlated negatively with CUSE, Full-Scale IQ, Exposure (frequency of use and familiarity with computing platforms), Perceptual Reasoning, and Response Inhibition. Finally, Mental Load was lower for higher Full-Scale IQ and Attention Control. In summary, in line with our hypothesis, all three task-level outcomes varied with general cognitive ability.

3.2. The contribution of cognitive abilities is comparable to that of experience

To better gauge the independent contributions of abilities vs. experience vs. demographic factors, we conducted hierarchical linear regression analyses with four blocks of predictors:

- 1. DEMOGRAPHIC FACTORS (Age, Gender, and Education),
- 2. EXPERIENCE (Exposure, Familiarity, and CUSE),
- 3. COGNITIVE ABILITIES (Full-Scale IQ), and
- 4. EXECUTIVE FUNCTIONS (Response Inhibition and Attention Control).

Fig. 3 presents an overview of the regression analysis results. A detailed breakdown is provided in a table in Appendix A.

Task Success was better predicted by COGNITIVE ABILITIES than by EXPERIENCE. This result is surprising, as competence in computer use is traditionally attributed to acquired skill (Iñiguez-Berrozpe and Boeren, 2020; Wicht et al., 2021). Slightly more than 10% of the variance in Task Success was explained jointly by COGNITIVE ABILITIES (7.9%) and EXECUTIVE FUNCTIONS (2.8%), whereas experience accounted for 6.9% (see Table 1). In the full model, we found that Full-Scale IQ (p=.019), Response Inhibition (p=.030), and CUSE (p=.020) were significant predictors of higher Task Success (see Appendix A). To sum up, cognitive abilities are at least as good a predictor as prior experience for an individual's ability to complete everyday tasks on computers.

Less surprisingly, DEMOGRAPHIC FACTORS – Age in particular – exhibited the strongest effect on Task Success. These explained nearly half of the variance in that outcome (48.6%; Table 1), with Age (p < .001; see Appendix A) standing out as the sole statistically significant demographic variable. Younger participants were better at completing the tasks than older ones were, while the effects of Gender and Education were non-significant (see Appendix A). We conclude then that, while demographic factors explained most of the variance, cognitive abilities

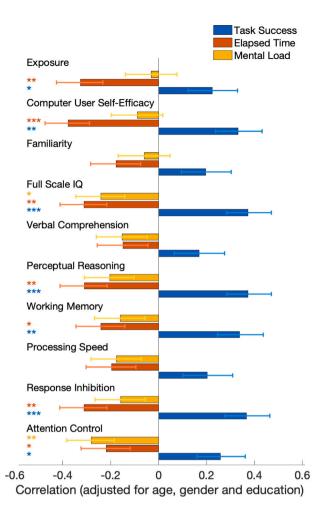


Fig. 2. Correlations of predictive factors with three outcome variables, adjusted for *Age, Gender* and *Education. Full-Scale IQ* and *Attention Control* correlated significantly with all three. Error bars indicate standard error. Significance levels: * p < .05, *** p < .01, *** p < .001.

Т	ask Success	Elapsed Time	Mental Load
Age	-0.44	*** 0.38	▲ 0.32
Gender		•	
Education - mid	•	•	-0.51
Education - high	•	•	-0.71
Exposure	•	** ▼ -0.22	
Familiarity	•	•	•
Computer User Self-Effi	cacy 🛕 0.17	** ▼ -0.22	
Full Scale IQ	▲ 0.19	* -0.19	
Response Inhibition	▲ 0.19	•	
Attention Control	•		•

Fig. 3. An overview of the results from linear regression modeling. Red downward-pointing triangles denote a statistically significant negative association, while green upward-pointing triangles indicate a statistically significant positive one. Gray circles denote non-significant association. The size of the symbols refers to standardized beta coefficients also printed numerically next to the triangles. Age was a strong predictor of all three outcomes and Education was a strong predictor of Mental Load. The predictive effect of COGNITIVE ABILITIES and Exposure/CUSE on Task Success and Elapsed Time were similar in magnitude. Significance levels: * p < .05, ** p < .01, *** p < .001.

Table 1
A breakdown of results from linear hierarchical regression analysis, addressing each outcome separately for the four hierarchical models. Addition of variables improved the models almost in all cases (p-values in the last column). The additional variance explained by the model (ΔR^2) varied from 0.6% to 12.3%. Overall, experience and cognition (cognitive abilities combined with executive functions) explained roughly similar amounts of variance. cognition explained 7.0–10.7%, experience explained 0.6–12.3%, and demographic factors explained 30%—49% of the variance.

						Model Co	omparis	ons		
		Adj. R ²	AIC	BIC	р	ΔR²	F	df1	df2	р
Task Success										
demographic factors	1	0.49	-124	-109	<.001					
+ experience	2	0.54	-131	-109	<.001	0.069	4.34	3	80	0.007
+ cognitive abilites	3	0.62	-147	-122	<.001	0.079	18.13	1	79	0.000
+ executive functions	4	0.65	-151	-121	<.001	0.028	3.47	2	77	0.036
Elapsed Time										
demographic factors	1	0.37	780	795	<.001					
+ experience	2	0.48	766	789	<.001	0.123	6.84	3	80	0.000
+ cognitive abilites	3	0.55	754	779	<.001	0.071	13.84	1	79	0.000
+ executive functions	4	0.57	753	783	<.001	0.023	2.27	2	77	0.110
Mental Load										
demographic factors	1	0.30	684	699	<.001					
+ experience	2	0.28	690	712	<.001	0.006	0.24	3	80	0.867
+ cognitive abilites	3	0.32	686	711	<.001	0.040	5.13	1	79	0.026
+ executive functions	4	0.33	686	715	<.001	0.030	1.96	2	77	0.148

played a larger role than prior experience in accounting for success in completing the tasks.

Elapsed Time behaved in a similar way to Task Success: with DEMOGRAPHIC FACTORS explaining 36.9% of the variance, EXPERIENCE 12.3%, COGNITIVE ABILITIES 7.1%, and EXECUTIVE FUNCTIONING 2.3% (see Table 1). The statistically significant variables here were $Age\ (p < .001)$, $CUSE\ (p = .0063)$, $Exposure\ (p = .0046)$, and $Full\text{-Scale}\ IQ\ (p = .036;$ see Appendix A). Users who were younger and users who were more experienced completed the tasks more swiftly than older or less experienced participants. Higher $Full\text{-Scale}\ IQ$ and $CUSE\ predicted\ faster\ completion. Gender, Education, Familiarity, Response Inhibition, and Attention Control\ displayed\ no\ significant\ effects\ on\ Elapsed\ Time.$ In comparison to $Task\ Success$, prior experience had a slightly stronger effect, and cognitive abilities were a less prominent predictor of the time taken.

Mental Load was most strongly influenced by Age and Education, with older participants reporting a greater mental burden than younger ones and with more advanced education being associated with lower Mental Load values. In fact, these were the only statistically significant predictors in the full model (p=.008 for Age; p=.032 for a medium vs. low Education Level and p=.003 for high vs. low; see Appendix A). Full-Scale IQ was significantly associated with Mental Load in Model 3 (p=.026), but not in Model 4 with executive functions included. Demographic factors explained 30.1% of the variance, while experience covered only 0.6%, cognitive abilities 4.0%, and executive functioning 3.0% (see Table 1). To sum up, cognitive abilities had a proportionately smaller effect on mental load than on task success or elapsed time.

