AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Bulletin

ISSN: 0033-2909

https://doi.org/10.1037/bul0000443

The Unpleasantness of Thinking: A Meta-Analytic Review of the Association Between Mental Effort and Negative Affect

Louise David[1], Eliana Vassena[1, 2], and Erik Bijleveld[1]

[1] Behavioural Science Institute, Radboud University

[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center

Influential theories in psychology, neuroscience, and economics assume that the exertion of mental effort should feel aversive. Yet, this assumption is usually untested, and it is challenged by casual observations and previous studies. Here, we meta-analyze (a) whether mental effort is generally experienced as aversive and (b) whether the association between mental effort and aversive feelings depends on population and task characteristics. We meta-analyzed a set of 170 studies (from 125 articles published in 2019–2020; 358 different tasks; 4,670 unique subjects). These studies were conducted in a variety of populations (e.g., health care employees, military employees, amateur athletes, college students; data were collected in 29 different countries) and used a variety of tasks (e.g., equipment testing tasks, virtual reality tasks, cognitive performance tasks). Despite this diversity, these studies had one crucial common feature: All used the NASA Task Load Index to examine participants' experiences of effort and negative affect. As expected, we found a strong positive association between mental effort and negative affect. Surprisingly, just one of our 15 moderators had a significant effect (effort felt somewhat less aversive in studies from Asia vs. Europe and North America). Overall, mental effort felt aversive in different types of tasks (e.g., tasks with and without feedback), in different types of populations (e.g., university-educated populations and non-university-educated populations), and on different continents. Supporting theories that conceptualize effort as a cost, we suggest that mental effort is inherently aversive.

Public Significance Statement

In practice, employers and educators often stimulate employees and students to exert mental effort. On the surface, this seems to work well: Employees and students are indeed often observed to opt for mentally effortful activities. One may be tempted to conclude from this observation that employees and students may readily learn to enjoy mental effort. Our results suggest that this conclusion would be false: Our meta-analysis shows that mental effort feels unpleasant across a wide range of populations and tasks. This insight is important for professionals (e.g., engineers, educators) who design tasks, tools, interfaces, materials, and instructions. When employees and students are required to exert substantial mental effort, it is sensible to support or reward them (e.g., by providing structure, by balancing demanding tasks with tasks that foster engagement, or by highlighting achievements).

Keywords: effort paradox, feeling of effort, subjective effort, cognitive effort, NASA Task Load Index

Supplemental materials: https://doi.org/10.1037/bul0000443.supp

Across scientific disciplines, people are often assumed to be effort avoiders. In psychology, this assumption is embodied in the classic "law of less work." Rooted in classic work on animal learning (Hull, 1943; Tsai, 1932), the "law of less work"—or the general assumption that people minimize their expenditure of effort—has had an immense influence on modern psychology. For example, it is now widely accepted that people often rely on heuristics and stereotypes, allowing them to expend less mental effort (Shah & Oppenheimer, 2008). Also,

Blair T. Johnson served as action editor.

Louise David https://orcid.org/0000-0003-1973-4568

Eliana Vassena https://orcid.org/0000-0002-5761-6813

Erik Bijleveld https://orcid.org/0000-0001-5892-3236

Louise David is now at the School of Health Professions Education, Maastricht University.

The authors thank Deepshikha Prasad and Fenna Andriessen for their help with data processing. Preregistration, coding protocol, data, code, and analysis scripts are available at https://osf.io/mktbr/.

Louise David played an equal role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, writing–original draft, and writing–review and editing. Eliana Vassena played an equal role in conceptualization, investigation, methodology, supervision, and writing–review and editing. Erik Bijleveld played an equal role in conceptualization, data curation, formal analysis, investigation, methodology, supervision, visualization, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Erik Bijleveld, Behavioural Science Institute, Radboud University, P.O. Box 9104, 6500HE Nijmegen, The Netherlands. Email: erik.bijleveld@ru.nl

it is now widely accepted that people exert mental effort strategically: They refrain from exerting mental effort unless their effort is compensated by a sufficiently valuable reward (Shenhav et al., 2017). More broadly, the "law of less work" is a cornerstone assumption in the biological and social sciences. For example, in neuroscience and economics (Holmstrom & Milgrom, 1994; Silvetti et al., 2023), effort is often modeled as a *cost*: a quantity that people try to minimize.

The underlying assumption shared across these domains is that effort, including mental effort, is inherently aversive. Yet, direct evidence for this assumption is scarce, as experimental studies often quantify the subjective value of mental effort by observing people's choices but not their experiences. It is unclear whether people's choices are a reliable proxy for the experienced unpleasantness of mental effort. Thus, a core question in this area remains unanswered: Does mental effort really feel aversive?

On the one hand, there are good reasons to think that mental effort should feel aversive. For example, negative affect is involved in initiating cognitive control (Botvinick, 2007; Dignath et al., 2020; Dreisbach & Fischer, 2012; Inzlicht et al., 2015), and cognitive control is assumed to be effortful (Morsella et al., 2009; Silvestrini et al., 2023). On the other hand, somewhat paradoxically, several lines of research suggest that mental effort can feel pleasant rather than aversive (Cacioppo et al., 1996; Eisenberger, 1992; for an overview, see Inzlicht et al., 2018). In this research, we systematically review and meta-analyze prior studies on the association between effort (as experienced during mental tasks) and negative affect. We estimate the strength of this mental effort–negative affect association, and we examine how it varies across tasks and populations. By doing so, we aim to draw a broad conclusion as to whether—and if so, when and for whom—mental effort feels aversive.

Why Mental Effort May Be Inherently Aversive

In an early essay titled "The Feeling of Effort," James (1880) proposed that mental effort arises when people are confronted with internal conflicts (e.g., between different representations or action plans) that they are attempting to resolve. According to this perspective, the experience of mental effort arises when people make decisions between alternatives that involve "mixed good and evil" (p. 22), such as when people decide to get out of bed on a cold morning. Relatedly, building on a set of pioneering experiments, Ach (1910/2006) proposed that acts of will (e.g., choosing to ignore a previously learned response rule) are accompanied by feelings of effort and tension (e.g., manifested as clenching the teeth, pressing together the lips). Thus, the phenomenology of effort has been a topic of interest in psychology for well over a century.

The idea that animals (including people) tend to minimize effort became mainstream in the 1930s and 1940s. For example, based on a large set of animal studies, Tsai (1932) wrote:

> The *law of minimum effort* [emphasis added] states that among several alternatives of behavior leading to equivalent satisfaction of some potent organic needs, the animal, within the limits of its discriminative ability, tends to finally select that which involves the least expenditure of energy. (p. 2)

This idea was later incorporated in Hull's (1943) attempt to formulate a general, mechanistic theory of behavior. In the book "Principles of Behavior," Hull (1943) formulated what he called the *law of less work*:

> If two or more behavior sequences, each involving a different amount of energy consumption or work (W), have been equally well reinforced an equal number of times, the organism will gradually learn to choose the less laborious behavior sequence leading to the attainment of the reinforcing state of affairs. (p. 294)

This line of research established the conservation of resources as one of the basic principles of psychology: All else being equal, when given the choice, animals will minimize the expenditure of effort.

In the decades that followed, several lines of research used the resource conservation principle to explain effort, both physical and mental, in humans. We highlight three influential research traditions. First, Kahneman (1973) conceptualized mental effort as a limited resource that must be allocated strategically during cognitive processes. In line with this conceptualization, experiments showed that mental effort—operationalized as pupil dilation—scales with task difficulty (suggesting that people expend the amount of effort that is necessary to perform well, but not more; Kahneman & Beatty, 1966) and responds to reward (suggesting that people invest effort, especially when their investment is likely to pay off; Bijleveld et al., 2009; Kahneman et al., 1968; Kahneman & Peavler, 1969). The idea that mental effort is a limited resource, in turn, is a close cousin of the well-established proposal that people often apply *heuristics*—simple processes that replace complex algorithms—when they make judgments and decisions (Bless & Fiedler, 2004; Newell & Simon, 1972; Shah & Oppenheimer, 2008; Tversky & Kahneman, 1974). By using heuristics, people make often-reasonable decisions while minimizing the expenditure of mental effort (for a synthesis, see Shah & Oppenheimer, 2008).

Second, Brehm and Self (1989) extended the discourse around mental effort by introducing *motivational intensity theory*, which was designed to predict, with high precision, when people should exert effort versus when they should refrain from doing so. In essence, motivational intensity theory predicts that effort should scale with task difficulty, but only (a) as long as success is possible and only (b) as long as the expenditure of effort is justified by the value of the outcome. This model is supported by dozens of studies (e.g., Bouzidi et al., 2022; Falk et al., 2022; Richter et al., 2008; Richter & Gendolla, 2009; for narrative reviews, see Gendolla et al., 2012; Richter et al., 2016). For example, in one study (Richter et al., 2008), participants were assigned to either of four levels of difficulty (low, moderate, high, impossible) of a cognitive task. Results indicated that effort—operationalized as cardiovascular reactivity—increased with task difficulty across the first three difficulty levels; yet, people refrained from exerting effort when the task was impossible. Studies like these have led to a nuanced account of effort allocation (Richter et al., 2016), which is in line with the resource conservation principle: People only exert effort when the rewards at stake are attainable and sufficiently valuable. When they do exert effort, they expend no more effort than is demanded by the task.

Third, over the past 15 years, there has been an upsurge in research on *effort-based decision making* (for a review, see Kool & Botvinick, 2018). In this research tradition, resonating with classic research (Hull, 1943; Tsai, 1932), researchers study how people decide between two or more choice options that are associated with different amounts of required effort. Studies in this tradition yielded several insights into the nature of effort-based decision making. For example, effort-based decisions are underpinned by the dopamine pathways (Treadway, Buckholtz, et al., 2012) in combination with medial prefrontal cortex (Chong et al., 2017; Silvetti et al., 2018); effort-based decisions are biased in patients with depression and schizophrenia (Barch et al., 2014; Treadway, Bossaller, et al., 2012); effort-based decisions respond to reward (Kool et al., 2010); effort-based decisions are modulated by fatigue and sleep (Dora et al., 2022; Massar et al., 2019; Müller et al., 2021); effort-based decisions depend on environmental factors (Bijleveld & Knufinke, 2018); effort based-decisions are influenced by the order in which information about reward and effort requirements is presented (Vassena et al., 2019); and, effort-based decisions about physical versus mental effort are computationally similar (Matthews et al., 2023). Though there are some important challenges in this domain (e.g., different effort-based decision-making tasks show low intercorrelations; Mækelæ et al., 2023), all these studies consistently support the resource conservation principle. That is, these studies show that, all else being equal, people prefer choice options associated with less effort.

In sum, the assumption that people tend to minimize effort, including mental effort, is deeply ingrained in psychology. Based on this rich history of ideas, one may be tempted to conclude that mental effort should also *feel* unpleasant. This conclusion would be consistent with mainstream models of cognitive control (which assume that negative affect plays a role in triggering cognitive control, e.g., Dignath et al., 2020; Dreisbach & Fischer, 2012; cognitive control, in turn, is assumed to be effortful; Morsella et al., 2009; Silvestrini et al., 2023). Moreover, recent studies showed that mental effort is associated with tension in the *corrugator supercilii*, a facial muscle that is known to be linked to negative affect (Devine et al., 2023), and that, under some conditions, people even choose to endure physical pain rather than to expend mental effort (Vogel et al., 2020). Despite this body of research, there is also an argument to be made that effort can, in fact, be pleasant rather than aversive. We now turn to a discussion of this competing perspective.

Why Mental Effort May Not Be Inherently Aversive

There are three arguments to suggest that effort may—at least for some people in some situations—*not* feel aversive. First, a well-established line of research shows that people vary in their *need for cognition*, that is, their "tendency to engage in and enjoy effortful cognitive endeavors" (Petty et al., 2009, p. 318). For example, in one classic study, participants read a narrative text that involved several arguments. Findings indicated that, in a surprise memory test, participants high in need for cognition remembered more of these arguments, suggesting that they expended more mental effort while reading (Cacioppo et al., 1983; for further support, see Lassiter et al., 1991; Priester & Petty, 1995; Srull et al., 1985). Overall, this body of literature suggests that at least some people should enjoy mental effort. This suggests that mental effort is *not* inherently aversive.

Second, theory on *learned industriousness* assumes that, when people have repeatedly been rewarded for expending effort, effort becomes a *secondary reinforcer*. According to this assumption, it should be possible to instill a generalized willingness to exert effort in people. Several studies on mental effort, both classic and modern, support this principle (Clay et al., 2022; Eisenberger et al., 1985; Lin et al., 2024). Based on these findings, one could hypothesize that people for whom effort expenditure has become a secondary reinforcer should experience effort as less aversive—or, perhaps, even as pleasant (Eisenberger, 1992). This line of reasoning, too, suggests that mental effort may *not* be inherently aversive.

Third, in many cultures, the expenditure of mental and physical effort has a positive rather than a negative connotation. For example, a recent line of cross-cultural studies showed that the expenditure of effort tends to be perceived (by others) as a signal of moral character (Celniker et al., 2023). Similarly, several religions emphasize that "working hard" is a moral virtue (e.g., Islam: Ali & Al-Owaihan, 2008; Protestantism: van Hoorn & Maseland, 2013), potentially causing billions of people around the globe to have positive associations with the exertion of effort. Thus, also from this perspective, one could argue that effort—including mental effort—is *not* inherently aversive.

To summarize, three strands of literature suggest that, at least for some people in some situations, the exertion of mental effort may be rewarding in and by itself. So, despite the long history of research on effort, it is still controversial whether mental effort is inherently aversive. This controversy is also fueled by casual observations—for example, if mental effort is aversive, why do millions of people play chess?

Prior Reviews and Meta-Analyses on the Aversiveness of Mental Effort

Before we lay out our empirical approach, we discuss some previous reviews and meta-analyses that have influenced the debate on the aversiveness on effort.

In a narrative review, Eisenberger (1992) synthesized ∼100 studies on learned industriousness. Most of these experiments used hungry animals as subjects. To give a typical example: In one experiment (Eisenberger et al., 1979), one group of rats was repeatedly rewarded for completing a high-effort sequence of behaviors (running back and forth in an alley five times; $n = 5$). Another group of rats was rewarded for completing a low-effort sequence (running back and forth in an alley once; $n = 5$). Findings indicated that the rats that were rewarded for high effort exerted more effort on a new, unrelated task (pressing a lever). This experiment, along with many others, suggests that effort can become a *secondary reinforcer*, which implies that the aversiveness of effort can be diminished through reward learning. It is important to note that Eisenberger (1992) focused exclusively on behavior and not on subjective experiences. Thus, although extensive and thorough, this review provides no direct evidence that rewards may change the experienced unpleasantness of effort.

In a systematic review of over 100 studies on the *need for cognition*, Cacioppo et al. (1996) made an argument for the validity of a self-report instrument designed to capture this construct, the *need for cognition scale*. This self-report instrument requires people to indicate their agreement with items such as "I prefer my life to be filled with puzzles I must solve." Much more directly than Eisenberger (1992) did, Cacioppo et al. (1996) claimed that, for some people, effort should feel *pleasant* and not just *less aversive*. Specifically, people who are high in need for cognition should enjoy the expenditure of mental effort. To support their claim, Cacioppo et al. (1996) reviewed several studies that examined the association between need for cognition and ratings of task enjoyment (as measured directly after a cognitive task). Some of these studies showed the expected positive correlation, whereas others showed null results. On balance, Cacioppo et al. (1996) provided sufficient ground to formulate a more nuanced hypothesis: Mental effort may be aversive, just not for everyone (e.g., not for people higher in need for cognition, such as people with a higher education).

By contrast to these two landmark reviews that suggest that effort is *not* inherently aversive, a large set of reviews have reinvigorated the classic proposal that people tend to minimize the expenditure of effort, implying that effort is costly and aversive after all. We mention several representative reviews here: In a narrative review, Richter et al. (2016) synthesized 30 years of research on motivational intensity theory; findings generally supported the idea that people avoid investing more effort than necessary. In further narrative reviews, Shenhav et al. (2017) and Silvestrini et al. (2023) evaluated a range of putative neural and computational mechanisms that may underpin the costs of mental effort. In a systematic review, Shah and Oppenheimer (2008) catalogued various ways in which people minimize the expenditure of mental effort through heuristics. In a systematic review and meta-analysis, Torka et al. (2021) synthesized how people choose to exert effort—or choose to refrain from exerting effort—when working in teams. Finally, in several systematic reviews and meta-analyses, occupational health researchers have examined the association between effort expenditure at work and health outcomes; findings indicated that *effort–reward imbalance* (i.e., the sustained combination of high effort and low reward) predicts mental and physical illnesses in the long run (e.g., Dragano et al., 2017; Rugulies et al., 2017). Together, though this set of prior reviews generally does not directly address people's subjective experiences while exerting effort, they are consistent with the assumption that effort is inherently costly and aversive.

Thus, the controversy in the literature around the aversiveness of mental effort also emerges from our survey of prior reviews. That is, some reviews suggest that the aversiveness of effort varies between people and situations, whereas others suggest effort is inherently aversive. This tension was previously described in a narrative review by Inzlicht et al. (2018), who coined the term *effort paradox*: On the one hand, people have a clear tendency to avoid effort; on the other hand, at least some people may enjoy effort, at least sometimes. In the present research, we go beyond this previous work by offering a quantitative synthesis. We meta-analyze a substantial body of studies to better understand if—and if so, under what conditions—mental effort is experienced as unpleasant.