3.3. Contributions from cognitive components are strong but task-specific

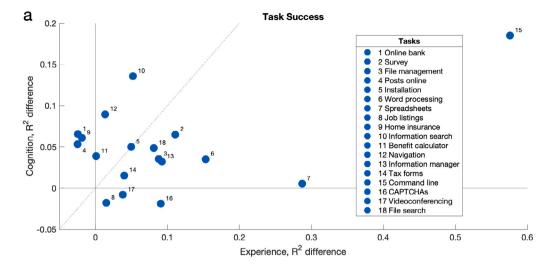
Large differences in the participants' performance and load were found across the 18 tasks. Mean *Task Success* ranged from 30% to 88.3%, mean *Elapsed Time* was between 56.4 and 157.1 s, and mean *Mental Load* lay in the 25.5–59.0% range. To gauge the task-specific effects of cognitive abilities, we performed a similar regression analysis, as outlined above, separately for each task. This enabled judging the relative contributions of cognition (*Full-Scale IQ* + EXECUTIVE FUNCTIONS) and EXPERIENCE to variability in task performance while controlling for DEMOGRAPHIC FACTORS.

We found that the explanatory power of COGNITION and EXPERIENCE varied considerably from task to task. Fig. 4 depicts the contribution of COGNITION (on the y-axis) and of EXPERIENCE (on the x-axis) to the three outcomes for each of the 18 tasks. The dashed diagonal divides the tasks between those better explained by COGNITION (above the diagonal) and those linked more closely to EXPERIENCE (below the diagonal).

Cognition explained as much as 16% of the variance in Task Success (task 10 in Fig. 4a), while the corresponding figures for EXPERIENCE were as high as 60% (task 15 in Fig. 4a). Cognition showed the strongest explanatory role for Task Success with regard to the command-line, information-search, and navigation tasks (Fig. 4a). Cognition explained the most variance in Elapsed Time for the survey, online post service, installation, information search, and command-line tasks (Fig. 4b), and in Mental Load, these abilities explained most of the variance seen in the tasks denoted as navigation, tax form, video conference, and commandline (Fig. 4c). Experience, on the other hand, explained most of task success variance in the word-processing, spreadsheet, and commandline tasks; for Elapsed Time, it accounted for the majority of the variance in the last two of these and the file search task; and it explained Mental Load most strongly in the conditions of the word-processing and the command-line task. To sum up, the roles of prior experience and cognitive abilities are strongly dependent on the computerized task at hand.

Three tasks stood out sharply from the rest in their profile, as Fig. 4 illustrates: The information search task (task 10) was strongly affected by cognition, with experience playing only a minor role. In marked contrast, the spreadsheet task (7) was the opposite: it was heavily dependent on experience, with barely any contribution from cognition. Use of the command-line (15) was strongly affected by both (potential outlier in the upper right corner in Fig. 4a). Other tasks showed weaker or more mixed contributions from various factors. These findings may be explained in part by task design: task 15, involving the commandline interface, was clearly the most difficult of the tasks, requiring rapidly internalizing the given UNIX commands and understanding that one can control the computer via text commands without accessing a graphical user interface. Similarly, handling of the spreadsheet task clearly benefited from prior exposure to use of formulae. Finally, because the information-search problem required interpreting a painting and generating relevant query terms for a search engine, solving it brought cognitive abilities to the fore.

Having ascertained that cognitive abilities' role varied greatly between tasks, we studied whether that role grows with task difficulty. Such a pattern should be reflected as a higher role of Full-Scale IQ in tasks with lower completion rates, longer completion times, and heavier mental load. Interestingly, the effect size (std coefficient) of Full-Scale IQ did not decrease as a function of Task Success (see Appendix B). For Elapsed Time and Mental Load, Full-Scale IQ played a larger part in seemingly easier tasks — i.e., tasks that were completed quickly and with lower Mental Load. Thus the effect of task difficulty



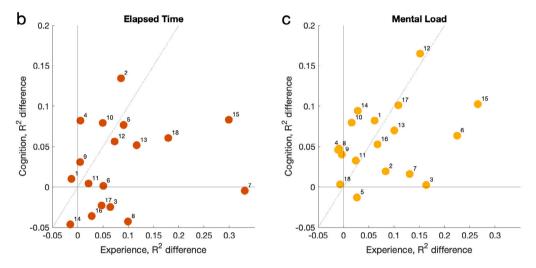


Fig. 4. A comparison of the role of COGNITION (COGNITIVE ABILITIES coupled with EXECUTIVE FUNCTIONS) and of EXPERIENCE (proxied by Exposure, Familiarity, and Computer User Self-Efficacy) across all tasks for (a) Task Success, (b) Elapsed Time, and (c) Mental Load. Cognition and experience differed in their effects on Task Success, task-dependently. For some tasks (e.g., 10 and 12; tasks above diagonal), Cognition explained more variance than experience did, while the reverse was true for some other tasks (e.g., 6 and 7; tasks below diagonal). In a few cases, both or neither were relevant (tasks above zero on both axes, and tasks below zero on both axes, respectively).

on Full-Scale IQ was not found (Task Success) or was the opposite of what we expected (Elapsed Time and Mental Load).

Since the task difficulty did not play a significant role, we further investigated the differences between the tasks relative to all cognitive variables. Tasks differed in their association with the Wechsler Adult Intelligence Scale (WAIS) subscales, *Full-Scale IQ*, *Response Inhibition* and *Attention Control*. All associations (standardized beta coefficients of the regression models) are presented in Fig. 5.

The different subscales independently predicted *Task Success* in several tasks (Fig. 5a). For example, *Verbal Comprehension* predicted success in information search and command-line tasks, *Perceptual Reasoning* in the installation task, and *Working Memory* in survey and spreadsheet tasks. Furthermore, the two variables assessing executive functions predicted success in online banking, post service, and command-line tasks. This suggests, that cognitive domains differently contribute to performance, depending on the task.

Instead of evaluating individual associations between predictors and outcomes, we can evaluate which kind of cognitive profiles predicted good performance and low mental load. For example, in the survey task, high scores in *Working Memory, Processing Speed* and *Verbal Comprehension* subscales predicted high success rate (Fig. 5a), fast completion time (Fig. 5b), and low mental load (Fig. 5c), respectively.

3.4. Cognitive components have an independent effect across tasks

Finally, to understand the effects coupled with *specific* tasks, we conducted linear mixed effects analyses for *Task Success* with task as a random effect. We then sought the best model — that is, the combination of predictor variables that fits the data best (i.e., yielding the lowest AIC value; see Table 2).

Our results suggest that COGNITIVE ABILITIES, EXECUTIVE FUNCTIONS, and DEMOGRAPHIC FACTORS all make clear and independent contributions. The best models incorporated one cognitive variable (either Working Memory, Full-Scale IQ or Perceptual Reasoning) and had either one (Response Inhibition) or both variables for executive functions (see Table 2). We found only small differences between the five best performing models; however, all of those models were clearly better than the ones relying on demographic variables alone or on demographic variables combined with Processing Speed or Verbal Comprehension. In sum, although Age explains much of the variance in task performance, a significant amount is left unexplained; to account for the patterns in the data, we need additional cognitive factors. The top models also proved to be better than models encompassing all of the WAIS subscales, probably due to the subscales' relatively high mutual correlation.