The Present Meta-Analysis

Our research addresses the controversy in the literature by tackling two questions: First, is mental effort generally experienced as aversive? Second, what sample and task characteristics moderate the experienced aversiveness of mental effort?

We examine these questions by meta-analyzing a substantial set of recent studies in which a sample of healthy adults carried out some cognitive task and then reported how much mental effort and how much negative affect they experienced during that task. A challenge for any meta-analysis on the link between mental effort and negative affect is that (self-reported) mental effort and negative affect are often considered to be secondary or exploratory measures. So, although both constructs are routinely included in behavioral research, they are often reported only as an afterthought (e.g., only briefly in the main text, not in the abstract or keywords). This makes it hard to systematically search for and then identify studies that include measures of both mental effort and negative affect. In the present study, we solve this challenge by focusing our search-and-inclusion strategy on a self-report instrument called the NASA Task Load Index (NASA-TLX; Hart, 2006; Hart & Staveland, 1988), which captures both mental effort and negative affect.

We capitalize on the fact that the NASA-TLX has gained traction in different scientific disciplines (e.g., ergonomics, psychology, and computer science). As such, we will examine mental effort and negative affect in a wide range of populations (e.g., American physicians, Indian fighter pilots, Japanese students) and a wide range of tasks. As the body of studies that uses the NASA-TLX is enormous, we conduct a rapid review in which we include only recent studies. Our main hypothesis is that, across tasks and populations, the feeling of effort should be associated with negative affect. Nonetheless, we also expect this association to vary across populations and tasks.

First, as reviewed above, studies on learned industriousness show that people who are repeatedly rewarded for exerting effort develop a tendency to exert greater effort in the future (Clay et al., 2022; Eisenberger et al., 1985; Lin et al., 2024; for experiments on animals, see Eisenberger et al., 1979). Speculatively, this learning process may train people to *enjoy* effort due to its prior association with reward. Based on this work, we reasoned that the link between mental effort and negative affect should depend on people's *learning history*. To test this idea, we examined four moderators: *education level* (higher educated people may have been rewarded more frequently for mental effort), *work experience* (people who worked longer in a certain job had more opportunities to get rewarded for mental effort), *skill-task fit* (people who were trained to do a specific task may have been rewarded to expend effort, especially in that task), and *continent/country* (educational systems are different across the globe; some reward effort more explicitly than others). To illustrate the latter point: Country-level differences may emerge because there are cultural differences in the value placed on hard work (Ali & Al-Owaihan, 2008; van Hoorn & Maseland, 2013), but also because some governments have been inspired by research on *growth mindset* and have therefore decided to encourage teachers to recognize and value pupils' effort (for meta-analytic reviews, see Burnette et al., 2023; Macnamara & Burgoyne, 2023; for an example of an application, see Western Cape Education Department, 2023).

Second, some tasks are, by design, more pleasant than others. Most notably, occupational psychology has a long history of studying what task parameters promote motivation and job satisfaction. Specifically, job characteristics theory starts out from the assumption that tasks (or jobs) that provide people with meaning, responsibility, and knowledge of one's own performance should be most conducive to motivation and job satisfaction (Fried & Ferris, 1987; Hackman & Oldham, 1976; Oldham & Fried, 2016). Based on this theory, we examined six design features of tasks: whether a task requires a variety of different activities and skills (*skill variety*), whether it has a clear start and finish (*task identity*), whether performance affects other people or otherwise has meaningful consequences (*task significance*), whether a task involves some degree of autonomous decision making (*control*), whether a task provides feedback on the consequences of actions (*monitoring feedback*), and whether a task provides feedback on performance (*performance feedback*). We are aware that research has examined more job characteristics beyond the ones proposed by job characteristics theory (e.g., related to social aspects of work, Morgeson & Humphrey, 2006). We chose to use the original characteristics from job characteristics theory, as these characteristics stick closely to objective task parameters, making it possible for us to code them based on the method sections of research articles. Though we based our choice of moderators on job characteristics theory, it is interesting to note that these design features resonate with the recent trend of *gamification*—that is, the addition of gamelike elements to learning platforms, aiming to enhance student engagement (Dalmina et al., 2019; Zainuddin et al., 2020). That is, designers often add gamelike elements such as "quests" and "levels" (which increase *skill variety* and *task identity*), performance feedback as compared to peers (which increases *task significance*), the possibility for the user to customize task elements (which increase *control*), progress tracking (which is a form of *monitoring feedback*), and point scoring systems (which is a form of *performance feedback*). We test the prediction that in tasks that have these design features, effort is less likely to translate into aversive feelings.

Method

Transparency and Openness

We preregistered our hypotheses, procedure, coding scheme, and analysis plan at https://osf.io/mktbr/ (David et al., 2024). Data and analysis scripts are stored there as well. All deviations from our preregistration are reported in the following section (under the Deviations From Preregistration section).

Inclusion Criteria

1. Article Type: We only included articles that were peer-reviewed, that were written in English, and that reported original data. Conference proceedings were included; dissertations were excluded.
2. Sample Size: We only included samples with a sample size greater than 10 to make sure we would spend our finite coding time on relatively robust studies. If an article reported several independent samples (e.g., two separate studies or two conditions of a between-subjects design), we coded these separately.
3. Participants: We only included samples of healthy participants who were not under the influence of a pharmacological or severe psychological treatment (e.g., induction of stress, mental fatigue, or sleep deprivation).
4. Procedure: We only included studies in which the NASA-TLX was administered directly after a single, discrete task that was described in the article. Per this criterion, we excluded studies that used the NASA-TLX to probe how people experienced a full working day, their job in general, or any activity that took longer than 1 day. If the same group of participants carried out multiple tasks consecutively, we coded all these tasks separately (if each was followed by a NASA-TLX measure).
5. Task: We only included tasks that required at least some cognitive effort (i.e., not tasks that consisted only of physical exercise).
6. Measures: We only included tasks that reported means and standard deviations (or standard errors or 95% confidence intervals) of both the *effort* and *frustration* items of the NASA-TLX.

Systematic Search Strategy

We searched for articles using the NASA-TLX via the online database Scopus. We chose to use Scopus as this database allowed us to search the main text of articles (i.e., not just the abstract, title, and keywords, which often do not mention measurement instruments by name). We considered documents published between 2015 and 2020 found using the search term: ALL ("NASA-TLX" OR "NASA Task Load Index"). This search yielded 5,061 documents.

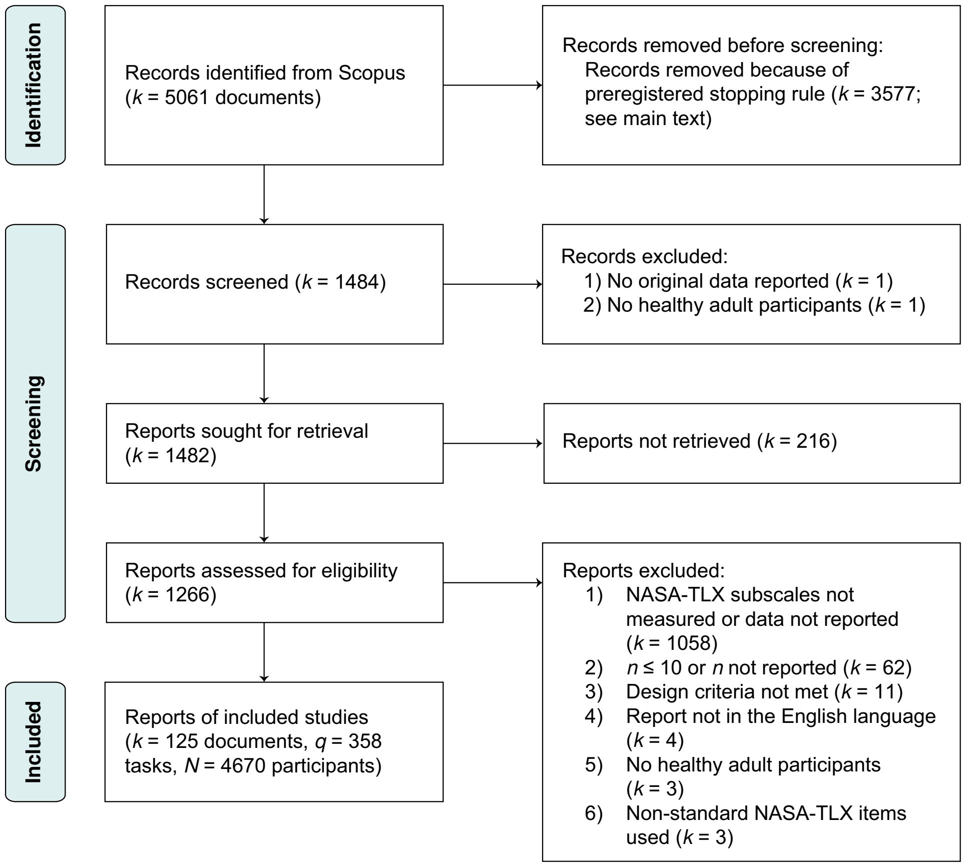
We applied our inclusion criteria in two steps (see Figure 1). First, two raters carried out an initial screening, checking (a) whether we had online access to the article through our university library and (b) whether Criterion 6 was met. Second, if the article passed this initial screening, one rater checked the remaining inclusion criteria. As our literature search yielded more articles than we could process given our resources, we preregistered an a priori stopping rule. We decided to start with the most recent article and then to code articles in reverse chronological order until either we processed all articles or it had become April 1, 2021, whichever came first. In line with this stopping rule, we processed the most recent 1,484 articles from our search. From these articles, 358 tasks (from 170 independent samples, from 125 articles; for a full list, see Supplemental Material; for a list of documents that we did not have digital access to at the time of searching, see https://osf.io/mktbr/) met our inclusion criteria. All these tasks were included in our analysis. In total, our meta-analysis was based on 9,144 NASA-TLX administrations from 4,670 unique individuals.

Calculating Effect Sizes

Our key outcome of interest was whether people experienced negative affect. We operationalized this outcome with the *frustration* item from the NASA-TLX, which is typically phrased as "How insecure, discouraged, irritated, stressed, and annoyed were you?" The main independent variable of this research is the feeling of effort, which we assessed with the *effort* item from the NASA-TLX, typically phrased as "How hard did you have to work to accomplish your level of performance?" Commonly, people respond to both items on a visual analog scale (with the anchors *very low* to *very high*) divided into 20 equal-width intervals. Based on their response, people typically receive a score between zero and 100 in increments of five.

We coded effort and frustration on the task level (i.e., each data point reflected a group of people that carried out the same task). To do so, we extracted raw means and standard deviations of the effort and frustration scales for all 358 tasks. If standard errors or confidence intervals were reported (instead of standard

Figure 1
Flow of Study Reports Into the Meta-Analytic Review



Note. NASA-TLX = NASA Task Load Index. See the online article for the color version of this figure.

deviations), we converted these into standard deviations. If means and standard deviations were not on a 1–100 scale, we converted them to that scale through linear transformation (Cohen et al., 1999). In most articles, these statistics were reported either in the main text or in tables. Yet, in some articles, these statistics were reported in plots. When we encountered a plot, we used *webplotdigitizer* (Rohatgi, 2021) to extract the relevant statistics. If it was not clear which dispersion measure was reported or which scale endpoints were used, we contacted the authors to provide clarification. If the authors did not respond, we excluded the article from our sample (k = 3). Based on the means and standard deviations for frustration, we calculated sampling variances with the *escalc* function of the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2021; Wickham et al., 2019).

To enhance the reuse potential of our dataset, we also coded the four remaining items of the NASA-TLX (*mental demand*, *physical demand*, *temporal demand*, and *performance*). As these were not our primary focus, we coded these only when they were presented in the main text or in tables (i.e., we did not digitize plots). Data from these items (q = 187 tasks; henceforth, we use the symbol q to denote sets of tasks) are available at https://osf.io/mktbr/.

Coding

Procedure

Table 1 presents an overview of all moderators. The first author coded all moderators for all tasks. To estimate the reproducibility of her coding, we created a detailed coding protocol. Using this protocol, an independent coder coded all moderators and control variables for 10 articles. To assess interrater reliability, we calculated Cohen's k for categorical variables and Pearson correlations for continuous variables. Initially, Cohen's k ranged from .36 (fair) to 1 (perfect; Landis & Koch, 1977); Pearson correlations were $\geq$.99. After discussing these results among the team, we increased the level of detail of the protocol (e.g., we added some more explanation and examples). The second coder then coded 10 new articles using the improved protocol. Cohen's k now varied between .69 (substantial) and 1 (perfect), with an average of .90 (almost perfect); Pearson correlations were $\geq$.99. Our final protocol is available at https://osf.io/mktbr/.

Learning History Moderators

Level of Education. We coded the highest level of education that participants in the sample typically received. Most samples

Table 1
Overview of Moderator Variables

Moderator	Level of coding	Description
Learning history moderators		
1. Education	Sample	Nonuniversity secondary education (32); university education[a] (199)
2. Work experience	Sample	*Continuous*, in years [M = 10.3; SD = 9.3] (58)
3. Skill–task fit	Task	Low fit (248); high fit[a] (109)
4a. Continent	Sample	Europe (158); North America (116); Asia (72)
4b. Country	Sample	United States (83); Germany (40); Canada (33)
Task design moderators		
5. Skill variety	Task	Repetitive (218); varied[a] (140)
6. Monitoring feedback	Task	Yes[a] (227); no or unknown (131)
7. Performance feedback	Task	Throughout[a] (47); No (309)
8. Control	Task	High control[a] (117); Low control (241)
9. Task significance	Task	High/medium significance[a] (170); Low significance (188)
10. Task identity	Task	High identity[a] (139); Low identity (219)
Exploratory moderators		
11. Age	Sample	*Continuous*, in years [M = 28.7; SD = 9.9] (277)
12. Gender	Sample	*Continuous*, proportion females [M = .36; SD = .22] (267)
13. Duration of task	Task	*Continuous,* in minutes [M = 35.2; SD = 54.7] (166)
14. Physical activity	Task	Light activity (112); no activity (228)
15. Group setting	Task	Individual (184); observers present (131); together with others (43)

Note. For categorical moderators, the Description column reports all moderator categories that were included in our analysis. Categories with <30 tasks were excluded from analysis; these categories do not appear in this table, but they are described in the main text. The number between round brackets is the number of tasks (denoted q in the main text) in the category. For continuous moderators (labeled *Continuous*), the Description column reports the unit of analysis, descriptive statistics (between square brackets), and the number of tasks for which we coded this moderator (between round brackets).
[a] Indicates moderator categories in which we expect a weaker (or negative) association between mental effort and negative affect.

consisted of university-educated participants (q = 199 tasks), followed by participants with some nonuniversity secondary education (q = 32), and participants who completed high school (q = 17). In some cases, studies provided no explicit information on participants' educational background. In these cases, we could sometimes infer the typical educational level from participants' occupation. If this was not feasible, we excluded the sample from analysis (q = 110). College samples were coded as university-educated. We excluded the high school category from analysis due to the small number of samples.

Work Experience. For most samples, we found no information on the mean work experience that participants had (q = 300). For the remaining samples (q = 58), we coded work experience in years.

Skill–Task Fit. We coded tasks as low fit when participants' skills—acquired either through formal education or experience—were unrelated to the task (q = 248). For example, in one study, participants (who hardly ever traveled by train) had to find their way in a virtual-reality version of the Saint-Michel Notre Dame train station in Paris (Armougum et al., 2020). This task was unrelated to participants' acquired skills, so we coded this task as low fit. Conversely, tasks were coded as high fit when participants could rely on previously acquired skills during the task (q = 109). For example, in one study that we coded as high fit, well-trained fighter pilots completed a flight simulation session (Mohanavelu et al., 2020). We could not code skill–task fit for one task (q = 1).

Continent and Country. We coded the country in which the data were collected. If no information was given, we assumed that data collection took place in the country in which most of the article's authors were based. Data came from the United States (q = 83), Germany (q = 40), Canada (q = 33), China (q = 22), United Kingdom (q = 20), Italy (q = 16), Japan (q = 13), Norway (q = 12), Netherlands (q = 11), Australia (q = 10), Denmark (q = 10), Poland (q = 9), Finland (q = 8), France (q = 8), Iran (q = 7), India (q = 7), Spain (q = 7), South Korea (q = 7), Sweden (q = 5), Indonesia (q = 5), Saudi Arabia (q = 4), Austria (q = 4), Belgium (q = 4), Malaysia (q = 3), Brazil (q = 2), Hong Kong (q = 2), Taiwan (q = 2), and Portugal (q = 1). One research team collected data in both Switzerland and Germany (q = 3).

We analyzed this moderator in two ways. First, we analyzed countries on the continent level, leading us to include Europe (q = 158), North America (q = 116), and Asia (q = 72). Second, to provide a more fine-grained analysis, we also analyzed this moderator on the country level, including the three countries with the most data points (i.e., the United States, Germany, and Canada; q ≥ 33 each).

Task Design Moderators

Skill Variety. We coded tasks as low skill variety if one component of a task was repeated several times or if a task involved short, standardized routines that were repeated continuously (q = 218). For example, in one task, participants performed ±2,900 trials of a computerized stimulus categorization task (Szychowska & Wiens, 2020). This task was coded as low variety, as all trials were very similar. We coded a task as high skill variety if a task consisted of several qualitatively different components or routines (q = 140). For example, in one study, novice surgical residents had to conduct a live robot-assisted laparoscopic surgery (Gerull et al., 2020). As this procedure required various steps (e.g., making incisions in the abdomen, inserting a medical instrument with a camera attached, controlling the camera, suturing, and communicating with staff), we coded this task as high variety.

Monitoring Feedback. We coded tasks as monitoring feedback present when participants received some sort of nonvalenced feedback from an external source on how they were doing on a task ($q = 227$). This feedback had to be related to the task goal. For example, in a driving simulation task (Zeller et al., 2020), participants drove a virtual car for 2 hr while receiving direct visual feedback on their actions (e.g., they could see their current speed on their dashboard). Similarly, in the study on surgical residents described above (Gerull et al., 2020), participants could see what they were doing on a monitor while they were handling the camera. We coded tasks as monitoring feedback absent when participants received no such external feedback ($q = 131$).

Performance Feedback. We coded tasks as performance feedback present when any sort of explicit performance feedback was given during the task ($q = 47$). Performance feedback could take different forms. In some tasks, some display element turned green or red, indicating good or poor performance, respectively (Blundell et al., 2020a). In other tasks, a virtual agent gave verbal performance feedback (e.g., "Wow, you are really good at this"; Wang, Buchweitz, et al., 2020). In yet other tasks, performance feedback was represented as the amount of money participants had accumulated (Dickinson et al., 2020). In most cases, performance feedback was absent ($q = 309$). In rare cases, participants received performance feedback, but only at the end of the task ($q = 2$). We excluded the latter level from analysis.

Control. We coded tasks as high control when participants could independently decide how to plan or carry out the task and whether there was room to take independent decisions. In other words, tasks were coded as high control if the task required (or at least allowed) participants to use their own judgment ($q = 117$). For example, in one task, participants had to drive back and forth to a certain location in a driving simulation (Milleville-Pennel & Marquez, 2020). Participants could choose between various routes (e.g., a longer route in a rural area vs. a shorter route in the city during rush hour). We coded tasks as *low control* when participants were not able to take independent decisions during the task, for example, if the task was fully scripted or if participants could only take minor decisions during the task ($q = 241$). For example, in one task, beginner golfers practiced their golf swings at a driving range 10 times (Woźniak et al., 2020). As they merely followed scripted instructions, we coded this task as low control.

Task Significance. We coded tasks as high significance when the participants' task behavior either affected other people or affected some real-world outcome ($q = 22$). All real-life performance situations fell into this category. For example, in one study, surgeons operated on real patients (Mendes, Costa, et al., 2020). We coded all simulations or training situations as medium significance ($q = 148$). For example, in one study, fighter pilots completed a simulated mission (Mohanavelu et al., 2020). We coded all other tasks as low significance ($q = 188$).

Task Identity. We coded tasks as high task identity if the task had a clear start and endpoint, that is, when participants were required to complete an entire piece of work from beginning to end ($q = 139$). For example, in one task (Ciumedean et al., 2020), participants played a game in virtual reality in which the goal was to escape from a prison. The task ended when they succeeded. We coded tasks as low task identity when tasks did not have clear-cut start and endpoints or when participants only performed part of a larger task or product ($q = 219$). For example, in one study, participants were asked to listen to and then categorize numerous short audio samples (Ishibashi et al., 2020). This task had no clear-cut start and endpoints and was therefore coded as low task identity.

Exploratory Moderators

Age. We coded the mean age of participants in the sample when this was reported ($q = 277$).

Gender. We coded the number of females and males in the sample when this was reported ($q = 267$). In our analysis, we used the proportion of females as a moderator.

Duration. We coded the duration of the task that participants had to conduct before they filled in the NASA-TLX. We only coded this moderator if the duration was explicitly stated in the article ($q = 166$).

Physical Activity. We coded the amount of physical activity that was necessary during the task in order to account for possible spill-over effects between physical and mental effort (Preston & Wegner, 2009). We categorized physical activity using a taxonomy that is commonly used in research on physical exercise (Piercy et al., 2018). We coded tasks as sedentary when people carried out the task in a sitting posture, for example, behind a desk ($q = 228$). We coded tasks as light physical activity when people carried out the task while standing and/or walking ($q = 112$). We coded tasks as moderate-to-vigorous physical activity when tasks required more intense physical exertion ($q = 18$).

Group Setting. We coded the group setting in which the task was conducted. We did this to account for effects of other people being present and for effects of working in teams (vs. individually). We coded tasks as individual–alone when participants were alone in a room (e.g., a cubicle), working on the task by themselves ($q = 184$). We coded tasks as individual–observers present when other people were physically present, but only in an observing role ($q = 131$). We coded tasks as *with others* when tasks were done together with other people ($q = 43$).

Reporting Transparency

As an index of the reporting transparency of the original studies, we coded whether the articles included data availability statements and, if they did, if the participant-level data were publicly accessible. Most of the articles ($k = 109$ articles, $q = 308$ tasks) did not include a data availability statement. Sixteen articles did include a data availability statement. Of these 16 articles ($q = 50$), eight articles ($q = 28$) mentioned that the data were available upon request. The remaining eight articles ($q = 22$) included a link to a public repository, from which the original data could be downloaded.

Analytic Strategy

We conducted the meta-analysis with the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2021; Wickham et al., 2019). We computed the best linear unbiased predictor for effort. Then, to examine Research Question 1, we adopted the best linear unbiased predictor for *effort* as the main predictor in a multilevel mixed-effects metaregression model using the raw mean (and the corresponding sampling variance) of *frustration* as our outcome measure. To examine Research Question 2, we added the moderators to this model. We only included moderator categories that consisted of ≥30 tasks. Each moderator was tested and interpreted individually. As such, each model included an intercept, the main effect of effort, the main effect of the moderator, and the Moderator × Effort interaction. To facilitate interpretation, we then estimated (and plotted) the effect of effort on frustration separately for each category of each categorical moderator. For continuous moderators, we estimated (and plotted) the effect of effort separately for several representative values of that moderator. We selected these values based on visual inspection of that moderator's distribution.

Articles often reported multiple samples; samples often carried out multiple tasks. Thus, our data had a nested structure, which we took into account by using a multilevel mixed-effects meta-analysis. First, to account for dependency among negative affect scores within each article, we included the article number in our random effects structure. This level assumes that scores within one article can be more similar than scores from other articles. Second, to account for a dependency among scores within studies, we included the study ID. This level assumes that scores within one study are more similar than scores across studies. Third, we added a unique identifier of each task within each independent sample to our random-effects structure. Thus, our random-effects structure was specified as "∼ 1 | ArticleID/SampleID/MeasureID." As an additional measure to account for the unknown structure of dependency within our data, we report cluster robust tests and confidence intervals (Pustejovsky & Tipton, 2018) applying the *clubSandwich* package (Pustejovsky, 2022).

A potential challenge for our meta-analysis stems from the fact that associations between questionnaire items may be inflated by response biases (e.g., people who have a stronger tendency to agree with items may score higher on both effort *and* negative affect, inflating the correlation between the two items; Baumgartner & Steenkamp, 2001; Bless & Fiedler, 2004). Such response biases may inflate correlations *within* individual samples. It is not a priori clear, however, if and how such response biases could have affected results from our meta-analysis. After all, we analyzed our data on an aggregate level (i.e., we analyzed sample means, not individual responses), and it is not a given that individual-level associations are mirrored by group-level associations (Kievit et al., 2013). Thus, we conducted a set of simulations to examine whether correlations within samples (which may be affected by response biases) could have plausibly biased our meta-analytic results. We described these simulations in the Supplemental Material.

Though analyses of publication bias are common in articles that report meta-analyses, we decided against presenting publication bias analyses for two reasons. First, the effect size of interest in our meta-analysis (i.e., the magnitude of the effect of effort on negative affect) was computed across studies. This approach is somewhat uncommon in that most meta-analyses focus on effect sizes that are computed within studies (e.g., standardized mean differences). As a result, common techniques for studying publication bias (e.g., funnel plots; Vevea and Hedges' weight function model) do not work in our case, as our effect size of interest necessarily includes effort as a between-studies predictor. Second, the raw NASA-TLX mean scores that we meta-analyzed were typically a byproduct of the original studies. That is, it seems unlikely that a study would be selected (or rejected) for publication based on the raw mean of any of the NASA-TLX dimensions. Thus, whereas publication bias may well exist in the body of literature we analyzed, it is unlikely that such bias affected our results (for a similar line of reasoning, see Buecker et al., 2021).

Deviations From Preregistration

In addition to examining all moderators separately, we planned to test all learning history moderators together in one model and all task design moderators in another to take into account correlations between moderators. However, during coding, it turned out that we could not code all moderators for all studies in a meaningful way, leading to missing data spread out over moderators. This especially affected the learning history moderators. If we had followed our plan (while excluding cases with missing data list-wise), we would have needed to exclude 93% of tasks in the learning history model. We felt this analysis would not be worthwhile, so we refrained from carrying it out. We did follow our plan for the task design moderators (we excluded <1% of tasks).

We planned and attempted to code two additional moderators. First, in cases where participants received performance feedback, we attempted to code the valence of this feedback (positive feedback only, negative feedback only, or both). It turned out that, in most tasks, participants received no performance feedback ($q = 309$). Only in two cases, participants received only positive feedback; in 11 cases, only negative feedback; and in 36 cases, both kinds of feedback. Therefore, we could not analyze positive versus negative feedback in a meaningful way. Second, we attempted to code whether tasks included performance-contingent incentives. It turned out that performance-contingent incentives were used only rarely ($q = 11$). Thus, we decided to drop these two moderators.

Moreover, we did not preregister the moderators that we labeled "exploratory moderators" nor the analysis in the section labeled "exploratory."

Results

Description of Included Studies

We included studies on a wide range of topics that used a variety of approaches, samples, and tasks. As the NASA-TLX is a popular tool in ergonomics, many of the included studies aimed to test how people experienced some kind of equipment or software (e.g., tools used in surgical procedures, various types of consumer electronics, flight simulator software). Several other studies were done within the standard experimental psychology paradigm. In these studies, participants carried out a computer task under controlled laboratory conditions. Studies were conducted in 27 different countries, mainly in Europe, North America, and Asia.

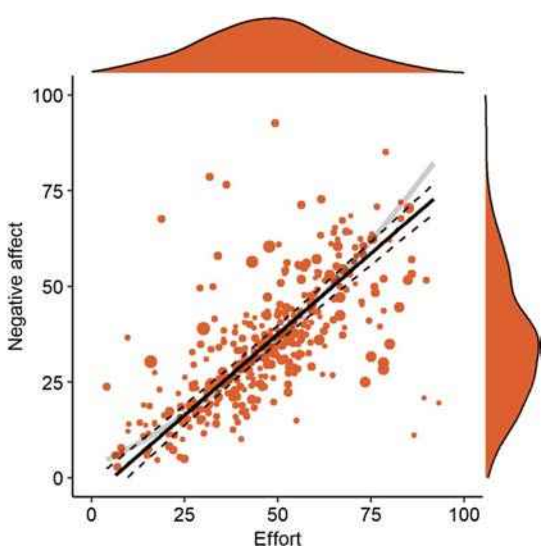
Table 1 presents descriptive statistics for all moderator variables that we coded. We also explored the associations between all moderators (reported in detail in the Supplemental Tables S1–S3). The most notable finding from this exploration was that four of the task design moderators—skill variety, control, task significance, and task identity—were correlated. That is, when a task was coded as having high skill variety, that task was more likely to be coded as high control (Cramer's $V = .66$), high task significance (Cramer's $V = .66$), and high task identity (Cramer's $V = .68$). In our moderator analysis, we dealt with these associations by testing the task design moderators individually but also all together in one model, which allowed us to explore whether the associations between these four moderators affected our main conclusions.

The Aversiveness of Effort

We first meta-analyzed all raw negative affect means (without including any predictors). On average, participants rated negative affect below the midpoint of the scale ($M = 34.6$, $SE = 1.3$, 95% CI [32.0, 37.1]). The prediction interval was [3.5, 65.6], suggesting that we can expect the true mean of negative affect in future, similar studies to lie anywhere in the bottom two thirds of the NASA-TLX scale. The Q-test was significant, $Q(357) = 53268.8$, $p < .001$, suggesting that the variability in observed negative affect was larger than would be expected based on sampling variability alone. I^2 was 98.3%, suggesting that almost all observed variance could be attributed to variance in true means rather than to sampling variance. Specifically, 52.2% of the variance was between-articles variance ($\sigma^2 = 132.3$); 8.4% was within-articles but between-samples variance ($\sigma^2 = 21.3$); 37.7% was within-samples but between-tasks variance ($\sigma^2 = 95.7$).

To test our main hypothesis, we proceeded by adding effort as a predictor to the model. As predicted, the effect of effort was significant, $\beta = 0.85$, $SE = 0.06$, 95% CI [0.73, 0.96], $t(33.2) = 14.7$, $p < .001$. The effect was large: with each point increase in effort, negative affect increased by 0.85 point on average (Figure 2). To check the robustness of this association, we explored the impact of influential cases. To that end, we excluded 16 tasks that had a Cook's distance larger than 0.4 or a $df\beta$ value outside the (−0.2, 0.2) range and reran our model, which did not substantially change our results, $\beta = 0.82$, $SE = 0.05$, 95% CI [0.72, 0.92], $t(20.2) = 17.0$, $p < .001$.

Figure 2
Plot of the Relationship Between Mean Effort and Mean Negative Affect



Note. Dots represent tasks. Dots are scaled to sample size; the smallest dot represents a sample size of 11; the largest dot represents a sample size of 114. The solid black line reflects the estimate from the multilevel metaregression model described in the main text. Black, dashed lines reflect the 95% confidence interval around that estimate. The light gray line in the background reflects the estimate from the quadratic model described in the main text. See the online article for the color version of this figure.

We finally tested whether the association between mental effort and negative affect may be better described as a curve (rather than a line). We did this to test the possibility that the link between effort and negative affect is U-shaped, such that (very) low and (very) high levels of effort both feel aversive. To that end, we ran a meta-analytic model that included both a linear and a quadratic term for effort. The quadratic model fit the data somewhat better than the linear model (Akaike information criterion$_{\text{linear}}$ = 2639.3, Akaike information criterion$_{\text{quadratic}}$ = 2633.0, Bayesian information criterion$_{\text{linear}}$ = 2658.7, Bayesian information criterion$_{\text{quadratic}}$ = 2656.3, likelihood ratio test = 8.3, $p = .004$). Nevertheless, the quadratic term was not significant, $\beta = 0.005$, $SE = .003$, $t(23.2) = 1.7$, $p = .104$. For descriptive purposes, we plotted this model's estimates in Figure 2 as a light gray line.

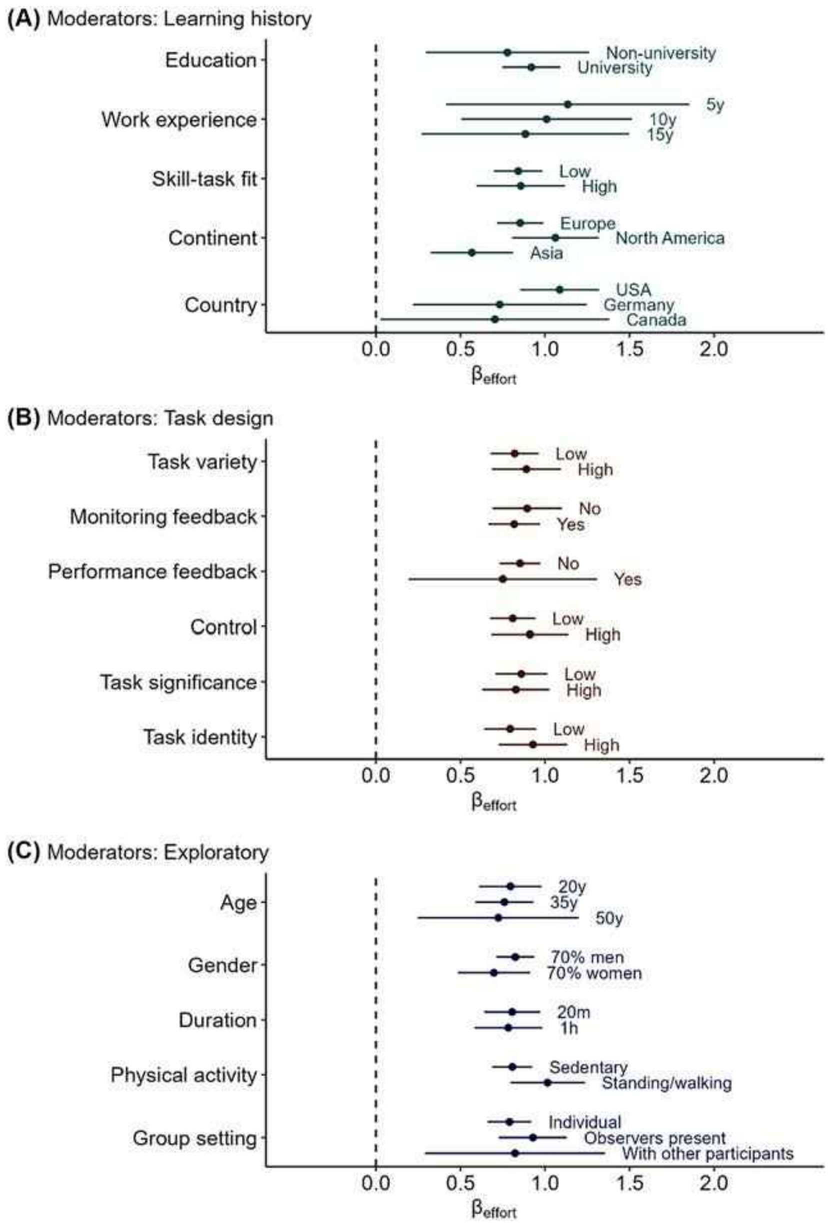
Moderator Analysis

Results from our moderator analyses are summarized in Figure 3 (and in more detail in the Supplemental Tables S4–S7). We first examined the moderators related to learning history. These moderators did not significantly interact with effort (Supplemental Table S4), except for continent. That is, we found no evidence that the association between effort and negative affect depended on education ($\beta_{\text{Effort} \times \text{Education}} = 0.14$, 95% CI [−0.33, 0.61], $p = .515$), work experience ($\beta_{\text{Effort} \times \text{Work-Experience}} = 0.03$, 95% CI [−0.11, 0.06], $p = .362$), or skill–task fit ($\beta_{\text{Effort} \times \text{Skill–Task-Fit}} = .01$, 95% CI [0.28, 0.31], $p = .918$). However, we found that effort was less strongly associated with negative affect in studies conducted in Asia, compared to studies from Europe ($\beta_{\text{Effort} \times \text{Continent}} = -0.29$, 95% CI [−0.55, −0.02], $p = .046$) and North America ($\beta_{\text{Effort} \times \text{Continent}} = -0.49$, 95% CI [−0.83, −0.16], $p = .006$; see Figure 4). Importantly, within each level of each moderator, there was a clear relationship between effort and negative affect ($\beta > .56$), including in studies conducted in Asia. So, effort felt aversive regardless of education, work experience, skill–task fit, or geographical location.