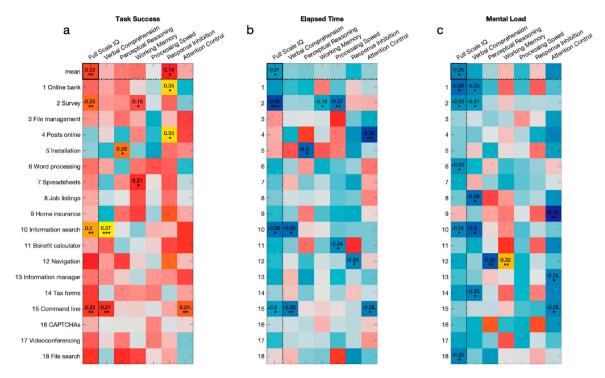


Fig. 5. Associations (standardized beta coefficient) between (a) Task Success, (b) Elapsed Time, and (c) Mental Load and cognition variables (the WAIS Full-Scale IQ and subscales, as well as executive functions). Tasks differed markedly in their patterns of association with the various cognitive predictors. Significance levels: *p < .05, **p < .01, ***p < .001.

Table 2
A comparison of models. All models included DEMOGRAPHIC FACTORS and EXPERIENCE as covariates. Additional variables in the model are listed in the first column. Sorting is by AIC; the models performing best (at the top) featured one variable representing COGNITIVE ABILITIES (Full-Scale IQ, Working Memory or Perceptual Reasoning) and one (Response Inhibition) or both EXECUTIVE FUNCTIONS Variables.

	Parameters	log-Likelihood	AIC	BIC	AIC rank	BIC rank
Working Memory + Response Inhibition	12	5.29	13.42	77.83	1	2
Working Memory + Executive Functions	13	5.80	14.41	84.18	2	7
Full Scale IQ + Response Inhibition	12	4.68	14.64	79.04	3	3
Full scale IQ + Executive Functions	13	5.42	15.16	84.93	4	8
Perceptual Reasoning + Executive Functions	13	5.25	15.51	85.28	5	9
Verbal Comprehension + Perceptual Reasoning + Working						
Memory + Processing Speed + Executive Functions	16	7.25	17.50	103.38	6	16
Full Scale IC	11	1.63	18.74	77.78	7	1
Executive Functions	12	2.26	19.47	83.88	8	6
Verbal Comprehension + Executive Functions	13	3.08	19.84	89.62	9	11
Verbal Comprehension + Perceptual Reasoning + Working						
Memory + Processing Speed	14	3.49	21.03	96.17	10	15
Processing Speed + Executive Functions	13	2.28	21.44	91.21	11	13
Working Memory	11	0.08	21.83	80.87	12	4
Perceptual Reasoning	11	-0.04	22.08	81.12	13	5
Processing Speed	11	-4.81	31.62	90.65	14	12
Verbal Comprehension	11	-5.36	32.72	91.76	15	14
Demographics + Experience	10	-7.35	34.69	88.36	16	10

To get some numeric estimate of the effect of the task itself, we compared the marginal \mathbb{R}^2 values when the task was either a random effect or a fixed factor in the best model above. Overall the model explained 47.1% and 45.9% of the variance of *Task Success* and *Elapsed Time*, respectively, while the task explained 24.2% and 28.0% of the variance.

3.5. Eye- movement patterns are associated with success

To quantify how the participants visually explored and focused on the tasks' content, eye movements were recorded while participants conducted the tasks. A few examples of the fixation heatmaps for the lowest and highest general cognitive ability participants are shown side by side in Fig. 6. For all tasks and participants, we calculated the number of fixations occurring during the first 60 s of a given task, those fixations' average duration, the percentage of them falling beyond

the screen area, and the dispersion of within-screen fixations indicating exploration-oriented behavior. Following this, we augmented the linear mixed effects model with these variables alongside all of the other predictor variables (see Appendix C).

Eye-movement analyses revealed a significant association between average *Fixation Duration* and *Task Success* (p=.028): longer fixations predicted better performance. Also, low fixation counts and a wide spread of fixation locations showed a significant positive association with *Elapsed Time* (p=.011 and p=.001, respectively). That is, more exploration during the first minute of a task predicted needing more time to complete it, and fewer fixations within the first minute predicted smaller *Elapsed Time*. Likewise, *Fixation Count* and *Explorative Behavior* were significantly associated with *Mental Load* (p=.005 and p=.001, respectively). Dispersion pointed to increased load, and the load experienced declined with the number of fixations. Thus, the eye-tracking data showed that people who engaged more exploratively with

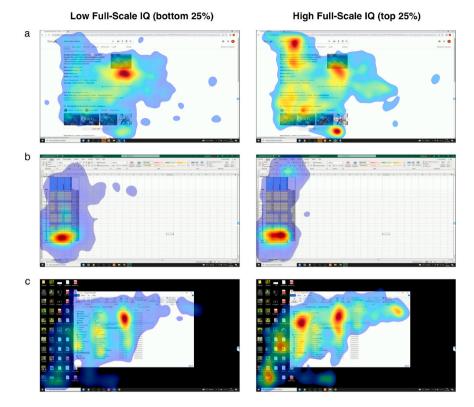


Fig. 6. Fixation maps for participants with 25% lowest (left) and highest (right) general cognitive ability overlaid on the user interface. Tasks from top to bottom: (a) information search, (b) spreadsheet calculations, and (c) local file search. The data are averaged over the subsample of participants, including fixations from the first minute of each given task.

the screen needed more time to complete the given task and perceived a larger mental burden. In contrast, those participants with fewer gaze fixations did not need as much time for completing the task and experienced less mental load. When the eye-movement factors were added to the model, the amount of variance explained in *Mental Load* rose by 2%. The corresponding figures for *Task Success* and *Elapsed Time* rose by only 0.6% and 0.5%, respectively. Therefore, while eye movements were significantly associated with all of our outcome metrics, they seem to be linked more closely to perceived mental load than to success levels.

4. Discussion

Our results provide the first clear evidence that cognitive abilities exert a significant, independent, and broad-based effect over people's ability to use computers. While previous experiments have limited their scope to isolated tasks, self-reporting, or narrow age ranges, our study design allowed investigating the effects of cognitive abilities across a wide spectrum of everyday tasks performed on computers while simultaneously accounting for other known predictors, such as experience and age. According to the results, cognitive abilities predict better overall performance and lower mental load.

Remarkably, the effect size of cognitive abilities is comparable to that of previous experience. In particular, its effect is comparable to that of general experience with computers and specific experience with the application domains in our study. This pattern was visible over a wide range of metrics used in the study to capture the different aspects of prior experience: interface and task familiarity, active computer use, exposure to relevant systems and interfaces, and the beliefs that one has about their capabilities to successfully use computers.

To concretize the importance of cognitive abilities, consider increasing Full-Scale IQ with 15 points — one standard deviation in the IQ scale. This correlates to an increase of 3.4 percentage points

for *Task Success*, a 5.2-second reduction in *Elapsed Time*, and 2.0-unit lower *Mental Load* values. These numbers might seem modest; however, their day-to-day cumulative effect is of practical significance. This corresponded to the effects of increasing age by about 6.5 years. Although the role of particular cognitive components varied across the range of tasks, working memory and executive functioning displayed the largest effects generally.

In what follows, we first look at what the results suggest about the cognitive demands of everyday computer applications. We then discuss implications to understanding of the digital divide.

4.1. Cognitive abilities predict general ability to use computers

Our results attest that strong cognitive abilities serve as a general predictor of successful computer use. *Full-Scale IQ* predicted better success rates, faster task completion, and lower perceived mental load. This is consistent with earlier work in which cognitive abilities correlated with performance in computerized settings specifically designed to be challenging, such as gaming (Bediou et al., 2018; Quiroga et al., 2015), and furthers earlier findings suggesting an ubiquitous effect of intelligence across domains (Deary et al., 2010; Kim, 2008; Schmidt and Hunter, 1998).

4.2. Working memory and executive functioning play a special role in computer use

Beyond demonstrating the role of general intelligence as a significant predictor of successful computer use, our results spotlight the importance of working memory and executive functioning in particular. The model that functioned best, though only by a small margin, employed *Working Memory*, instead of *Full-Scale IQ*, as the construct for cognitive ability. Accordingly, our research indicates that working memory capacity in particular explains large amounts of the variability observed in success with computer use. Since most computer-use

tasks require active maintenance of information – for example, about instructions and goals – the significant role of working memory is understandable. *Full-Scale IQ* is an aggregate measurement, and our tasks' apparently quite weak dependence on other WAIS subscales, like *Verbal Comprehension* and *Processing Speed*, rendered the combined score no better at explaining the data.

We found that successful computer use requires executive functions, too. Many aspects of engaging with computers demand selecting and tracking one's goals while also planning and executing actions in the correct order (Polson et al., 1992). We measured these functions by means of measurements specifically probing response inhibition and attention control. Our data show that the role of executive functions is crucial; adding them to our models always yielded a better fit to the data. It is noteworthy also that the response-inhibition task captured the variability in executive functions more fully than the one centered on attention control. One possible explanation for the strong link between response inhibition and successful computer use is the common presence of distractors. We measured response inhibition with an anti-saccade task, in which participants needed to inhibit saccades to an irrelevant target. User interfaces present numerous task-irrelevant visual elements that may distract the user. The ability to avoid these is advantageous for completing the task. The Response Inhibition variable directly measures user ability to suppress or inhibit eye movements toward an irrelevant target.

4.3. Processing speed may have no or a negligible effect in complex compound tasks

It is interesting that, although the tasks were time-limited, *Processing Speed* was not a significant predictor of performance. On the surface, this appears counter-intuitive. One explanation is that performing well in such computer-based tasks, which take minutes rather than seconds to complete, depends more on complex cognitive processing than it does on rather mechanical execution of a simple, repetitive task of the sort represented by the WAIS *Processing Speed* subtests. The non-significance of processing speed in this respect indicates that typical computer-use tasks draw more on complex cognitive processing than in settings such as typical mobile games.

4.4. Different predictors for task performance and mental load

In addition to task performance (success and completion time) we estimated the mental load participants experienced while performing the tasks. The task-specific results suggest that *Verbal Comprehension* and *Attention Control* were more often associated with mental load than with success or time. This likely reflects participants ability to understand instructions and focus on task-relevant information. Failing in these increased the experienced stress.

Furthermore, in the hierarchical regression analysis (see Appendix A) the association of Full-Scale IQ with Task Success and Elapsed Time remained significant after adding executive functions to the analysis. However, the association of Mental Load with Full-Scale IQ turned to non-significant after adding executive functions. Mental Load, but not Task Success or Elapsed time, was also strongly associated with the level of education. Thus, the association of cognition with Mental Load is less clear than the association of cognition with task performance. We speculate that experiencing stress while using computers is more strongly dependent on the ability of understanding how to solve the task or the ability of trying to solve the task rather than solving the task incorrectly or slowly.

4.5. Cognitive abilities may contribute to the digital divide

Our results enrich the current understanding of the digital divide. While early conceptions of it highlighted the importance of access to technology, which mirrors socioeconomic inequalities, more recent work has shifted focus toward the role of acquired skills (Hargittai et al., 2019; OECD, 2012; Wicht et al., 2021). Moreover, the two are linked; those with access develop stronger skills and hence benefit more from computing, with the effects accumulating over the years and decades.

Our finding suggests that a heretofore unrecognized gap exists: people with more advanced cognitive abilities may benefit more fully from computing than others, with the long-term effects remaining unknown. While training in computer use can reduce gulfs of this nature, the gap might not be entirely eliminated in cases of complex technology. The effect size of cognitive abilities was not directly related to success in completing the tasks, suggesting that the gap is not solely a result of task difficulty or complexity. Until we understand the influences, the gap may remain highly challenging to close. Although age was the most prominent determinant of performance in our study, it was followed by cognitive abilities and computer-use experience in virtually equal measure.

This finding was somewhat unexpected, since we designed our tasks to be similar to common computer-use operations. At the same time, the tasks varied considerably in how much performance was associated with cognitive abilities and how much with experience. Across all 18 tasks, we did not find any evidence that success with only difficult activities might be predicted by *Full-Scale IQ*; importantly, every one of the tasks was hard enough to allow cognition to explain performance differences. Hence, future work on digital skills should include cognitive abilities as a factor of interest.

4.6. User interface design should consider cognitive skills

We conclude that improving user interfaces is critical for attempts to guarantee that the benefits of computing are spread equally. Making computers more available or training people to use them, alone, will not suffice if the user interfaces pose high cognitive demands. Our main finding suggests that calls for design that takes individual abilities into account have not been heeded (Wobbrock et al., 2018). Everyday tasks with computers are not only frustrating (Bessiere et al., 2006; Hertzum and Hornbæk, 2023), but so difficult that a person's cognitive abilities are predictive of their task completion rates. Our more detailed results suggest what to prioritize in efforts to address this. Specifically, our models suggest that design should focus on minimizing reliance on executive functions and working memory. What does this mean in practice?

Executive functions are critical in recalling information and inhibiting irrelevant responses. To minimize executive load, user interfaces should rely not on people's recall but on ability to recognize items (Norman, 2013). Presenting options visually with recognizable graphics and labels can help tap this opportunity. However, a known tradeoff is that it may tax *other* cognitive abilities. As more elements are displayed, more elements compete for attention, and more searching and navigating will be needed. Thus, to lower executive load, designers need to, at the same time, minimize the number of elements on display while making them sufficiently recognizable.

Working memory, on the other hand, entails keeping track of intermediate results while performing tasks (Proctor and Vu, 2007). We saw strong effects of working memory on form-filling tasks and tasks that require comparing data. To remedy this, instead of users *mentally* carrying over results from one step to another, or performing operations on them, user interfaces could *externalize* such information (Scaife and Rogers, 1996). However, externalization can be hard to realize, as it can make the user interface more cluttered. Interaction techniques could help users by supporting visualization and manipulation

of intermediate results at will. For example, interaction techniques could help users more easily transfer information to a form from other documents. One trade-off here concerns learnability: learning to use such techniques will take some practice.

Our results suggest that today's interfaces rely on extensive scanning and exploration. They present vast amounts of information, of several types, all at once. This demands attention control from users. User interfaces should better guide users' attention, such that information gets handled smoothly, in appropriate order (Polson et al., 1992). What can be done to address this? Reducing "bloat", or the number of unnecessary features, can help (McGrenere and Moore, 2000). Further, promising results have been obtained by designing task-centric user interfaces as opposed to feature-centric interfaces (Lafreniere et al., 2014). In addition, predictive models of visual saliency are emerging that could be used to design attention-guiding interfaces (Jiang et al., 2023), although it is an open question how to calibrate them based on individual differences in cognitive abilities. An orthogonal approach is to leverage users' prior knowledge. Exploiting prior knowledge may help reduce the load on executive functioning. User interfaces that are "intuitive" are such that users can exploit their existing knowledge to predict how tasks are completed (Polson et al., 1992). In practice this means using metaphors, analogs, and conceptual models that are familiar from everyday life.