We next examined task design moderators. None of the moderators related to task characteristics significantly interacted with effort (Supplemental Table S5). That is, we found no evidence that the association between effort and negative affect depended on task variety ($\beta_{\text{Effort} \times \text{Task-Variety}} = 0.07$, 95% CI [−0.16, 0.30], $p = .539$), monitoring feedback ($\beta_{\text{Effort} \times \text{Monitoring-Feedback}} = -0.08$, 95% CI [−0.32, 0.17], $p = .530$), performance feedback ($\beta_{\text{Effort} \times \text{Performance-Feedback}} = -0.10$, 95% CI [−0.65, 0.44], $p = .658$), control ($\beta_{\text{Effort} \times \text{Control}} = 0.10$, 95% CI [−0.14, 0.34], $p = .388$), task significance ($\beta_{\text{Effort} \times \text{Task-Significance}} = -0.03$, 95% CI [−0.27, 0.21], $p = .784$), or task identity ($\beta_{\text{Effort} \times \text{Task-Identity}} = 0.13$, 95% CI [−0.11, 0.38], $p = .261$). Also here, within each level of each moderator, effort was associated with negative affect ($\beta > .75$; Figure 3). So, effort felt aversive on varied and repetitive tasks, on tasks with feedback and with no feedback, on tasks with high and low control, on tasks with high and low significance, and on tasks with high and low task identity. These results did not meaningfully change when we tested all six task design moderators together in one model (Supplemental Table S6).

Finally, we examined our exploratory moderators: age, gender, task duration, physical activity, and group setting. None of these

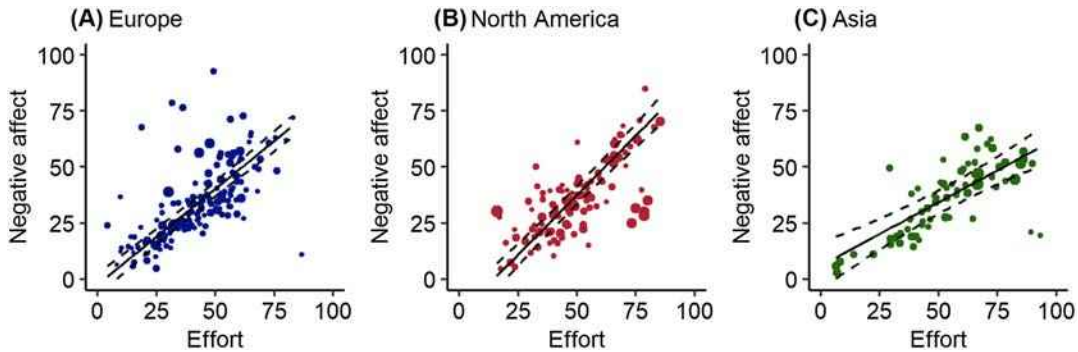
Figure 3
Results From Moderator Analysis



Note. The horizontal axis represents the metaregression parameter for the effect of effort on negative affect, separately for different moderator categories (for continuous moderators: for different representative values of the moderator). Error bars reflect 95% confidence intervals around the estimate. y = year; h = hour; m = minute. See the online article for the color version of this figure.

moderators significantly interacted with effort (Supplemental Table S7). That is, we found no evidence that the association between effort and negative affect depended on age ($\beta_{\text{Effort} \times \text{Age}}$ = 0.00, 95% CI [−0.02, 0.01], p = .730), gender ($\beta_{\text{Effort} \times \text{Gender}}$ = −0.32, 95% CI [−0.74, 0.11], p = .124), duration ($\beta_{\text{Effort} \times \text{Duration}}$ = 0.00, 95% CI [~0.00, ~0.00], p = .594), physical activity ($\beta_{\text{Effort} \times \text{Physical-Activity}}$ = 0.21, 95% CI [−0.02, 0.44], p = .069), or group setting (individual vs. observers present: $\beta_{\text{Effort} \times \text{Group-Setting}}$ = 0.14, 95% CI [−0.09, 0.37], p = .227; individual vs. together with others, $\beta_{\text{Effort} \times \text{Group-Setting}}$ = 0.03, 95% CI [−0.48, 0.54], p = .891). Here too, for each level of each moderator, effort was associated with negative affect (β > .69; Figure 3). So, effort felt aversive regardless of people's age and gender; regardless of whether tasks were short or long; regardless of whether people sat, stood, or walked; and regardless of whether people were alone or were watched or joined by other people.

Figure 4
The Association Between Effort and Negative Affect on Three Continents



Note. Solid lines represent estimates from the model described in the main text. Dashed lines indicate 95% confidence intervals. Dots are scaled to sample size. See the online article for the color version of this figure.

Robustness Analysis (Exploratory)

So far, we found that mental effort is associated with negative affect across tasks and populations. In principle, this finding is consistent with the idea that mental effort is inherently aversive. Still, it is important to note that meta-analysis is an observational technique. So, we cannot exclude the possibility that the association between effort and negative affect is spurious—that is, due to a third variable (Lipsey, 2003). Specifically, if subjective mental effort and negative affect are both triggered by some common cause (e.g., some property of the task or the population), this could explain the association, at least in part. To test this possibility, we ran a model in which we predicted negative affect from effort, as before. However, we now also included the 10 moderators for which we had >90% valid data points: skill–task fit, continent, all six task design moderators, physical activity, and group setting. We found that the association between effort and negative affect was in the same range as it was in our main analysis, $\beta = 0.88$, $SE = 0.06$, 95% CI [0.76, 1.00], $t(26.0) = 14.9$, $p < .001$ (for details, see Supplemental Table S8). Thus, the association between mental effort and negative affect cannot be explained by the possibility that both were caused by any of these 10 variables.

A further potential threat to our conclusions is that the association between effort and negative affect may be inflated by response biases (e.g., acquiescence bias). To assess whether response biases may have affected our conclusions, we conducted computer simulations. Findings suggest that, even if these response biases would have been extremely strong in the original samples, our main conclusion would not change. Details are reported in the Supplemental Material.

Discussion

This meta-analysis shows that mental effort is strongly associated with negative affect across populations and tasks. As for populations, mental effort felt aversive among university and nonuniversity educated people for experienced and inexperienced workers, regardless of whether people received specific training for the task at hand. Moreover, mental effort felt aversive in Europe, in North America, and, to a lesser extent, in Asia. As for tasks, mental effort felt aversive in varied and repetitive tasks, in tasks with and without feedback, regardless of whether people had control over how to plan or carry out the task, regardless of whether the task affected some real-life outcome, and regardless of whether the task had a clear beginning and end. Together, the link between mental effort and negative affect was ubiquitous, suggesting that mental effort is inherently aversive.

Theoretical Implications

Our study provides a new, crucial piece of support for models—for example, from psychology, economics, and cognitive neuroscience—that assume that mental effort is costly (Douglas & Shepherd, 2002; Holmstrom & Milgrom, 1994; Shenhav et al., 2017; Silvetti et al., 2018; Vassena, Deraeve, & Alexander, 2017; Vassena, Holroyd, & Alexander, 2017; Verguts et al., 2015; Vogel et al., 2020). That is, across populations and tasks, our study showed that people's subjective experiences corroborate the assumption that mental effort is perceived as a cost. At the same time, our conclusion that mental effort is inherently aversive also raises a problem (Inzlicht et al., 2018): What should we do with the idea that mental effort sometimes can feel positive? Or, to use a real-world example, if mental effort is inherently unpleasant, why do millions of people play chess?

Our preferred post hoc explanation is that high-mental-effort activities may be pleasant despite the effort, not because of it. Specifically, the utility of daily-life activities, like playing chess, can be conceptualized as a compound of costs (e.g., effort costs; opportunity costs) and benefits (e.g., monetary rewards; social rewards; mastery- and challenge-related rewards). It may often happen that this compound (sometimes called the *integrative value signal* or the *net value*; Apps & Ramnani, 2014; Vassena et al., 2015) turns out to be positive on balance. Similarly, people may *learn* that exerting mental effort—at least in the context of some specific activities—is likely to lead to reward. To return to our chess example, if the benefits of chess outweigh the costs, people may choose to play chess and even self-report that they enjoy chess. Crucially, we suggest that this does *not* imply that people enjoy the mental effort that is involved.

This explanation—which we call the *integrative value account*—can be reconciled with research on the need for cognition, which suggests that there are individual differences in the tendency to seek out and enjoy mentally effortful activities (Cacioppo et al., 1996). Still, this reconciliation requires the assumption that people higher in the need for cognition either (a) place greater value on the rewards typically associated with mentally effortful activities (e.g., feelings of mastery, self-efficacy, and competence; see Gheza et al., 2023) or (b) experience a higher probability of these rewards (e.g., because they perform relatively well at mentally effortful activities).[1] Under this assumption, people higher in need for cognition assign higher expected value to mentally effortful activities and, thus, are more likely to seek out mentally effortful activities and are more likely to receive and enjoy the rewards associated with these activities. Yet, even for people high in need for cognition, the expenditure of mental effort may feel unpleasant. This interpretation is viable given that several items of the need for cognition scale emphasize reward (e.g., "I really enjoy a task that involves coming up with new solutions to problems"; Cacioppo & Petty, 1982). Also, previous experiments suggest that people high in need for cognition are more motivated to avoid negative consequences (such as failure) during mental tasks that they expect to be difficult (Steinhart & Wyer, 2009).

To illustrate the latter point in terms of our chess example, people higher in the need for cognition may place higher value on the rewarding aspects of chess (e.g., feeling competent; winning games; accumulating rating points). Also, people higher in the need for cognition may experience such chess-related rewards more frequently because they tend to be better at chess. For these reasons, people high in need for cognition may be more likely to seek out and enjoy chess. Yet, even for them, we predict that mental effort will still feel unpleasant.

The integrative value account is consistent with a growing body of computational and neuroimaging research (Silvestrini et al., 2023). Specifically, most computational models of effort allocation include a cost function that assumes that higher effort equals higher cost. In these models, this cost is typically traded off against prospective rewards. As a result, tasks that require more effort have a lower net value. Thus, consistent with our interpretation and with classic work in psychology (Brehm & Self, 1989; Hull, 1943; Kahneman & Peavler, 1969; Richter et al., 2016), modern computational models suggest that people exert effort despite the inherent cost and only when these costs are compensated by a sufficiently valuable reward.

Though computational models of effort converge on the idea that effort is a cost, they differ in what other decision-relevant variables they include. For example, some emphasize the uncertainty of the outcome; others incorporate the volatility of the environment, the presence of punishments, or the value of obtaining information (Gottlieb et al., 2020; Shenhav et al., 2013; Silvetti et al., 2011, 2023). In this sense, *value* can be conceptualized as a multifaceted construct that is influenced by a range of extrinsic and intrinsic factors, which together can balance out the cost of exerting effort. Importantly, the precise nature of the cost of effort remains debated (Shenhav et al., 2017). One could speculate that exerting effort on one task makes it impossible to invest effort in some other task at the same time (i.e., there are *opportunity costs*; Dora et al., 2022; Kurzban et al., 2013). Relatedly, one recent model suggests that deep focus on one task reduces people's ability to multitask (Musslick & Cohen, 2021). Alternatively, one could argue that the exertion of effort carries some biological cost within the neural systems involved in cognitive control (Holroyd, 2016; Silvetti et al., 2023; Wiehler et al., 2022). Thus, although our integrative value account is broadly consistent with modern computational models, we should note that these models have not yet reached consensus on *why* mental effort is costly.

From a neural perspective, the computation of integrative value has mostly been associated with the dorsal anterior cingulate cortex (Silvetti et al., 2023; Vassena et al., 2014). Interestingly, though, cognitively demanding tasks also elicit activity in the anterior insula (Engström et al., 2015; Otto et al., 2014), a brain region that is implicated in interoception (Simmons et al., 2013). Speculatively, this system—that includes the dorsal anterior cingulate cortex and anterior insula—may integrate objective performance outcomes, environmental states, and bodily states, which ultimately results in the experienced aversiveness of effort.

Regardless of the specific neural mechanisms, the line of reasoning laid out above implies that there is a dissociation between choices (to pursue some high-effort activity) and feelings (during that high-effort activity). This putative dissociation is important, especially from a clinical perspective. Specifically, prior research suggests that people with depression and schizophrenia (vs. healthy controls) tend to avoid high-effort activities (Culbreth et al., 2018). In this area, the dominant paradigm focuses exclusively on people's choices: Researchers tend to use laboratory tasks in which participants choose between two behavioral options that vary in effort requirements and reward value (Culbreth et al., 2018; for an exception, see Brinkmann & Gendolla, 2007). However, as choices (to pursue some high-effort activity) and feelings (during that high-effort activity) may be dissociable, this approach may not tell us a lot about patients' subjective experience when they exert effort. It is an interesting avenue for future research to examine the aversiveness of mental effort directly, using subjective measures, in clinical samples.

Finally, it is important to note that people may justify their effort expenditure after the fact. That is, when people do something unpleasant to attain some goal (e.g., exert effort, endure pain, undergo humiliation), they may later infer that that goal must have been very valuable to them (Aronson & Mills, 1959). After all, why else would they have carried out the aversive action? This notion of *effort justification* has its roots in dissonance theory, and—although the exact mechanisms are still under debate (e.g., Zentall, 2010)—modern approaches (e.g., related to the *effort heuristic*, Kruger et al., 2004; Ziano et al., 2023, and the so-called *Ikea effect*, Norton et al., 2012) resonate with this idea. In any case, effort justification models reconcile two empirical facts, that is, the finding that effort is experienced as aversive (e.g., this meta-analysis; see also Devine et al., 2023; Vogel et al., 2020), and the finding that experiencing effort causes people to value the outcome of their effort more. From this perspective, too, it makes sense that, even though mental effort may be inherently aversive, people sometimes seek out mentally effortful activities.

Alternative Models

We will now consider four ways in which one can account for our findings without assuming that mental effort is inherently aversive: First, all studies that we included were instances of human-subject research. In such research, participants typically follow a series of instructions for which they get some external reward, such as a monetary payment. So, most tasks that we included in our study could still be said to be *extrinsically motivated*. Conversely, it is possible that effort can only feel pleasant during *intrinsically motivated* activities, that is, during activities that people carry out

[1] We thank Sean Devine for suggesting this interpretation.

naturally without some external goal or reward, during which people typically feel more task enjoyment. Though intuitively plausible, we do not think this account is viable. Specifically, it relies on the assumption that external rewards (e.g., money) decrease intrinsic motivation, and this assumption has often been challenged (Cerasoli et al., 2014; Gerhart & Fang, 2015). Therefore, even if all participants were paid in the included studies, there should still be meaningful variation in intrinsic motivation. Specifically, *self-determination theory* (Ryan & Deci, 2000) predicts more intrinsic motivation in tasks that allow people to make their own decisions (which should satisfy the *need for autonomy*), in which people receive feedback (which should satisfy the *need for competence*), and in which people work together with others (which should satisfy the *need for relatedness*). We found no evidence that these moderators mattered. Similarly, we found no evidence for a role of task variety and task significance, which may also be linked to intrinsic motivation (e.g., Bailey & Madden, 2016). So, our findings provide no hints that intrinsic motivation dampens the aversiveness of effort.

Second, although participants were externally rewarded for taking part in the included studies, they were usually not paid as a function of their task performance (see the Deviations From Preregistration section). Instead, in most studies, participants received a fixed monetary payment for their time (e.g., as is common in experimental psychology), or they participated as a part of their work tasks (e.g., as is common in applied research). Thus, regardless of how much mental effort was demanded by the task, participants' compensation was similar. This feature of the included studies leaves open the possibility that participants who took part in more demanding studies (and thus experienced more mental effort) were more likely to feel underrewarded. One could thus argue that the negative affect that we observed in our meta-analysis stemmed from a mental effort–reward discrepancy, not from mental effort per se. Under this account, mental effort is not inherently aversive; instead, mental effort only becomes aversive when coupled with low reward (Dragano et al., 2017; Kurzban et al., 2013; Rugulies et al., 2017; Siegrist, 1996). We cannot rule out this possibility. Future research is needed to test directly whether performance-contingent rewards can dampen the aversiveness of effort (but see Garrison et al., 2024). A general challenge for this future research is that it may be hard to empirically distinguish the (decreased) aversiveness of mental effort from the (increased) experience of reward associated with the task.