We admit that while these are crucial goals toward more accessible interfaces, these recommendations are not novel. However, our results may help rethink priorities. While the role of cognitive abilities *in general* have been recognized (e.g., Johnson (2020)), our results suggest the primacy of two components *in particular*. They call for more work to understand how to address them in practice while accounting for other relevant factors like cultural differences.

4.7. Design processes need to include users with diverse cognitive skills

Beyond user interface design, user-centered design processes could be reconsidered. User groups should cover diverse abilities. Presently, user interfaces get designed mostly under a "one size fits all" policy, often with focus on commercially interesting markets: groups who are likely to adopt new products early. Our results suggest that users with lower cognitive abilities need to be considered as a target segment. However, there are complexities in designing for diverse user groups. An improvement in one group may cause decreases in another. One promising avenue to tackle this is via stronger focus on learnability. The focus in design is often on ease-of-use, which may dismiss the opportunity to support the growth of skills over time. The different groups' needs for learning need to be accounted for (Sarkar, 2023). Spreadsheet computing is a point in case: the task showed an exceptionally strong effect of prior experience while the effect of cognitive abilities was virtually non-existent. Command-line interaction, by contrast, required both cognitive abilities and experience. User interface design could strive to minimize the effect of cognitive abilities by supporting learning via methods like scaffolding (Soloway et al., 1994).

Evaluation practice should acknowledge that cognitive abilities have a significant effect in measurements of usability. What is achievable for a user with high cognitive abilities may be out of reach for some other user. Alas, literature on usability testing is virtually silent about cognitive abilities as a factor (Himmelsbach et al., 2019). Therefore, new evaluation practices should be developed that encourage testing of products with more diverse samples in order to assure more equitable usability. However, acquiring estimates of cognitive abilities is often impractical. Running a test battery may take longer than a usability study. At the moment this remains an open problem. However, we warn against the practice of convenience samples, where colleagues or students are recruited because they are available. This may inadvertently produce biased estimates of usability.

4.8. Limitations

We note some limitations of the study design. First, we did not quite reach our target sample size of N=100, as we struggled to recruit older males (see Appendix D). Second, in completing the computerized tasks, the participants could not choose to use the operating system they have more experience in. However, most of the tasks were designed to be independent of the operating system, and the effects of familiarity with the operating system and task-related interfaces were accounted for in our analyses. Third, we did not have any baseline computer task, because we were not able to identify a single baseline task for the 18 experimental tasks we had. Especially in terms of eye-movement patterns, this could have helped to control for individual differences. Fourth, although the tasks were ecologically valid and similar to reallife computer use, the three-minute time limit for each task may have had an effect on some participants' performance.

4.9. Conclusion

Our results suggest that contemporary user interfaces are getting so complex that their design is starting to affect inclusivity. Our data show that cognitive abilities predict people's ability to complete challenging but normal tasks with computers. Users who score higher in cognitive abilities are faster and more successful. This study is the first to show that such effects are broad and can be independent of other factors, specifically the effect of experience. Contrary to conventional wisdom, being experienced with computers is insufficient to eliminate this effect. This effect is large enough to have practical significance. We furthermore found that, of the specialized abilities, working memory and executive functioning had the largest effects. We conclude that more research is needed to understand two questions: first, what are the consequences of these effects on an individual's ability to function in the information society and, second, how can we design user interfaces to decrease the role of cognitive abilities, in particular executive functioning and working memory?

CRediT authorship contribution statement

Erik Lintunen: Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation. Viljami Salmela: Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Petri Jarre: Writing – original draft, Methodology, Investigation, Conceptualization. Tuukka Heikkinen: Writing – original draft, Methodology, Investigation, Conceptualization. Markku Kilpeläinen: Writing – original draft, Validation, Supervision, Investigation, Conceptualization. Markus Jokela: Writing – original draft, Methodology, Conceptualization. Antti Oulasvirta: Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data is shared on the project's OSF page https://osf.io/7b6uv.

Acknowledgments

This work was supported by the Research Council of Finland (flag-ship program: Finnish Center for Artificial Intel- ligence, FCAI, grants 328400, 345604, 341763; Subjective Functions, grant 357578), by Aalto University's Department of Information and Communications Engineering, and by University of Helsinki's Department of Psychology and Logopedics.

Table 3
Hierarchical linear regression model variables for the three outcomes, and for Models 2–4. Adding a COGNITIVE ABILITIES variable (Full-Scale IQ) to the analyses, under Model 3, reduced the effect exhibited by Education and Age. With the full model (Model 4), adding EXECUTIVE FUNCTIONS variables (Response Inhibition and Attention Control) reduced the effect visible from Full-Scale IQ. Under the full model, the statistically significant predictors were Age, Education, Computer User Self-Efficacy, Full-Scale IQ, and Response Inhibition, depending on the outcome.

	MODEL 2			MODEL 3			MODEL 4		
	Stand. est	SE	р	Stand. est.	SE	р	Stand. est.	SE	р
Task Success									
Intercept	0.578	0.111	<.001	0.177	0.138	0.203	0.124	0.136	0.365
Age	-0.619	0.001	<.001	-0.577	0.001	<.001	-0.444	0.001	<.001
Gender	0.066	0.024	0.664	0.066	0.022	0.633	0.008	0.022	0.952
Education, mid	0.429	0.030	0.023	0.243	0.028	0.166	0.287	0.027	0.094
Education, high	0.242	0.029	0.184	0.038	0.027	0.825	0.093	0.027	0.578
Exposure	0.107	0.001	0.165	0.137	0.001	0.053	0.129	0.001	0.062
Familiarity	0.070	0.001	0.428	0.091	0.001	0.259	0.119	0.001	0.131
Computer User Self-Efficacy	0.195	0.001	0.018	0.185	0.001	0.014	0.170	0.001	0.020
Full Scale IQ				0.297	0.001	0.000	0.189	0.001	0.019
Response Inhibition							0.193	0.069	0.030
Attention Control							0.089	0.020	0.251
Elapsed Time									
Intercept	131.121	18.095	<.001	189.784	23.041	<.001	196.898	23.060	<.001
Age	0.540	0.164	<.001	0.500	0.154	<.001	0.381	0.186	0.000
Gender	0.112	3.977	0.489	0.112	3.692	0.455	0.163	3.689	0.278
Education, mid	-0.283	4.875	0.156	-0.106	4.676	0.579	-0.145	4.625	0.443
Education, high	-0.222	4.748	0.252	-0.028	4.592	0.881	-0.076	4.571	0.681
Exposure	-0.198	0.156	0.017	-0.226	0.145	0.004	-0.220	0.144	0.005
Familiarity	-0.035	0.165	0.706	-0.055	0.154	0.528	-0.081	0.153	0.353
Computer User Self-Efficacy	-0.247	0.115	0.005	-0.237	0.107	0.004	-0.224	0.106	0.006
Full Scale IQ				-0.283	0.141	0.000	-0.186	0.162	0.036
Response Inhibition							-0.169	11.788	0.086
Attention Control							-0.085	3.323	0.319
Mental Load									
Intercept		11.713	0.003	59.823	15.666	0.000	60.441	15.739	0.000
Age	0.453	0.106	0.000	0.423	0.105	0.000	0.325	0.127	0.008
Gender	-0.012	2.575	0.951	-0.012	2.511	0.950	0.012	2.518	0.947
Education, mid	-0.616	3.156	0.010	-0.483	3.179	0.042	-0.508	3.157	0.032
Education, high	-0.847	3.073	0.000	-0.701	3.122	0.003	-0.711	3.120	0.003
Exposure	-0.007	0.101	0.945	-0.028	0.099	0.765	-0.038	0.098	0.686
Familiarity		0.107	0.778	-0.046	0.105	0.670	-0.065	0.104	0.543
Computer User Self-Efficacy	-0.064	0.075	0.528	-0.057	0.073	0.567	-0.034	0.073	0.730
Full Scale IQ				-0.212	0.096	0.026	-0.135	0.111	0.218
Response Inhibition							-0.033	8.046	0.787
Attention Control							-0.202	2.268	0.059

Appendix A. Regression model results

See Table 3.

Appendix B. Standardized coefficient of full-scale IQ

See Fig. 7.

Appendix C. Fixation statistics

See Table 4.

Appendix D. Participant sample

Expressed Gender	Age	N
Female	[20, 35)	17
Female	[35, 50)	17
Female	[50, 65]	17
Male	[20, 35)	17
Male	[35, 50)	13
Male	[50, 65]	7
Other	[20, 65]	0

Appendix E. Task descriptions

See Table 5.

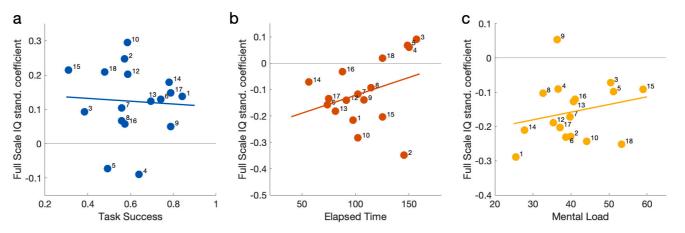


Fig. 7. The relationship between the role of cognitive abilities (captured by the standardized coefficient of Full-Scale IQ) and indices of task difficulty: average Task Success (A), average Elapsed Time (B), and average Mental Load (C). Note that the rising slope of the regression line in panes B and C indicates a diminishing absolute coefficient — i.e., reduction in the role of Full-Scale IQ.

Table 4
Statistics from a linear mixed effects analysis, where eye-movement variables (Number of Fixations, Duration of Fixations, Out-of-screen Fixations, and Dispersion of Fixations) were added to the model. Longer mean duration of fixations predicted better Task Success; lower number of fixations predicted lower Elapsed Time (faster completion) and lower Mental Load in completing the task; and higher dispersion of fixations predicted higher Elapsed Time (slower completion) and higher Mental Load.

Task Success					Elapsed Time				Mental Load			
	Estimate	SE	df	р	Estimate	SE	df	р	Estimate	SE	df	p
Intercept	0.639	0.034	19.900	<.001	107.656	6.124	20.000	<.001	39.190	2.118	33.400	<.001
Education, low	0.057	0.027	77.300	0.035	-2.668	4.709	77.400	0.573	-6.924	3.245	78.700	0.036
Education, high	0.039	0.027	79.600	0.151	-0.245	4.732	79.700	0.959	-10.109	3.243	79.500	0.003
Age	-0.005	0.001	82.400	<.001	0.637	0.180	82.500	<.001	0.274	0.123	80.700	0.029
Exposure	0.002	0.001	77.600	0.079	-0.452	0.159	77.800	0.006	0.085	0.110	78.800	0.439
Computer User Self-Efficacy	0.001	0.001	80.100	0.183	-0.149	0.102	80.200	0.149	0.047	0.070	79.600	0.5
Familiarity	0.003	0.000	1392.700	<.001	-0.463	0.045	1402.900	<.001	-0.256	0.018	1419.100	<.001
Verbal Comprehension	0.001	0.001	79.000	0.416	-0.076	0.146	79.100	0.602	-0.031	0.100	79.500	0.757
Perceptual Reasoning	0.003	0.001	78.000	0.013	-0.262	0.173	78.200	0.134	-0.116	0.119	79.000	0.334
Working Memory	0.002	0.001	77.200	0.035	-0.181	0.161	77.400	0.266	-0.037	0.111	78.600	0.737
Processing Speed	0.001	0.001	77.300	0.325	0.001	0.148	77.500	0.997	-0.003	0.102	78.700	0.979
Response Inhibition	0.161	0.069	78.200	0.022	-11.407	12.125	78.300	0.35	-0.390	8.335	79.100	0.963
Attention Control	0.012	0.021	77.300	0.546	-3.864	3.620	77.500	0.289	-3.173	2.494	78.700	0.207
Number of fixations	0.000	0.000	1183.400	0.504	-0.071	0.028	1212.000	0.011	-0.031	0.011	1412.800	0.005
Duration of fixations	0.363	0.165	762.600	0.028	-13.799	28.508	801.800	0.628	-8.354	11.928	1392.900	0.484
Out-of-screen fixations	0.046	0.171	1294.100	0.787	-7.879	29.445	1315.800	0.789	13.748	11.656	1432.100	0.238
Dispersion of fixations	-0.076	0.061	1399.600	0.212	34.574	10.455	1410.800	<.001	21.161	4.041	1319.600	<.001

Table 5

Task d	lescriptions.	
1	Online banking	The participant is asked to use an online banking service with the given log-in details, then asked to find a specific bank statement and to download this onto the test computer.
2	Survey creation	The participant is logged in to an online service for creating surveys. The task is to create a questionnaire, given questions and answers. The participants is then asked to generate a shareable link for their questionnaire.
3	File management	Given a set of files, the participant is asked to create a folder and to differentiate PDF files from others before compressing the folder and uploading the compressed file into the cloud repository.
4	Online postal service	The participant is asked to log in to an online postal service, given log-in details, then asked to create a new bill with given details, mark the bill paid, and finally archive it.
5	A software install	The participant is asked to find the installer file for named software. After downloading and installing the correct version, the user is asked to configure settings.
6	Word-processing	The task is to "insert" page numbering for an empty multi-page text document. The participant is given specific instructions for formatting.
7	Spreadsheet calculations	The participant is given a template for a working hour list, and the task is to calculate sums of columns and percentages of total working hours into specific cells.
8	A job listing	The participant is given details for a job opening, and the task is to draft a listing using the specified online service. The listing must include all details given, and it must be listed in the correct category.
9	Home insurance	The participant is given details of a home that has been rented gets presented with a website that offers home insurance. The task is to find a quote calculator and calculate an estimated annual price for the property. The participant is then asked to generate a PDF file of the basket and to download this onto the test computer.
10	Information search	The participant is shown an image of a painting. The task is to find its name and the work that it has been named after.
11	Benefit calculators	The participant is directed to the website of the Social Insurance Institution of Finland. For the given details, the task is to find the appropriate calculator and get it to supply specific details about the relevant benefit.
12	Navigation	The participant is directed to an online navigation service and is asked to find a route that fulfills given criteria.
13	An information manager	The participant is presented with information on a task in need of completion. This task is to be added to personal information manager software. The task is then marked as completed, and a status report is sent to a given e-mail address.
14	A tax form	The participant is given the name of a tax form of interest. The task is to find the form, a PDF version of it, and fill in all instances of a named field appearing in the form.
15	Command-line processing	The participant is presented with a command-line interface and is given the essential commands needed in this task. The task is to find and navigate to the target directory, open a text file in that directory, and identify a piece of information from the file on the basis of given criteria.
16	CAPTCHAs	The participant is given a set of nine CAPTCHA puzzles. The task is to complete all of these in the given time.
17	Remote work	The participant is presented with a commonly used video-conferencing and group communication software. The task is to start a new meeting, given specific criteria.
18	Local file search	The participant is given the name of a file to be found from the test computer, given some hints to its location.