Third, prior research suggests that people often hold folk-psychological beliefs, or *lay theories*, about mental work. For example, research has documented that people often hold beliefs about intelligence ("You cannot change your intelligence"; Dweck & Yeager, 2019), concentration ("You can change how much you mind wander"; Zedelius et al., 2021), and self-regulation ("After mental activity, your energy is depleted and you must rest to get it refueled again"; Job et al., 2015). Although such lay theories may not always be accurate, they affect people's judgments in various ways (Zedelius et al., 2017). Speculatively, it could be the case that many people hold the lay theory that difficult tasks—or tasks that require mental effort—tend to be unpleasant. If true, mental effort may not be inherently aversive. Instead, its aversiveness would result from people's folk-psychological expectations. A challenge for this account, however, is that lay theories should be expected to vary between populations (Haslam, 2017), and we found little evidence for this idea (but see the following section, under the Diminished aversiveness of effort in studies from Asia section).

Fourth, in all studies that we included, mental effort was passively observed rather than manipulated. As a result, our findings cannot be used to make claims about the causal chain of processes that mediate the mental effort–negative affect relationship. As mentioned in the introduction, a well-established model is that negative affect plays a role in initiating cognitive control (Botvinick, 2007; Dignath et al., 2020; Dreisbach & Fischer, 2012; Inzlicht et al., 2015), which is widely assumed to be effortful (Silvestrini et al., 2023). Under this account, mental effort is inherently aversive. Nevertheless, we cannot exclude other mechanisms. For example, one could speculate that when people experienced negative affect during task performance, they may *infer* from this feeling that the task must have been effortful. Under this account, mental effort is not inherently aversive—rather, the feeling of mental effort results from negative affect. Experimental work that measures the physiological correlates of mental effort (e.g., using designs akin to Bogdanov et al., 2022; Devine et al., 2023) is necessary to further unravel the causal chain of processes that link effort and negative affect.

Diminished Aversiveness of Effort in Studies From Asia

The finding that mental effort felt less aversive in studies conducted in Asia is intriguing. This finding fits the general idea that the aversiveness of effort depends on people's learning history. In Asian countries, especially in China, it is relatively common for high school students to spend $\geq$60 hr per week on school work (based on a large-scale international study of 15 year olds; Organisation for Economic Co-operation and Development, 2017). Moreover, in Japan, it is common for children to spend time in *juku* (private tutoring schools) after regular school hours to cram for exams (Lowe, 2015). So, speculatively, continent-level differences in exposure to mental effort in educational settings may explain our finding. Yet, we cannot exclude two further explanations. First, it may be the case that words like "effort" and "annoyed" have different connotations in different languages (Steele, 2020), leading people to respond differently to translated items. Second, it may be the case that researchers from different continents favor tasks with different characteristics. This explanation seems somewhat unlikely, though, as we coded six core task characteristics, which did not significantly affect the aversiveness of mental effort.

Strengths and Limitations

We highlight that we analyzed sample means rather than individual responses. Thus, our findings showed that when *a sample* of participants felt more effort on average, *that sample* also tended to feel more negative affect on average. A clear advantage of this approach is that our findings cannot be explained by individual-level response biases (see simulations in Supplemental Material). However, a drawback of this approach is that we need to be cautious to generalize our conclusions to the individual level. That is, our meta-analysis does not show that when one person experiences more effort, that person will also experience more negative affect (some previous studies do show this, e.g., Hart & Staveland, 1988).

A potential criticism of our study may stem from the fact that, in ergonomics, the NASA-TLX is often used to measure a single construct called "workload." "Workload" is typically computed as the sum (or mean) of effort, negative affect, perceived performance, and three items that measure perceived task demands. If one accepts "workload" as a meaningful construct, one could argue that effort and negative affect are simply part of the same construct—and thus, that the association between effort and negative affect is trivial. To argue against this line of reasoning, we note that the compound concept of "workload" is controversial, even in ergonomics (e.g., the concept of "workload" lacks precision and explains behavior only superficially; Dekker & Hollnagel, 2004). Moreover, we note that mental effort and negative are clearly dissociable, in that people can readily experience negative affect without experiencing mental effort (e.g., when they feel bored, sad, or lethargic; see Russell & Carroll, 1999). Indeed, from a psychological perspective, it makes little sense to lump effort and negative affect together in one measure. In our view, these constructs are distinct, and it is worthwhile to study their association (Inzlicht et al., 2018).

As the number of studies that used the NASA-TLX is large, we decided to start coding the most recent studies and then work backward in time until our resources ran out, including all articles to which we had digital access. Although this strategy was successful—that is, it led us to include a substantial number of studies with substantial diversity—we should note that (a) we ended up including only studies that were published in 2019 and 2020, and (b) there were several articles that we discarded because we could not access them (see the Method section). Also, we note that we only included articles written in English. Altogether, we cannot exclude the possibility that our search-and-inclusion strategy led to biases (e.g., because more recent studies systematically differ from older studies; because the studies we could access systematically differ from the studies we could not access; because studies reported in English systematically differ from studies reported in other languages). Thus, future rapid reviews on the NASA-TLX could consider different sampling strategies (e.g., sampling randomly from the literature to also include older studies). Similarly, we cannot be certain that the tasks we included are fully representative of tasks that are used in the real world, nor can we be certain that we captured all relevant moderators.

The availability of the primary data from the studies we included was suboptimal. Specifically, only eight out of 125 articles included a link to a public repository from which the data could be downloaded. Thus, there is ample room for improvement regarding transparency and openness in this research area. Although we were able to extract our measures of interest from the articles that did not publicly share their data, it would greatly benefit future meta-analyses if researchers who use the NASA-TLX would make their participant-level data more accessible. This would allow the research community to study the link between mental effort and negative affect in a more fine-grained manner. For example, this could be done by examining whether the aversiveness of effort is subject to circadian or seasonal variations or whether it varies with personality traits (such as conscientiousness; Bates, 2024).

A final limitation of our study is that effort was measured solely using self-report. Although physiological effort (during mental activity) and the feeling of effort are correlated (Bijleveld, 2018), further research is necessary to examine the relationship between physiological and behavioral measures of effort on the one hand and negative affect on the other. Also, the NASA-TLX has limitations. That is, although negative emotions are heterogeneous and separable (Walters & Simons, 2022), the NASA-TLX lumps together several types of task-related negative affect (feeling insecure, discouraged, irritated, stressed, and annoyed) in one dimension. Thus, one could argue that the NASA-TLX is too coarse. Conversely, one could also argue the opposite, that is, that the NASA-TLX is too narrow. After all, the NASA-TLX does not capture all forms of task-related negative affect (e.g., it does not capture boredom and fatigue). We were aware of these limitations from the start of this study. Nevertheless, despite its shortcomings, the NASA-TLX is so widely used that it has allowed us to assess the relationship between effort and negative affect on a large, near-global scale. This would not have been feasible with any other meta-analytic approach. We thus feel that the advantages of the NASA-TLX outweigh the disadvantages, but we note that future research needs to look at the aversiveness of mental effort in complementary ways as well—in particular, by using different strategies to measure effort and negative affect.

Conclusion

The main finding from this meta-analysis is that mental effort is strongly associated with negative affect. We found this association in all types of tasks that we studied, including tasks that have motivating features (e.g., tasks in which people have autonomy, tasks in which people receive feedback, and tasks in which performance has real-world consequences). Furthermore, we found this association in all types of populations that we studied, including populations in which mental effort likely had been rewarded in the past (e.g., experienced professionals, university-educated people). In sum, the association between mental effort and negative affect proved to be robust and generalizable across a wide range of tasks and populations.

This meta-analysis provides a new, central piece of evidence for models that assume that mental effort is costly. This assumption is made in various fields (e.g., psychology, neuroscience, and economics), but it is not often tested. Our meta-analysis supports the assumption that mental effort is costly by showing that mental effort is consistently accompanied by negative affect.

This meta-analysis speaks to the open challenge that is commonly referred to as the effort paradox (Inzlicht et al., 2018): If mental effort is consistently unpleasant, why do people still voluntarily pursue mentally effortful activities? For example, why do millions of people play chess? Our results suggest that there may be a dissociation between choices and feelings: When people choose to pursue mentally effortful activities, this should not be taken as an indication that they enjoy mental effort per se. Perhaps people choose mentally effortful activities despite the effort, not because of it.

References

References marked with an asterisk indicate documents included in our meta-analysis.

*Abd Rahman, N. I., Md Dawal, S. Z., & Yusoff, N. (2020). Driving mental workload and performance of ageing drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *69*, 265–285. https://doi.org/10.1016/j.trf.2020.01.019

Ach, N. (2006). *Über den Willen* (T. Herz, Trans.) [On Volition]. Verlag von Quelle und Meyer. Retrieved January 16, 2024, from https://www.uni-konstanz.de/kogpsych/ach.htm (Original work published 1910)

*Alagi, H., Navarro, S. E., Hergenhan, J., Music, S., & Hein, B. (2020). *Teleoperation with tactile feedback based on a capacitive proximity sensor array* [Conference session]. 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Dubrovnik, Croatia. https://doi.org/10.1109/I2MTC43012.2020.9128701

Ali, A. J., & Al-Owaihan, A. (2008). Islamic work ethic: A critical review. *Cross Cultural Management*, *15*(1), 5–19. https://doi.org/10.1108/13527600810848791

Apps, M. A. J., & Ramnani, N. (2014). The anterior cingulate gyrus signals the net value of others' rewards. *The Journal of Neuroscience*, *34*(18), 6190–6200. https://doi.org/10.1523/JNEUROSCI.2701-13.2014

*Armougum, A., Gaston-Bellegarde, A., Marle, C. J.-L., & Piolino, P. (2020). Expertise reversal effect: Cost of generating new schemas. *Computers in Human Behavior*, *111*, Article 106406. https://doi.org/10.1016/j.chb.2020.106406

*Aromaa, S., Väätänen, A., Aaltonen, I., Goriachev, V., Helin, K., & Karjalainen, J. (2020). Awareness of the real-world environment when using augmented reality head-mounted display. *Applied Ergonomics*, *88*, Article 103145. https://doi.org/10.1016/j.apergo.2020.103145

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*(2), 177–181. https://doi.org/10.1037/h0047195

*Arsintescu, L., Chachad, R., Gregory, K. B., Mulligan, J. B., & Flynn-Evans, E. E. (2020). The relationship between workload, performance and fatigue in a short-haul airline. *Chronobiology International*, *37*(9–10), 1492–1494. https://doi.org/10.1080/07420528.2020.1804924

*Ataya, A., Kim, W., Elsharkawy, A., & Kim, S. (2020). Gaze-head input: Examining potential interaction with immediate experience sampling in an autonomous vehicle. *Applied Sciences*, *10*(24), Article 9011. https://doi.org/10.3390/app10249011

*Badesa, F. J., Diez, J. A., Catalan, J. M., Trigili, E., Cordella, F., Nann, M., Crea, S., Soekadar, S. R., Zollo, L., Vitiello, N., & Garcia-Aracil, N. (2019). Physiological responses during hybrid BCNI control of an upper-limb exoskeleton. *Sensors*, *19*(22), Article 4931. https://doi.org/10.3390/s19224931

Bailey, C., & Madden, A. (2016). What makes work meaningful—Or meaningless. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/what-makes-work-meaningful-or-meaningless/

Barch, D. M., Treadway, M. T., & Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: Reduced effort allocation predicts amotivation and functional impairment. *Journal of Abnormal Psychology*, *123*(2), 387–397. https://doi.org/10.1037/a0036299

Bates, T. C. (2024). Conscientiousness makes effort less frustrating. *Personality and Individual Differences*, *226*, Article 112690. https://doi.org/10.1016/j.paid.2024.112690

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, *38*(2), 143–156. https://doi.org/10.1509/jmkr.38.2.143.18840

*Beege, M., Nebel, S., Schneider, S., & Rey, G. D. (2021). The effect of signaling in dependence on the extraneous cognitive load in learning environments. *Cognitive Processing*, *22*(2), 209–225. https://doi.org/10.1007/s10339-020-01002-5

*Berg, M., Bayazit, D., Mathew, R., Rotter-Aboyoun, A., Pavlick, E., & Tellex, S. (2020). *Grounding language to landmarks in arbitrary outdoor environments* [Conference session]. 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France. https://doi.org/10.1109/ICRA40945.2020.9197068

*Bier, L., Emele, M., Gut, K., Kulenovic, J., Rzany, D., Peter, M., & Abendroth, B. (2019). Preventing the risks of monotony related fatigue while driving through gamification. *European Transport Research Review*, *11*(1), Article 44. https://doi.org/10.1186/s12544-019-0382-4

Bijleveld, E. (2018). The feeling of effort during mental activity. *Consciousness and Cognition*, *63*, 218–227. https://doi.org/10.1016/j.concog.2018.05.013

Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opener: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. *Psychological Science*, *20*(11), 1313–1315. https://doi.org/10.1111/j.1467-9280.2009.02443.x

Bijleveld, E., & Knufinke, M. (2018). Exposure to bright light biases effort-based decisions. *Behavioral Neuroscience*, *132*(3), 183–193. https://doi.org/10.1037/bne0000244

Bless, H., & Fiedler, K. (2004). *Social cognition: How individuals construct social reality*. Psychology Press. https://doi.org/10.4324/9781315784731

*Blundell, J., Scott, S., Harris, D., Huddlestone, J., & Richards, D. (2020a). With flying colours: Pilot performance with colour-coded head-up flight symbology. *Displays*, *61*, Article 101932. https://doi.org/10.1016/j.displa.2019.101932

*Blundell, J., Scott, S., Harris, D., Huddlestone, J., & Richards, D. (2020b). Workload benefits of colour coded head-up flight symbology during high workload flight. *Displays*, *65*, Article 101973. https://doi.org/10.1016/j.displa.2020.101973

*Bogaard, T., Wielemaker, J., Hollink, L., Hardman, L., & van Ossenbruggen, J. (2020). Understanding user behavior in digital libraries using the MAGUS session visualization tool. In M. Hall, T. Merčun, T. Risse, & F. Duchateau (Eds.), *Digital libraries for open knowledge* (Vol. 12246, pp. 171–184). Springer International Publishing. https://doi.org/10.1007/978-3-030-54956-5_13

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex*, *32*(19), 4255–4270. https://doi.org/10.1093/cercor/bhab480

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 356–366. https://doi.org/10.3758/CABN.7.4.356

Bouzidi, Y. S., Falk, J. R., Chanal, J., & Gendolla, G. H. E. (2022). Choosing task characteristics oneself justifies effort: A study on cardiac response and the critical role of task difficulty. *Motivation Science*, *8*(3), 230–238. https://doi.org/10.1037/mot0000269

*Braarud, P. Ø. (2020). An efficient screening technique for acceptable mental workload based on the NASA Task Load Index—Development and application to control room validation. *International Journal of Industrial Ergonomics*, *76*, Article 102904. https://doi.org/10.1016/j.ergon.2019.102904

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*(1), 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545

Brinkmann, K., & Gendolla, G. H. E. (2007). Dysphoria and mobilization of mental effort: Effects on cardiovascular reactivity. *Motivation and Emotion*, *31*(1), 71–82. https://doi.org/10.1007/s11031-007-9054-0

Buecker, S., Mund, M., Chwastek, S., Sostmann, M., & Luhmann, M. (2021). Is loneliness in emerging adults increasing over time? A preregistered cross-temporal meta-analysis and systematic review. *Psychological Bulletin*, *147*(8), 787–805. https://doi.org/10.1037/bul0000332

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M., & Simon, S. (2023). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin*, *149*(3–4), 174–205. https://doi.org/10.1037/bul0000368

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197–253. https://doi.org/10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, *45*(4), 805–818. https://doi.org/10.1037/0022-3514.45.4.805

*Caggianese, G., Capece, N., Erra, U., Gallo, L., & Rinaldi, M. (2020). Freehand-steering locomotion techniques for immersive virtual environments: A comparative evaluation. *International Journal of Human–Computer Interaction*, *36*(18), 1734–1755. https://doi.org/10.1080/10447318.2020.1785151

*Cai, S., Ke, P., Narumi, T., & Zhu, K. (2020). *ThermAirGlove: A pneumatic glove for thermal perception and material identification in virtual reality* [Conference session]. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, United States. https://doi.org/10.1109/VR46266.2020.00044

*Calvi, A., D'Amico, F., Ciampoli, L. B., & Ferrante, C. (2020). Evaluation of driving performance after a transition from automated to manual control: A driving simulator study. *Transportation Research Procedia*, *45*, 755–762. https://doi.org/10.1016/j.trpro.2020.02.101

*Caria, M., Todde, G., Sara, G., Piras, M., & Pazzona, A. (2020). Performance and usability of smartglasses for augmented reality in precision livestock farming operations. *Applied Sciences*, *10*(7), Article 2318. https://doi.org/10.3390/app10072318

*Cecotti, H., Day-Scott, Z., Huisinga, L., & Gordo-Pelaez, L. (2020). *Virtual reality for immersive learning in art history* [Conference session]. 2020 6th International Conference of the Immersive Learning Research Network (iLRN), San Luis Obispo, CA, United States. https://doi.org/10.23919/iLRN47897.2020.9155108

Celniker, J. B., Gregory, A., Koo, H. J., Piff, P. K., Ditto, P. H., & Shariff, A. F. (2023). The moralization of effort. *Journal of Experimental Psychology: General*, *152*(1), 60–79. https://doi.org/10.1037/xge0001259

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*(4), 980–1008. https://doi.org/10.1037/a0035661

*Chacko, S. M., Granado, A., & Kapila, V. (2020). An augmented reality framework for robotic tool-path teaching. *Procedia CIRP*, *93*, 1218–1223. https://doi.org/10.1016/j.procir.2020.03.143

*Chan, W. P., Sakr, M., Quintero, C. P., Croft, E., & Van der Loos, H. F. M. (2020). *Towards a multimodal system combining augmented reality and electromyography for robot trajectory programming and execution* [Conference session]. 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Naples, Italy. https://doi.org/10.1109/RO-MAN47096.2020.9223526

*Chen, Y., Katsuragawa, K., & Lank, E. (2020). Understanding viewport- and world-based pointing with everyday smart devices in immersive augmented reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376592

Chong, T. T.-J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLOS Biology*, *15*(2), Article e1002598. https://doi.org/10.1371/journal.pbio.1002598

*Ciumedean, C.-B., Patras, C., Cibulskis, M., Váradi, N., & Nilsson, N. C. (2020). *Mission impossible spaces: Using challenge-based distractors to reduce noticeability of self-overlapping virtual architecture* [Conference session]. Symposium on Spatial User Interaction, Association for Computing Machinery. https://doi.org/10.1145/3385959.3418453

Clay, G., Mlynski, C., Korb, F. M., Goschke, T., & Job, V. (2022). Rewarding cognitive effort increases the intrinsic value of mental labor. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(5), Article e2111785119. https://doi.org/10.1073/pnas.2111785119

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*(3), 315–346. https://doi.org/10.1207/S15327906MBR3403_2

Culbreth, A. J., Moran, E. K., & Barch, D. M. (2018). Effort-cost decision-making in psychosis and depression: Could a similar behavioral deficit arise from disparate psychological and neural mechanisms? *Psychological Medicine*, *48*(6), 889–904. https://doi.org/10.1017/S0033291717002525

*Czarnek, G., Strojny, P., Strojny, A., & Richter, M. (2020). Assessing engagement during rescue operation simulated in virtual reality: A psychophysiological study. *International Journal of Human–Computer Interaction*, *36*(5), 464–476. https://doi.org/10.1080/10447318.2019.1655905

Dalmina, L., Barbosa, J. L. V., & Vianna, H. D. (2019). A systematic mapping study of gamification models oriented to motivational characteristics. *Behaviour & Information Technology*, *38*(11), 1167–1184. https://doi.org/10.1080/0144929X.2019.1576768

*Datta, R. R., Schönhage, S., Dratsch, T., Toader, J., Müller, D. T., Wahba, R., Kleinert, R., Thomas, M., Dieplinger, G., Stippel, D. L., Bruns, C. J., & Fuchs, H. F. (2021). Learning curve of surgical novices using the single-port platform SymphonX: Minimizing OR trauma to only one 15-mm incision. *Surgical Endoscopy*, *35*(9), 5338–5351. https://doi.org/10.1007/s00464-020-07998-3

David, L., Vassena, E., & Bijleveld, E. (2024, May 23). *The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect* (pre-registration, data, and code). PsyArXiv. https://osf.io/mktbr

*de Melo, C. M., Kim, K., Norouzi, N., Bruder, G., & Welch, G. (2020). Reducing cognitive load and improving warfighter problem solving with intelligent virtual assistants. *Frontiers in Psychology*, *11*, Article 554706. https://doi.org/10.3389/fpsyg.2020.554706

Dekker, S., & Hollnagel, E. (2004). Human factors and folk models. *Cognition Technology and Work*, *6*(2), 79–86. https://doi.org/10.1007/s10111-003-0136-9

*DeLucia, P. R., & Greenlee, E. T. (2022). Tactile vigilance is stressful and demanding. *Human Factors*, *64*(4), 732–745. https://doi.org/10.1177/0018720820965294

*Desolda, G., Matera, M., & Lanzilotti, R. (2020). Metamorphic data sources: A user-centric paradigm to consume linked data in interactive workspaces. *Future Generation Computer Systems*, *102*, 992–1015. https://doi.org/10.1016/j.future.2019.09.032

Devine, S., Vassena, E., & Otto, A. R. (2023). More than a feeling: Physiological measures of affect index the integration of effort costs and rewards during anticipatory effort evaluation. *Cognitive, Affective & Behavioral Neuroscience*, *23*(4), 1129–1140. https://doi.org/10.3758/s13415-023-01095-3

*Dickinson, P., Gerling, K., Wilson, L., & Parke, A. (2020). Virtual reality as a platform for research in gambling behaviour. *Computers in Human Behavior*, *107*, Article 106293. https://doi.org/10.1016/j.chb.2020.106293

Dignath, D., Eder, A. B., Steinhauser, M., & Kiesel, A. (2020). Conflict monitoring and the affective-signaling hypothesis—An integrative review. *Psychonomic Bulletin & Review*, *27*(2), 193–216. https://doi.org/10.3758/s13423-019-01668-9

*Ding, Y., Cao, Y., Duffy, V. G., Wang, Y., & Zhang, X. (2020). Measurement and identification of mental workload during simulated computer tasks with multimodal methods and machine learning. *Ergonomics*, *63*(7), 896–908. https://doi.org/10.1080/00140139.2020.1759699

*Do, S., & Lee, B. (2020). *Improving reliability of virtual collision responses: A cue integration technique* [Conference session]. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery. https://doi.org/10.1145/3313831.3376819

*Dodwell, A., & Trick, L. M. (2020). The effects of secondary tasks that involve listening and speaking on young adult drivers with traits associated with autism spectrum disorders: A pilot study with driving simulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *69*, 120–134. https://doi.org/10.1016/j.trf.2019.12.011

Dora, J., van Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J., & Bijleveld, E. (2022). The effect of opportunity costs on mental fatigue in labor/leisure trade-offs. *Journal of Experimental Psychology: General*, *151*(3), 695–710. https://doi.org/10.1037/xge0001095

Douglas, E. J., & Shepherd, D. A. (2002). Self-employment as a career choice: Attitudes, entrepreneurial intentions, and utility maximization. *Entrepreneurship Theory and Practice*, *26*(3), 81–90. https://doi.org/10.1177/104225870202600305

Dragano, N., Siegrist, J., Nyberg, S. T., Lunau, T., Fransson, E. I., Alfredsson, L., Bjorner, J. B., Borritz, M., Burr, H., Erbel, R., Fahlén, G., Goldberg, M., Hamer, M., Heikkilä, K., Jöckel, K.-H., Knutsson, A., Madsen, I. E. H., Nielsen, M. L., Nordin, M., … the IPD-Work consortium. (2017). Effort–reward imbalance at work and incident coronary heart disease. *Epidemiology*, *28*(4), 619–626. https://doi.org/10.1097/EDE.0000000000000666

Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, *78*(2), 94–98. https://doi.org/10.1016/j.bandc.2011.12.003

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, *14*(3), 481–496. https://doi.org/10.1177/1745691618804166

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*(2), 248–267. https://doi.org/10.1037/0033-295X.99.2.248

Eisenberger, R., Carlson, J., Guile, M., & Shapiro, N. (1979). Transfer of effort across behaviors. *Learning and Motivation*, *10*(2), 178–197. https://doi.org/10.1016/0023-9690(79)90043-2

Eisenberger, R., Mitchell, M., & Masterson, F. A. (1985). Effort training increases generalized self-control. *Journal of Personality and Social Psychology*, *49*(5), 1294–1301. https://doi.org/10.1037/0022-3514.49.5.1294

Engström, M., Karlsson, T., Landtblom, A.-M., & Craig, A. D. (2015). Evidence of conjoint activation of the anterior insular and cingulate cortices during effortful tasks. *Frontiers in Human Neuroscience*, *8*, Article 1071. https://doi.org/10.3389/fnhum.2014.01071

*Ezcurdia, I., Arregui, A., Ardaiz, O., Ortiz, A., & Marzo, A. (2020). Content adaptation and depth perception in an affordable multi-view display. *Applied Sciences*, *10*(20), Article 7357. https://doi.org/10.3390/app10207357

Falk, J. R., Gollwitzer, P. M., Oettingen, G., & Gendolla, G. H. E. (2022). Incidental affective influences on effort-related cardiac response: The critical role of choosing task characteristics. *International Journal of Psychophysiology*, *177*, 76–82. https://doi.org/10.1016/j.ijpsycho.2022.04.010

*Filho, J. A. W., Stuerzlinger, W., & Nedel, L. (2020). Evaluating an immersive space–time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 514–524. https://doi.org/10.1109/TVCG.2019.2934415

Fried, Y., & Ferris, G. R. (1987). The validity of the job characteristics model: A review and meta-analysis. *Personnel Psychology*, *40*(2), 287–322. https://doi.org/10.1111/j.1744-6570.1987.tb00605.x

*Ganni, S., Li, M., Botden, S. M. B. I., Nayak, S. R., Ganni, B. R., Rutkowski, A.-F., Goossens, R. H. M., & Jakimowicz, J. (2020). Virtual operating room simulation setup (vorss) for procedural training in minimally invasive surgery—A pilot study. *Indian Journal of Surgery*, *82*(5), 810–816. https://doi.org/10.1007/s12262-020-02131-z

Garrison, K. E., Wahlers, J. N., & Schmeichel, B. J. (2024). *The influence of performance incentives on the subjective experience of mental effort* [Working paper]. University of Alabama.

Gendolla, G. H. E., Wright, R. A., & Richter, M. (2012). Effort intensity: Some insights from the cardiovascular system. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 420–438). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0024

Gerhart, B., & Fang, M. (2015). Pay, intrinsic motivation, extrinsic motivation, performance, and creativity in the workplace: Revisiting long-held beliefs. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 489–521. https://doi.org/10.1146/annurev-orgpsych-032414-111418

*Gerull, W., Zihni, A., & Awad, M. (2020). Operative performance outcomes of a simulator-based robotic surgical skills curriculum. *Surgical Endoscopy*, *34*(10), 4543–4548. https://doi.org/10.1007/s00464-019-07243-6

*Ghaleb, A. M., Khalaf, T. M., Ramadan, M. Z., Ragab, A. E., & Badwelan, A. (2020). Effect of cycling on a stationary bike while performing assembly tasks on human physiology and performance parameters. *International Journal of Environmental Research and Public Health*, *17*(5), Article 1761. https://doi.org/10.3390/ijerph17051761

Gheza, D., Kool, W., & Pourtois, G. (2023). Need for cognition moderates the relief of avoiding cognitive effort. *PLOS ONE*, *18*(11), Article e0287954. https://doi.org/10.1371/journal.pone.0287954

*Gilardi, F., De Falco, F., Casasanta, D., Andellini, M., Gazzellini, S., Petrarca, M., Morocutti, A., Lettori, D., Ritrovato, M., Castelli, E., Raponi, M., Magnavita, N., & Zaffina, S. (2020). Robotic technology in pediatric neurorehabilitation: A pilot study of human factors in an Italian pediatric hospital. *International Journal of Environmental Research and Public Health*, *17*(10), Article 3503. https://doi.org/10.3390/ijerph17103503

*Glimne, S., Brautaset, R., & Österman, C. (2020). Visual fatigue during control room work in process industries. *Work*, *65*(4), 903–914. https://doi.org/10.3233/WOR-203141

*Gold, B., & Windscheid, J. (2020). Observing 360-degree classroom videos—Effects of video type on presence, emotions, workload, classroom observations, and ratings of teaching quality. *Computers & Education*, *156*, Article 103960. https://doi.org/10.1016/j.compedu.2020.103960

*Golmohammadi, R., Darvishi, E., Faradmal, J., Poorolajal, J., & Aliabadi, M. (2020). Attention and short-term memory during occupational noise exposure considering task difficulty. *Applied Acoustics*, *158*, Article 107065. https://doi.org/10.1016/j.apacoust.2019.107065

Gottlieb, J., Cohanpour, M., Li, Y., Singletary, N., & Zabeh, E. (2020). Curiosity, information demand and attentional priority. *Current Opinion in Behavioral Sciences*, *35*, 83–91. https://doi.org/10.1016/j.cobeha.2020.07.016

*Greenhalgh, M., Blaauw, E., Deepak, N., St. Laurent, C. O. L. M., Cooper, R., Bendixen, R., Koontz, A. M., & Cooper, R. A. (2022). Usability and task load comparison between a robotic assisted transfer device and a mechanical floor lift during caregiver assisted transfers on a care recipient. *Disability and Rehabilitation. Assistive Technology*, *17*(7), 833–839. https://doi.org/10.1080/17483107.2020.1818137

*Greenlee, E. T., Lui, T. G., & Maw, E. L. (2021). Is physiobehavioral monitoring nonintrusive? An examination of transcranial doppler sonography in a vigilance task. *Human Factors*, *63*(7), 1256–1270. https://doi.org/10.1177/0018720820920118

*Gross, I. T., Whitfill, T., Redmond, B., Couturier, K., Bhatnagar, A., Joseph, M., Joseph, D., Ray, J., Wagner, M., & Auerbach, M. (2020). Comparison of two telemedicine delivery modes for neonatal resuscitation support: A simulation-based randomized trial. *Neonatology*, *117*(2), 159–166. https://doi.org/10.1159/000504853

*Guo, W., & Kim, J. H. (2020). How augmented reality influences student workload in engineering education. In C. Stephanidis, D. Harris, W.-C. Li, D. D. Schmorrow, C. M. Fidopiastis, P. Zaphiris, A. Ioannou, X. Fang, R. A. Sottilare, & J. Schwarz (Eds.), *HCI international 2020—Late breaking papers: Cognition, learning and games* (Vol. 12425, pp. 388–396). Springer International Publishing. https://doi.org/10.1007/978-3-030-60128-7_29

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*(2), 250–279. https://doi.org/10.1016/0030-5073(76)90016-7

*Hafiz, P., & Bardram, J. E. (2020). The ubiquitous cognitive assessment tool for smartwatches: Design, implementation, and evaluation study. *JMIR mHealth and uHealth*, *8*(6), Article e17506. https://doi.org/10.2196/17506

*Hangli, G., Hamada, T., Sumitomo, T., & Koshizuka, N. (2020). Intellevator: An intelligent elevator system proactive in traffic control for time-efficiency improvement. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 35535–35545. https://doi.org/10.1109/ACCESS.2020.2975020

Hart, S. G. (2006). Nasa-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of Nasa-Tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Haslam, N. (2017). The origins of lay theories: The case of essentialist beliefs. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health* (pp. 3–16). Springer. https://doi.org/10.1007/978-3-319-57306-9_1

Holmstrom, B., & Milgrom, P. (1994). The firm as an incentive system. *The American Economic Review*, *84*(4), 972–991. https://www.jstor.org/stable/2118041

Holroyd, C. B. (2016). The waste disposal problem of effortful control. In T. S. Braver (Ed.), *Motivation and cognitive control* (pp. 235–260). Routledge/Taylor & Francis Group.

*Hoppe, A. H., Marek, F., Camp, F., & Stiefelhagen, R. (2020). Extending movable surfaces with touch interaction using the VirtualTablet: An extended view. *Advances in Science. Technology and Engineering Systems Journal*, *5*(2), 328–337. https://doi.org/10.25046/aj050243

*Hsia, L., & Hwang, G. (2020). From reflective thinking to learning engagement awareness: A reflective thinking promoting approach to improve students' dance performance, self-efficacy and task load in flipped learning. *British Journal of Educational Technology*, *51*(6), 2461–2477. https://doi.org/10.1111/bjet.12911

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. Appleton-Century.

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*(3), 126–132. https://doi.org/10.1016/j.tics.2015.01.004

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

*Ishibashi, T., Nakao, Y., & Sugano, Y. (2020). *Investigating audio data visualization for interactive sound recognition* [Conference session]. Proceedings of the 25th International Conference on Intelligent User Interfaces, Association for Computing Machinery. https://doi.org/10.1145/3377325.3377483

*Jafari, M., Zaeri, F., Jafari, A. H., Payandeh Najafabadi, A. T., Al-Qaisi, S., & Hassanzadeh-Rangi, N. (2020). Assessment and monitoring of mental workload in subway train operations using physiological, subjective, and performance measures. *Human Factors and Ergonomics in Manufacturing*, *30*(3), 165–175. https://doi.org/10.1002/hfm.20831

*Jain, M., Tripathi, R., Bhansali, I., & Kumar, P. (2019). *Automatic generation and evaluation of usable and secure audio reCAPTCHA* [Conference session]. The 21st International ACM SIGACCESS Conference on Computers and Accessibility, Association for Computing Machinery. https://doi.org/10.1145/3308561.3353777

James, W. (1880). *The feeling of effort*. The Boston Society of Natural History.

Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2015). Implicit theories about willpower predict self-regulation and grades in everyday life. *Journal of Personality and Social Psychology*, *108*(4), 637–647. https://doi.org/10.1037/pspp0000014

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kahneman, D., & Peavler, W. S. (1969). Incentive effects and pupillary changes in association learning. *Journal of Experimental Psychology*, *79*(2, Pt. 1), 312–318. https://doi.org/10.1037/h0026912

Kahneman, D., Peavler, W. S., & Onuska, L. (1968). Effects of verbalization and incentive on the pupil response to mental activity. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *22*(3), 186–196. https://doi.org/10.1037/h0082759

*Keunecke, J. G., Gall, C., Birkholz, T., Moritz, A., Eiche, C., & Prottengeier, J. (2019). Workload and influencing factors in non-emergency medical transfers: A multiple linear regression analysis of a cross-sectional questionnaire study. *BMC Health Services Research*, *19*(1), Article 812. https://doi.org/10.1186/s12913-019-4638-4

*Khairat, S., Coleman, C., Ottmar, P., Bice, T., Koppel, R., & Carson, S. S. (2019). Physicians' gender and their use of electronic health records: Findings from a mixed-methods usability study. *Journal of the American Medical Informatics Association*, *26*(12), 1505–1514. https://doi.org/10.1093/jamia/ocz126

*Khalaf, A. S., Alharthi, S. A., Alshehri, A., Dolgov, I., & Toups Dugas, P. O. (2020). A comparative study of hand-gesture recognition devices for games. In M. Kurosu (Ed.), *Human–computer interaction. Multimodal and natural interaction* (Vol. 12182, pp. 57–76). Springer International Publishing. https://doi.org/10.1007/978-3-030-49062-1_4

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, Article 513. https://doi.org/10.3389/fpsyg.2013.00513

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. https://doi.org/10.1038/s41562-018-0401-9

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

*Krekhov, A., Cmentowski, S., Waschk, A., & Kruger, J. (2020). Deadeye visualization revisited: Investigation of preattentiveness and applicability in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 547–557. https://doi.org/10.1109/TVCG.2019.2934370

Kruger, J., Wirtz, D., Van Boven, L., & Altermatt, T. W. (2004). The effort heuristic. *Journal of Experimental Social Psychology*, *40*(1), 91–98. https://doi.org/10.1016/S0022-1031(03)00065-9

*Kruger, J. M., & Bodemer, D. (2020). *Different types of interaction with augmented reality learning material* [Conference session]. 2020 6th International Conference of the Immersive Learning Research Network (iLRN), San Luis Obispo, CA, United States. https://doi.org/10.23919/iLRN47897.2020.9155148

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

*Kurzhals, K., Göbel, F., Angerbauer, K., Sedlmair, M., & Raubal, M. (2020). *A view on the viewer: Gaze-adaptive captions for videos* [Conference session]. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery. https://doi.org/10.1145/3313831.3376266

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

*Large, D. R., Harrington, K., Burnett, G., & Georgiou, O. (2019). Feel the noise: Mid-air ultrasound haptics as a novel human-vehicle interaction paradigm. *Applied Ergonomics*, *81*, Article 102909. https://doi.org/10.1016/j.apergo.2019.102909

Lassiter, G. D., Briggs, M. A., & Slaw, R. D. (1991). Need for cognition, causal processing, and memory for behavior. *Personality and Social Psychology Bulletin*, *17*(6), 694–700. https://doi.org/10.1177/0146167291176013

*Lee, J.-I., Asente, P., Kim, B., Kim, Y., & Stuerzlinger, W. (2020). *Evaluating automatic parameter control methods for locomotion in multiscale virtual environments* [Symposium]. 26th ACM Symposium on Virtual Reality Software and Technology, Association for Computing Machinery. https://doi.org/10.1145/3385956.3418961

*Lee, L.-H., Zhu, Y., Yau, Y.-P., Braud, T., Su, X., & Hui, P. (2020). *One-thumb text acquisition on force-assisted miniature interfaces for mobile headsets* [Conference session]. 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom), Austin, TX, United States. https://doi.org/10.1109/PerCom45495.2020.9127378

*Léger, É., Reyes, J., Drouin, S., Popa, T., Hall, J. A., Collins, D. L., & Kersten-Oertel, M. (2020). MARIN: An open-source mobile augmented reality interactive neuronavigation system. *International Journal of Computer Assisted Radiology and Surgery*, *15*(6), 1013–1021. https://doi.org/10.1007/s11548-020-02155-6

*Lewis, B., & Vogel, D. (2020). Longer delays in rehearsal-based interfaces increase expert use. *ACM Transactions on Computer–Human Interaction*, *27*(6), Article 45. https://doi.org/10.1145/3418196

*Li, M., Ganni, S., Ponten, J., Albayrak, A., Rutkowski, A.-F., & Jakimowicz, J. (2020). *Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training* [Conference session]. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, United States. https://doi.org/10.1109/VR46266.2020.1581301697128

*Li, S., Zhang, T., Zhang, W., Liu, N., & Lyu, G. (2020). Effects of speech-based intervention with positive comments on reduction of driver's anger state and perceived workload, and improvement of driving performance. *Applied Ergonomics*, *86*, Article 103098. https://doi.org/10.1016/j.apergo.2020.103098

*Li, W.-C., Horn, A., Sun, Z., Zhang, J., & Braithwaite, G. (2020). Augmented visualization cues on primary flight display facilitating pilot's monitoring performance. *International Journal of Human–Computer Studies*, *135*, Article 102377. https://doi.org/10.1016/j.ijhcs.2019.102377

*Li, W.-C., Yan, Z., Zhang, J., Braithwaite, G., Court, S., Lone, M., & Thapa, B. (2020). Evaluating pilot's perceived workload on interacting with augmented reality device in flight operations. In D. Harris & W.-C. Li (Eds.), *Engineering psychology and cognitive ergonomics. Cognition and design* (Vol. 12187, pp. 332–340). Springer International Publishing. https://doi.org/10.1007/978-3-030-49183-3_26

Lin, H., Westbrook, A., Fan, F., & Inzlicht, M. (2024). An experimental manipulation increases the value of effort. *Nature Human Behaviour*, *8*, 988–1000. https://doi.org/10.1038/s41562-024-01842-7

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *Annals of the American Academy of Political and Social Science*, *587*(1), 69–81. https://doi.org/10.1177/0002716202250791

Lowe, R. J. (2015). Cram schools in Japan: The need for research. *Language Teaching*, *39*(1), 26–31. https://doi.org/10.37546/JALTTLT39.1-5

*Lowndes, B. R., Forsyth, K. L., Blocker, R. C., Dean, P. G., Truty, M. J., Heller, S. F., Blackmon, S., Hallbeck, M. S., & Nelson, H. (2020). NASA-TLX assessment of surgeon workload variation across specialties. *Annals of Surgery*, *271*(4), 686–692. https://doi.org/10.1097/SLA.0000000000003058

*Lu, F., Davari, S., Lisle, L., Li, Y., & Bowman, D. A. (2020). *Glanceable AR: Evaluating information access methods for head-worn augmented reality* [Conference session]. 2020 IEEE conference on virtual reality and 3D user interfaces (VR), Atlanta, GA, United States. https://doi.org/10.1109/VR46266.2020.00113

*MacCormick, D., & Zaman, L. (2020). *Echo: Analyzing gameplay sessions by reconstructing them from recorded data* [Conference session]. Proceedings of the Annual Symposium on Computer–Human Interaction in Play, Association for Computing Machinery. https://doi.org/10.1145/3410404.3414254

Macnamara, B. N., & Burgoyne, A. P. (2023). Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices. *Psychological Bulletin*, *149*(3–4), 133–173. https://doi.org/10.1037/bul0000352

Mækelæ, M. J., Klevjer, K., Westbrook, A., Eby, N. S., Eriksen, R., & Pfuhl, G. (2023). Is it cognitive effort you measure? Comparing three task paradigms to the Need for Cognition scale. *PLOS ONE*, *18*(8), Article e0290177. https://doi.org/10.1371/journal.pone.0290177

*Mandlekar, A., Booher, J., Spero, M., Tung, A., Gupta, A., Zhu, Y., Garg, A., Savarese, S., & Fei-Fei, L. (2019). *Scaling robot supervision to hundreds of hours with roboturk: Robotic manipulation dataset through human reasoning and dexterity* [Conference session]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China. https://doi.org/10.1109/IROS40897.2019.8968114

*Markov-Vetter, D., Luboschik, M., Islam, A. T., Gauger, P., & Staadt, O. (2020). *The effect of spatial reference on visual attention and workload during viewpoint guidance in augmented reality* [Symposium]. Symposium on Spatial User Interaction, Association for Computing Machinery. https://doi.org/10.1145/3385959.3418449

Massar, S. A. A., Lim, J., Sasmita, K., & Chee, M. W. L. (2019). Sleep deprivation increases the costs of attentional effort: Performance, preference and pupil size. *Neuropsychologia*, *123*, 169–177. https://doi.org/10.1016/j.neuropsychologia.2018.03.032

*Matsuda, Y., & Komuro, T. (2020). *Dynamic layout optimization for multi-user interaction with a large display* [Conference session]. Proceedings of the 25th International Conference on Intelligent User Interfaces, Association for Computing Machinery. https://doi.org/10.1145/3377325.3377481

Matthews, J., Pisauro, M. A., Jurgelis, M., Müller, T., Vassena, E., Chong, T. T.-J., & Apps, M. A. J. (2023). Computational mechanisms underlying the dynamics of physical and cognitive fatigue. *Cognition*, *240*, Article 105603. https://doi.org/10.1016/j.cognition.2023.105603

*Mazur, L. M., Adams, R., Mosaly, P. R., Stiegler, M. P., Nuamah, J., Adapa, K., Chera, B., & Marks, L. B. (2020). Impact of simulation-based training on radiation therapists' workload, situation awareness, and performance. *Advances in Radiation Oncology*, *5*(6), 1106–1114. https://doi.org/10.1016/j.adro.2020.09.008

*Medathati, N. V. K., Desai, R., & Hillis, J. (2020). *Towards inferring cognitive state changes from pupil size variations in real world conditions* [Symposium]. ACM Symposium on Eye Tracking Research and Applications, Association for Computing Machinery. https://doi.org/10.1145/3379155.3391319

*Meena, Y. K., Cecotti, H., Wong-Lin, K., & Prasad, G. (2019). Design and evaluation of a time adaptive multimodal virtual keyboard. *Journal on Multimodal User Interfaces*, *13*(4), 343–361. https://doi.org/10.1007/s12193-019-00293-z

*Mendes, H. C. M., Costa, C. I. A. B., da Silva, N. A., Leite, F. P., Esteves, A., & Lopes, D. S. (2020). PIÑATA: Pinpoint insertion of intravenous needles via augmented reality training assistance. *Computerized Medical Imaging and Graphics*, *82*, Article 101731. https://doi.org/10.1016/j.compmedimag.2020.101731

*Mendes, V., Bruyere, F., Escoffre, J. M., Binet, A., Lardy, H., Marret, H., Marchal, F., & Hebert, T. (2020). Experience implication in subjective surgical ergonomics comparison between laparoscopic and robot-assisted surgeries. *Journal of Robotic Surgery*, *14*(1), 115–121. https://doi.org/10.1007/s11701-019-00933-2

*Milleville-Pennel, I., & Marquez, S. (2020). Comparison between elderly and young drivers' performances on a driving simulator and self-assessment of their driving attitudes and mastery. *Accident Analysis and Prevention*, *135*, Article 105317. https://doi.org/10.1016/j.aap.2019.105317

*Mills, B., Dykstra, P., Hansen, S., Miles, A., Rankin, T., Hopper, L., Brook, L., & Bartlett, D. (2020). Virtual reality triage training can provide comparable simulation efficacy for paramedicine students compared to live simulation-based scenarios. *Prehospital Emergency Care*, *24*(4), 525–536. https://doi.org/10.1080/10903127.2019.1676345

*Mingardi, M., Pluchino, P., Bacchin, D., Rossato, C., & Gamberini, L. (2020). Assessment of implicit and explicit measures of mental workload in working situations: Implications for industry 4.0. *Applied Sciences*, *10*(18), Article 6416. https://doi.org/10.3390/app10186416

*Mohanavelu, K., Poonguzhali, S., Ravi, D., Singh, P. K., Mahajabin, M., Ramachandran, K., Singh, U. K., & Jayaraman, S. (2020). Cognitive workload analysis of fighter aircraft pilots in flight simulator environment. *Defence Science Journal*, *70*(2), 131–139. https://doi.org/10.14429/dsj.70.14539

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Morsella, E., Wilson, L. E., Berger, C. C., Honhongva, M., Gazzaley, A., & Bargh, J. A. (2009). Subjective aspects of cognitive control at different stages of processing. *Attention, Perception, & Psychophysics*, *71*(8), 1807–1824. https://doi.org/10.3758/APP.71.8.1807

Müller, T., Klein-Flügge, M. C., Manohar, S. G., Husain, M., & Apps, M. A. J. (2021). Neural and computational mechanisms of momentary fatigue and persistence in effort-based choice. *Nature Communications*, *12*(1), Article 4593. https://doi.org/10.1038/s41467-021-24927-7

*Muñoz, A., Martí, A., Mahiques, X., Gracia, L., Solanes, J. E., & Tornero, J. (2020). Camera 3D positioning mixed reality-based interface to improve worker safety, ergonomics and productivity. *CIRP Journal of Manufacturing Science and Technology*, *28*, 24–37. https://doi.org/10.1016/j.cirpj.2020.01.004

Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. *Trends in Cognitive Sciences*, *25*(9), 757–775. https://doi.org/10.1016/j.tics.2021.06.001

*Nadj, M., Maedche, A., & Schieder, C. (2020). The effect of interactive analytical dashboard features on situation awareness and task performance. *Decision Support Systems*, *135*, Article 113322. https://doi.org/10.1016/j.dss.2020.113322

*Navratil, G., Konturek, P., & Giannopoulos, I. (2020). Interacting with 3d models–3D-CAD vs. holographic models. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *6*, 129–134. https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-129-2020

Newell, A., & Simon, H. A. (1972). *Human problem solving* (pp. xiv, 920). Prentice-Hall.

*Nino, L., Marchak, F., & Claudio, D. (2020). Physical and mental workload interactions in a sterile processing department. *International Journal of Industrial Ergonomics*, *76*, Article 102902. https://doi.org/10.1016/j.ergon.2019.102902

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, *22*(3), 453–460. https://doi.org/10.1016/j.jcps.2011.08.002

*Nuamah, J. K., Seong, Y., Jiang, S., Park, E., & Mountjoy, D. (2020). Evaluating effectiveness of information visualizations using cognitive fit theory: A neuroergonomics approach. *Applied Ergonomics*, *88*, Article 103173. https://doi.org/10.1016/j.apergo.2020.103173

Oldham, G. R., & Fried, Y. (2016). Job design research and theory: Past, present and future. *Organizational Behavior and Human Decision Processes*, *136*, 20–35. https://doi.org/10.1016/j.obhdp.2016.05.002

Organisation for Economic Co-operation and Development. (2017). *PISA 2015 results: Students' well-being* (Vol. 3). OECD Publishing.

Otto, T., Zijlstra, F. R. H., & Goebel, R. (2014). Neural correlates of mental effort evaluation—Involvement of structures related to self-awareness. *Social Cognitive and Affective Neuroscience*, *9*(3), 307–315. https://doi.org/10.1093/scan/nss136

*Panganiban, A. R., Matthews, G., & Long, M. D. (2020). Transparency in autonomous teammates: Intention to support as teaming information. *Journal of Cognitive Engineering and Decision Making*, *14*(2), 174–190. https://doi.org/10.1177/1555343419881563

*Parent, M., Albuquerque, I., Tiwari, A., Cassani, R., Gagnon, J.-F., Lafond, D., Tremblay, S., & Falk, T. H. (2020). PASS: A multimodal database of physical activity and stress for mobile passive body/ brain-computer interface research. *Frontiers in Neuroscience*, *14*, Article 542934. https://doi.org/10.3389/fnins.2020.542934

Petty, R. E., Brinol, P., Loersch, C., & McCaslin, M. J. (2009). The need for cognition. In M. H. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 318–329). Guilford Press.

Piercy, K. L., Troiano, R. P., Ballard, R. M., Carlson, S. A., Fulton, J. E., Galuska, D. A., George, S. M., & Olson, R. D. (2018). The physical activity guidelines for americans. *JAMA*, *320*(19), 2020–2028. https://doi.org/10.1001/jama.2018.14854

*Ponathil, A., Ozkan, F., Welch, B., Bertrand, J., & Chalil Madathil, K. (2020). Family health history collected by virtual conversational agents: An empirical study to investigate the efficacy of this approach. *Journal of Genetic Counseling*, *29*(6), 1081–1092. https://doi.org/10.1002/jgc4.1239

Preston, J., & Wegner, D. M. (2009). Elbow grease: When action feels like work. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *The Oxford handbook of human action* (pp. 569–586). Oxford University Press.