Appendix F. Task screenshots

See Fig. 8.

Appendix G. Individual differences in task performance

See Fig. 9.

Appendix H. Task differences

See Fig. 10.

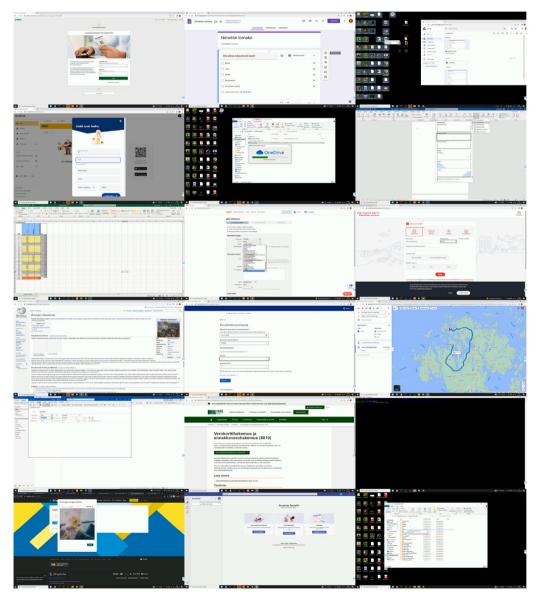


Fig. 8. Screenshots of the 18 tasks. From top to bottom, left to right: in the order given in Appendix E.

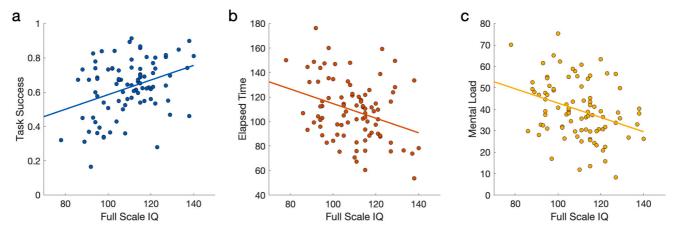


Fig. 9. There were large individual-to-individual differences in task performance. Stronger cognitive ability (Full-Scale IQ) correlated with (a) more successful completion of the tasks (Task Success), (b) shorter completion times (Elapsed Time), and (c) lower mental load (Mental Load).

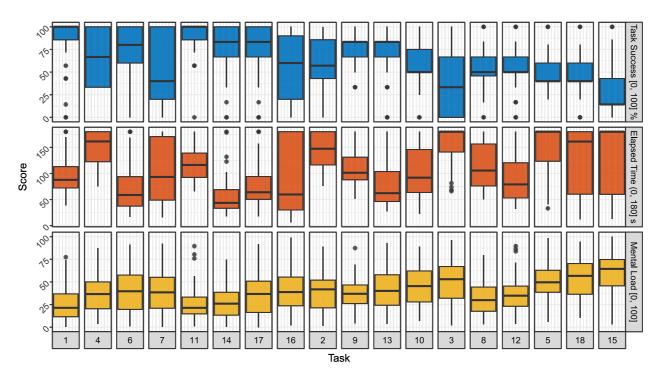


Fig. 10. Depiction of the 18 tasks' wide span of differences in Task Success, Elapsed Time, and Mental Load (where tasks are ordered by Task Success upper quartile, decreasing).

References

- Bediou, B., Adams, D.M., Mayer, R.E., Tipton, E., Green, C.S., Bavelier, D., 2018. Metaanalysis of action video game impact on perceptual, attentional, and cognitive skills. Psychol. Bull. 144 (1), 77.
- Bessiere, K., Newhagen, J.E., Robinson, J.P., Shneiderman, B., 2006. A model for computer frustration: The role of instrumental and dispositional factors on incident, session, and post-session frustration and mood. Comput. Hum. Behav. 22 (6), 941–961.
- Boot, W.R., Charness, N., Czaja, S.J., Sharit, J., Rogers, W.A., Fisk, A.D., Mitzner, T., Lee, C.C., Nair, S., 2015. Computer proficiency questionnaire: assessing low and high computer proficient seniors. Gerontologist 55 (3), 404–411.
- Card, S.K., Moran, T.P., Newell, A., 1980. Computer text-editing: An information-processing analysis of a routine cognitive skill. Cogn. Psychol. 12 (1), 32–74
- Cassidy, S., Eachus, P., 2002. Developing the computer user self-efficacy (cuse) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. J. Educ. Comput. Res. 26 (2), 133–153. http://dx.doi.org/10.2190/JGJR-0KVL-HRF7-GCNV.
- Choi, E.Y., Wisniewski, K.M., Zelinski, E.M., 2021. Information and communication technology use in older adults: A unidirectional or Bi-directional association with cognitive function? Comput. Hum. Behav. 121, 106813.
- Deary, I.J., Weiss, A., Batty, G.D., 2010. Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. Psychol. Sci. Public Interest 11 (2), 53–79.
- Dobrowolski, P., Hanusz, K., Sobczyk, B., Skorko, M., Wiatrow, A., 2015. Cognitive enhancement in video game players: The role of video game genre. Comput. Hum. Behav. 44, 59–63.
- Draheim, C., Tsukahara, J.S., Martin, J.D., Mashburn, C.A., Engle, R.W., 2021. A toolbox approach to improving the measurement of attention control.. J. Exp. Psychol. [Gen] 150 (2), 242–275. http://dx.doi.org/10.1037/xge0000783.
- Falck, O., Heimisch-Roecker, A., Wiederhold, S., 2021. Returns to ICT skills. Res. Policy 50 (7), 104064.
- Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., Hewitt, J.K., 2006. Not all executive functions are related to intelligence. Psychol. Sci. 17 (2), 172–179.
- Goldberg, J.H., Kotval, X.P., 1999. Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. 24 (6), 631–645. http://dx.doi.org/10. 1016/S0169-8141(98)00068-7.
- Hargittai, E., Piper, A.M., Morris, M.R., 2019. From internet access to internet skills: digital inequality among older adults. Univers. Access Inf. Soc. 18, 881–890.
- Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Hum. Ment. Workload 1 (3), 139–183.
- Hertzum, M., Hornbæk, K., 2023. Frustration: Still a common user experience. ACM Trans. Comput. Hum. Interact.
- Himmelsbach, J., Schwarz, S., Gerdenitsch, C., Wais-Zechmann, B., Bobeth, J., Tscheligi, M., 2019. Do we care about diversity in human computer interaction: A comprehensive content analysis on diversity dimensions in research. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–16.
- Iñiguez-Berrozpe, T., Boeren, E., 2020. Twenty-first century skills for all: Adults and problem solving in technology rich environments. Technol. Knowl. Learn. 25, 929–951.
- Jiang, Y., Leiva, L.A., Rezazadegan Tavakoli, H., RB Houssel, P., Kylmälä, J., Oulasvirta, A., 2023. UEyes: Understanding visual saliency across user interface types. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–21.
- Johnson, J., 2020. Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Guidelines. Morgan Kaufmann.
- Kieras, D., Polson, P.G., 1985. An approach to the formal analysis of user complexity. Int. J. Man-Mach. Stud. 22 (4), 365–394.
- Kim, K.H., 2008. Meta-analyses of the relationship of creative achievement to both iq and divergent thinking test scores. J. Creat. Behav. 42 (2), 106–130.
- Kohonen, I., Kuula-Luumi, A., Spoof, S.K., 2019. The Ethical Principles of Research with Human Participants and Ethical Review in the Human Sciences in Finland. Finnish National Board on Research Integrity TENK Guidelines 2019. Finnish National Board on Research Integrity TENK, URL: https://tenk.fi/en/ethical-review/ethical-review-human-sciences.
- Kokkinakis, A.V., Cowling, P.I., Drachen, A., Wade, A.R., 2017. Exploring the relationship between video game expertise and fluid intelligence. PLoS One 12 (11), e0186621.
- Kovacs, K., Conway, A.R.A., 2016. Process overlap theory: A unified account of the general factor of intelligence. Psychol. Inq. 27 (3), 151–177. http://dx.doi.org/10. 1080/1047840X.2016.1153946.
- Lafreniere, B., Bunt, A., Terry, M., 2014. Task-centric interfaces for feature-rich software. In: Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design. pp. 49–58.

- Lazar, J., Jones, A., Shneiderman, B., 2006. Workplace user frustration with computers: An exploratory investigation of the causes and severity. Behav. Inf. Technol. 25 (03), 239–251.
- Liversedge, S.P., Findlay, J.M., 2000. Saccadic eye movements and cognition. Trends Cogn. Sci. 4 (1), 6–14.
- McGrenere, J., Moore, G., 2000. Are we all in the same "bloat"? In: Graphics Interface, vol. 2000, pp. 187–196.
- Mitzner, T.L., Savla, J., Boot, W.R., Sharit, J., Charness, N., Czaja, S.J., Rogers, W.A., 2019. Technology adoption by older adults: Findings from the PRISM trial. Gerontologist 59 (1), 34–44.
- Moacdieh, N., Sarter, N., 2015. Display clutter: A review of definitions and measurement techniques. Hum. Factors 57 (1), 61–100. http://dx.doi.org/10.1177/0018720814541145
- Nair, S.N., Czaja, S.J., Sharit, J., 2007. A multilevel modeling approach to examining individual differences in skill acquisition for a computer-based task. J. Gerontol. Ser. B. Psychol. Sci. Soc. Sci. 62. 85–96.
- Nguyen, O.T., Jenkins, N.J., Khanna, N., Shah, S., Gartland, A.J., Turner, K., Merlo, L.J., 2021. A systematic review of contributing factors of and solutions to electronic health record-related impacts on physician well-being. J. Am. Med. Inform. Assoc. 28 (5), 974–984.
- Norman, D., 2013. The Design of Everyday Things: revised and Expanded Edition. Basic Books.
- OECD, 2012. Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills. OECD Publishing Paris.
- Ownby, R.L., Czaja, S.J., Loewenstein, D., Rubert, M., 2008. Cognitive abilities that predict success in a computer-based training program. Gerontologist 48 (2), 170–180.
- Pak, R., Rogers, W.A., Fisk, A.D., 2006. Spatial ability subfactors and their influences on a computer-based information search task. Hum. Factors 48 (1), 154–165.
- Polson, P.G., Lewis, C., Rieman, J., Wharton, C., 1992. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. Int. J. Man-Mach. Stud. 36 (5), 741–773.
- Proctor, R.W., Vu, K.P.L., 2007. Human information processing: an overview for human-computer interaction. Hum. Comput. Interact. Handb. 69–88.
- Quiroga, M.A., Diaz, A., Roman, F.J., Privado, J., Colom, R., 2019. Intelligence and video games: Beyond "brain-games". Intelligence 75, 85–94.
- Quiroga, M.Á., Escorial, S., Román, F.J., Morillo, D., Jarabo, A., Privado, J., Hernández, M., Gallego, B., Colom, R., 2015. Can we reliably measure the general factor of intelligence (g) through commercial video games? Yes, we can!. Intelligence 53, 1–7
- Sachdev, P.S., Blacker, D., Blazer, D.G., Ganguli, M., Jeste, D.V., Paulsen, J.S., Petersen, R.C., 2014. Classifying neurocognitive disorders: the DSM-5 approach. Nat. Rev. Neurol. 10 (11), 634–642. http://dx.doi.org/10.1038/nrneurol.2014.181.
- Sarkar, A., 2023. Should computers be easy to use? Questioning the doctrine of simplicity in user interface design. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–10.
- Scaife, M., Rogers, Y., 1996. External cognition: how do graphical representations work? Int. J. Hum.-Comput. Stud. 45 (2), 185–213.
- Schmidt, F.L., Hunter, J.E., 1998. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychol. Bull. 124 (2), 262.
- Schraagen, J.M., Chipman, S.F., Shalin, V.L., 2000. Cognitive Task Analysis. Psychology Press.
- Shapiro, C.M., Auch, C., Reimer, M., Kayumov, L., Heslegrave, R., Huterer, N., Driver, H., Devins, G.M., 2006. A new approach to the construct of alertness. J. Psychosom. Res. 60 (6), 595–603. http://dx.doi.org/10.1016/j.jpsychores.2006.04. 012.
- Sharit, J., Czaja, S.J., 1999. Performance of a computer-based troubleshooting task in the banking industry: Examining the effects of age, task experience, and cognitive abilities. Int. J. Cogn. Ergon. 3 (1), 1–22.
- Sharit, J., Hernández, M.A., Czaja, S.J., Pirolli, P., 2008. Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. ACM Trans. Comput.-Hum. Interact. 15 (1), 1–25.
- Soloway, E., Guzdial, M., Hay, K.E., 1994. Learner-centered design: The challenge for HCI in the 21st century. Interactions 1 (2), 36–48.
- Spearman, C., 1904. "General intelligence", objectively determined and measured. Am. J. Psychol. 15 (2), 201–292, URL: http://www.jstor.org/stable/1412107.
- Tun, P.A., Lachman, M.E., 2010. The association between computer use and cognition across adulthood: use it so you won't lose it? Psychol. Aging 25 (3), 560.
- Wagner, N., Hassanein, K., Head, M., 2010. Computer use by older adults: A multi-disciplinary review. Comput. Hum. Behav. 26 (5), 870–882.
- Wechsler, D., 2008a. Wechsler Adult Intelligence Scale, fourth ed. http://dx.doi.org/ 10.1037/t15169-000.
- Wechsler, D., 2008b. Wechsler adult intelligence scale-Fourth Edition (WAIS-IV), vol. 22, (no. 498), NCS Pearson, San Antonio, TX, p. 1.

- Westerman, S.J., Davies, D.R., Glendon, A.I., Stammers, R., Matthews, G., 1995. Age and cognitive ability as predictors of computerized information retrieval. Behav. Inf. Technol. 14 (5), 313–326.
- Wicht, A., Reder, S., Lechner, C.M., 2021. Sources of individual differences in adults' ICT skills: A large-scale empirical test of a new guiding framework. Plos One 16 (4), e0249574.
- Wilson, M., Scalise, K., Gochyyev, P., 2015. Rethinking ICT literacy: From computer skills to social network settings. Think. Ski. Creat. 18, 65–80.
- WMA, 2013. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. World Health Organization, URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.
- Wobbrock, J.O., Gajos, K.Z., Kane, S.K., Vanderheiden, G.C., 2018. Ability-based design. Commun. ACM 61 (6), 62–71.