Priester, J. R., & Petty, R. E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, *21*(6), 637–654. https://doi.org/10.1177/0146167295216010

Pustejovsky, J. E. (2022). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections* (Version 0.5.5) [Computer software]. https://CRAN.R-project.org/package=clubSandwich

Pustejovsky, J. E., & Tipton, E. (2018). Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics*, *36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004

*Ramirez Gomez, A., & Gellersen, H. (2020). *More than looking: Using eye movements behind the eyelids as a new game mechanic* [Symposium]. Proceedings of the Annual Symposium on Computer–Human Interaction in Play, Association for Computing Machinery. https://doi.org/10.1145/3410404.3414240

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

*Reiser, J. E., Wascher, E., & Arnau, S. (2019). Recording mobile EEG in an outdoor environment reveals cognitive-motor interference dependent on movement complexity. *Scientific Reports*, *9*(1), Article 13086. https://doi.org/10.1038/s41598-019-49503-4

*Reiser, J. E., Wascher, E., Rinkenauer, G., & Arnau, S. (2021). Cognitive-motor interference in the wild: Assessing the effects of movement complexity on task switching using mobile EEG. *European Journal of Neuroscience*, *54*(12), 8175–8195. https://doi.org/10.1111/ejn.14959

Richter, M., Friedrich, A., & Gendolla, G. H. E. (2008). Task difficulty effects on cardiac activity. *Psychophysiology*, *45*(5), 869–875. https://doi.org/10.1111/j.1469-8986.2008.00688.x

Richter, M., & Gendolla, G. H. E. (2009). The heart contracts to reward: Monetary incentives and preejection period. *Psychophysiology*, *46*(3), 451–457. https://doi.org/10.1111/j.1469-8986.2009.00795.x

Richter, M., Gendolla, G. H. E., & Wright, R. A. (2016). Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 3, pp. 149–186). Elsevier. https://doi.org/10.1016/bs.adms.2016.02.001

Rohatgi, A. (2021). *Webplotdigitizer* (Version 4.5). https://automeris.io/WebPlotDigitizer

*Romanowski, A., Chaniecki, Z., Koralczyk, A., Woźniak, M., Nowak, A., Kucharski, P., Jaworski, T., Malaya, M., Rózga, P., & Grudzień, K. (2020). Interactive timeline approach for contextual spatio-temporal ECT data investigation. *Sensors*, *20*(17), Article 4793. https://doi.org/10.3390/s20174793

Rugulies, R., Aust, B., & Madsen, I. E. (2017). Effort-reward imbalance at work and risk of depressive disorders. A systematic review and meta-analysis of prospective cohort studies. *Scandinavian Journal of Work, Environment & Health*, *43*(4), 294–306. https://doi.org/10.5271/sjweh.3632

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*(1), 3–30. https://doi.org/10.1037/0033-2909.125.1.3

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

*Saito, Y., & Raksincharoensak, P. (2019). Effect of risk-predictive haptic guidance in one-pedal driving mode. *Cognition Technology & Work*, *21*(4), 671–684. https://doi.org/10.1007/s10111-019-00558-3

*Schwarz, S., Regal, G., Kempf, M., & Schatz, R. (2020). *Learning success in immersive virtual reality training environments: Practical evidence from automotive assembly* [Conference session]. Proceedings of the 11th Nordic Conference on Human–Computer Interaction: Shaping Experiences, Shaping Society, Association for Computing Machinery. https://doi.org/10.1145/3419249.3420182

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

*Shuggi, I. M., Oh, H., Wu, H., Ayoub, M. J., Moreno, A., Shaw, E. P., Shewokis, P. A., & Gentili, R. J. (2019). Motor performance, mental workload and self-efficacy dynamics during learning of reaching movements throughout multiple practice sessions. *Neuroscience*, *423*, 232–248. https://doi.org/10.1016/j.neuroscience.2019.07.001

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology*, *1*(1), 27–41. https://doi.org/10.1037/1076-8998.1.1.27

Silvestrini, N., Musslick, S., Berry, A. S., & Vassena, E. (2023). An integrative effort: Bridging motivational intensity theory and recent neurocomputational and neuronal models of effort and control allocation. *Psycholological Review*, *130*(4), 1081–1103. https://doi.org/10.1037/rev0000372

Silvetti, M., Lasaponara, S., Daddaoua, N., Horan, M., & Gottlieb, J. (2023). A Reinforcement Meta-Learning framework of executive function and information demand. *Neural Networks*, *157*, 103–113. https://doi.org/10.1016/j.neunet.2022.10.004

Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, *5*, Article 75. https://doi.org/10.3389/fnhum.2011.00075

Silvetti, M., Vassena, E., Abrahamse, E., & Verguts, T. (2018). Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLOS Computational Biology*, *14*(8), Article e1006370. https://doi.org/10.1371/journal.pcbi.1006370

Simmons, W. K., Avery, J. A., Barcalow, J. C., Bodurka, J., Drevets, W. C., & Bellgowan, P. (2013). Keeping the body in mind: Insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Human Brain Mapping*, *34*(11), 2944–2958. https://doi.org/10.1002/hbm.22113

*Solanes, J. E., Muñoz, A., Gracia, L., Martí, A., Girbés-Juan, V., & Tornero, J. (2020). Teleoperation of industrial robot manipulators based on augmented reality. *International Journal of Advanced Manufacturing Technology*, *111*(3–4), 1077–1097. https://doi.org/10.1007/s00170-020-05997-1

*Souders, D. J., Charness, N., Roque, N. A., & Pham, H. (2020). Aging: Older adults' driving behavior using longitudinal and lateral warning systems. *Human Factors*, *62*(2), 229–248. https://doi.org/10.1177/0018720819864510

*Spinelli, R., Magagnotti, N., & Labelle, E. R. (2020). The effect of new silvicultural trends on mental workload of harvester operators. *Croatian Journal of Forest Engineering*, *41*(2), 177–190. https://doi.org/10.5552/crojfe.2020.747

Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(2), 316–345. https://doi.org/10.1037/0278-7393.11.2.316

Steele, J. (2020). *What is (perception of) effort? Objective and subjective effort during attempted task performance*. PsyArXiv. https://doi.org/10.31234/osf.io/kbyhm

Steinhart, Y., & Wyer, R. S., Jr. (2009). Motivational correlates of need for cognition. *European Journal of Social Psychology*, *39*(4), 608–621. https://doi.org/10.1002/ejsp.565

*Szychowska, M., & Wiens, S. (2020). Visual load does not decrease the auditory steady-state response to 40-Hz amplitude-modulated tones. *Psychophysiology*, *57*(12), Article e13689. https://doi.org/10.1111/psyp.13689

*Takizawa, R., Verhulst, A., Seaborn, K., Fukuoka, M., Hiyama, A., Kitazaki, M., Inami, M., & Sugimoto, M. (2019). *Exploring perspective dependency in a shared body with virtual supernumerary robotic arms* [Conference session]. 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), San Diego, CA, United States. https://doi.org/10.1109/AIVR46125.2019.00014

*Tokadli, G., Dorneich, M. C., Matessa, M., & Eda, S. (2019). *The evaluation of a playbook interface for human-autonomy teaming in single pilot operations* [Conference session]. 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC), San Diego, CA, United States. https://doi.org/10.1109/DASC43569.2019.9081620

Torka, A.-K., Mazei, J., & Hüffmeier, J. (2021). Together, everyone achieves more-or, less? An interdisciplinary meta-analysis on effort gains and losses in teams. *Psychological Bulletin*, *147*(5), 504–534. https://doi.org/10.1037/bul0000251

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology*, *121*(3), 553–558. https://doi.org/10.1037/a0028813

Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Kessler, R. M., & Zald, D. H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *The Journal of Neuroscience*, *32*(18), 6170–6176. https://doi.org/10.1523/JNEUROSCI.6459-11.2012

Tsai, L. S. (1932). *The laws of minimum effort and maximum satisfaction in animal behavior*. The National Research Institute of Psychology.

*Turner, C. J., Chaparro, B. S., & He, J. (2021). Typing on a smartwatch while mobile: A comparison of input methods. *Human Factors*, *63*(6), 974–986. https://doi.org/10.1177/0018720819891291

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

*Van Cutsem, J., De Pauw, K., Vandervaeren, C., Marcora, S., Meeusen, R., & Roelands, B. (2019). Mental fatigue impairs visuomotor response time in badminton players and controls. *Psychology of Sport and Exercise*, *45*, Article 101579. https://doi.org/10.1016/j.psychsport.2019.101579

van Hoorn, A., & Maseland, R. (2013). Does a protestant work ethic exist? Evidence from the well-being effect of unemployment. *Journal of Economic Behavior & Organization*, *91*, 1–12. https://doi.org/10.1016/j.jebo.2013.03.038

*Varas-Diaz, G., Kannan, L., & Bhatt, T. (2020). Effect of mental fatigue on postural sway in healthy older adults and stroke populations. *Brain Sciences*, *10*(6), Article 388. https://doi.org/10.3390/brainsci10060388

Vassena, E., Cobbaert, S., Andres, M., Fias, W., & Verguts, T. (2015). Unsigned value prediction-error modulates the motor system in absence of choice. *NeuroImage*, *122*, 73–79. https://doi.org/10.1016/j.neuroimage.2015.07.081

Vassena, E., Deraeve, J., & Alexander, W. H. (2017). Predicting motivation: Computational models of PFC can explain neural coding of motivation and effort-based decision-making in health and disease. *Journal of Cognitive Neuroscience*, *29*(10), 1633–1645. https://doi.org/10.1162/jocn_a_01160

Vassena, E., Deraeve, J., & Alexander, W. H. (2019). Task-specific prioritization of reward and effort information: Novel insights from behavior and computational modeling. *Cognitive, Affective & Behavioral Neuroscience*, *19*(3), 619–636. https://doi.org/10.3758/s13415-018-00685-w

Vassena, E., Holroyd, C. B., & Alexander, W. H. (2017). Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. *Frontiers in Neuroscience*, *11*, Article 316. https://doi.org/10.3389/fnins.2017.00316

Vassena, E., Silvetti, M., Boehler, C. N., Achten, E., Fias, W., & Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. *PLOS ONE*, *9*(3), Article e91008. https://doi.org/10.1371/journal.pone.0091008

Verguts, T., Vassena, E., & Silvetti, M. (2015). Adaptive effort investment in cognitive and physical tasks: A neurocomputational model. *Frontiers in Behavioral Neuroscience*, *9*, Article 57. https://doi.org/10.3389/fnbeh.2015.00057

*Verschueren, J. O., Tassignon, B., Proost, M., Teugels, A., Van Cutsem, J., Roelands, B., Verhagen, E., & Meeusen, R. (2020). Does mental fatigue negatively affect outcomes of functional performance tests? *Medicine and Science in Sports and Exercise*, *52*(9), 2002–2010. https://doi.org/10.1249/MSS.0000000000002323

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vogel, T. A., Savelson, Z. M., Otto, A. R., & Roy, M. (2020). Forced choices reveal a trade-off between cognitive effort and physical pain. *eLife*, *9*, Article e59410. https://doi.org/10.7554/eLife.59410

*Vurgun, N., Vongsurbchart, T., Myszka, A., Richter, P., & Rogula, T. (2021). Medical student experience with robot-assisted surgery after limited laparoscopy exposure. *Journal of Robotic Surgery*, *15*(3), 443–450. https://doi.org/10.1007/s11701-020-01129-9

Walters, K. J., & Simons, J. S. (2022). The structure of affect: A network analytic moderation approach. *Motivation and Emotion*, *46*(1), 110–125. https://doi.org/10.1007/s11031-021-09916-z

*Wang, I., Buchweitz, L., Smith, J., Bornholdt, L.-S., Grund, J., Ruiz, J., & Korn, O. (2020). *Wow, you are terrible at this! An intercultural study on virtual agents giving mixed feedback* [Conference session]. Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, Association for Computing Machinery. https://doi.org/10.1145/3383652.3423887

*Wang, Y., Hu, Y., & Chen, Y. (2021). An experimental investigation of menu selection for immersive virtual environments: Fixed versus handheld menus. *Virtual Reality*, *25*(2), 409–419. https://doi.org/10.1007/s10055-020-00464-4

*Wang, Y., Wang, K., Huang, Y., Wu, D., Wu, J., & He, J. (2020). An assessment method and toolkit to evaluate keyboard design on smartphones. *Journal of Visualized Experiments*, *164*, Article e61796. https://doi.org/10.3791/61796

*Warner, D. O., Nolan, M., Garcia-Marcinkiewicz, A., Schultz, C., Warner, M. A., Schroeder, D. R., & Cook, D. A. (2020). Adaptive instruction and learner interactivity in online learning: A randomized trial. *Advances in Health Sciences Education*, *25*(1), 95–109. https://doi.org/10.1007/s10459-019-09907-3

Western Cape Education Department. (2023). *Growth mindset (GM)*. https://wcedonline.westerncape.gov.za/growth-mindset-gm

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

*Widyanti, A., & Firdaus, M. (2020). Assessment of mental workload of flight attendants based on flight duration: An effort to provide safe working condition. *Aviation*, *23*(3), 97–103. https://doi.org/10.3846/aviation.2019.11847

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575.e5. https://doi.org/10.1016/j.cub.2022.07.010

*Wikström, V., Martikainen, S., Falcon, M., Ruistola, J., & Saarikivi, K. (2020). Collaborative block design task for assessing pair performance in virtual reality and reality. *Heliyon*, *6*(9), Article e04823. https://doi.org/10.1016/j.heliyon.2020.e04823

*Woźniak, M. P., Dominiak, J., Pieprzowski, M., Ładoński, P., Grudzień, K., Lischke, L., Romanowski, A., & Woźniak, P. W. (2020). Subtletee: Augmenting posture awareness for beginner golfers. *Proceedings of the ACM on Human–Computer Interaction*, *4*, Article 204. https://doi.org/10.1145/3427332

*Xu, W., Liang, H.-N., Zhang, Z., & Baghaei, N. (2020). Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. *Games for Health Journal*, *9*(6), 405–414. https://doi.org/10.1089/g4h.2019.0102

*Xu, X., Shi, H., Yi, X., Liu, W., Yan, Y., Shi, Y., Mariakakis, A., Mankoff, J., & Dey, A. K. (2020). *EarBuddy: Enabling on-face interaction via wireless earbuds* [Conference session]. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery. https://doi.org/10.1145/3313831.3376836

*Xu, X., Yu, C., Wang, Y., & Shi, Y. (2020). Recognizing unintentional touch on interactive tabletop. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(1), Article 33. https://doi.org/10.1145/3381011

*Yang, Y., Karreman, J., & de Jong, M. (2020). *Comparing the effects of paper and mobile augmented reality instructions to guide assembly tasks* [Conference session]. 2020 IEEE International Professional Communication Conference (ProComm), Kennesaw, GA, United States. https://doi.org/10.1109/ProComm48883.2020.00021

*Yi, H., Hong, J., Kim, H., & Lee, W. (2019). *DexController: Designing a VR controller with grasp-recognition for enriching natural game experience* [Symposium]. 25th ACM Symposium on Virtual Reality Software and Technology, Association for Computing Machinery. https://doi.org/10.1145/3359996.3364263

*Yu, D., Zhou, Q., Newn, J., Dingler, T., Velloso, E., & Goncalves, J. (2020). Fully-occluded target selection in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *26*(12), 3402–3413. https://doi.org/10.1109/TVCG.2020.3023606

*Yuksel, B. F., Fazli, P., Mathur, U., Bisht, V., Kim, S. J., Lee, J. J., Jin, S. J., Siu, Y.-T., Miele, J. A., & Yoon, I. (2020). *Human-in-the-loop machine learning to increase video accessibility for visually impaired and blind users* [Conference session]. Proceedings of the 2020 ACM Designing Interactive Systems Conference, Association for Computing Machinery. https://doi.org/10.1145/3357236.3395433

*Yuksel, B. F., Kim, S. J., Jin, S. J., Lee, J. J., Fazli, P., Mathur, U., Bisht, V., Yoon, I., Siu, Y.-T., & Miele, J. A. (2020). *Increasing video accessibility for visually impaired users with human-in-the-loop machine learning* [Conference session]. Extended Abstracts of the 2020 CHI Conference on

Human Factors in Computing Systems, Association for Computing Machinery. https://doi.org/10.1145/3334480.3382821

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review*, *30*, Article 100326. https://doi.org/10.1016/j.edurev.2020.100326

*Zargari Marandi, R., Fjelsted, C. A., Hrustanovic, I., Dan Olesen, R., & Gazerani, P. (2020). Eye movements in response to pain-related feelings in the presence of low and high cognitive loads. *Behavioral Sciences*, *10*(5), Article 92. https://doi.org/10.3390/bs10050092

Zedelius, C. M., Müller, B. C. N., & Schooler, J. W. (2017). *The science of lay theories: How beliefs shape our cognition, behavior, and health* (pp. xii, 382). Springer. https://doi.org/10.1007/978-3-319-57306-9

Zedelius, C. M., Protzko, J., & Schooler, J. W. (2021). Lay theories of the wandering mind: Control-related beliefs predict mind wandering rates in- and outside the lab. *Personality and Social Psychology Bulletin*, *47*(6), 921–938. https://doi.org/10.1177/0146167220949408

*Zeller, R., Williamson, A., & Friswell, R. (2020). The effect of sleep-need and time-on-task on driver fatigue. *Transportation Research Part F: Traffic Psychology and Behaviour*, *74*, 15–29. https://doi.org/10.1016/j.trf.2020.08.001

Zentall, T. R. (2010). Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*(5), 296–300. https://doi.org/10.1177/0963721410383381

Ziano, I., Yeung, S. K., Lee, C. S., Shi, J., & Feldman, G. (2023). “The effort heuristic” revisited: Mixed results for replications of Kruger et al. (2004)’s Experiments 1 and 2. *Collabra Psychology*, *9*(1), Article 87489. https://doi.org/10.1525/collabra.87489

Received September 23, 2022
Revision received May 23, 2024
Accepted May 24, 2024 ■

Supplementary information with *The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect*

by Louise David, Eliana Vassena, & Erik Bijleveld (2024)

Table of contents

Associations between moderators

Table S1. Association strength (Cramer's V) between all categorical moderators.

		1	3	4a	4b	5	6	7	8	9	10	14	15
1	Education (2)	1											
3	Skill–task fit (2)	.29	1										
4a	Continent (3)	.10	.16	1									
4b	Country (3)	.05	.15	1	1								
5	Skill variety (2)	.26	.56	.16	.25	1							
6	Monitoring feedback (2)	.09	.14	.09	.17	.13	1						
7	Performance feedback (2)	.01	.04	.21	.50	.05	.21	1					
8	Control (2)	.25	.48	.13	.14	.66	.10	.06	1				
9	Task significance (2)	.19	.66	.16	.21	.66	.22	.00	.48	1			
10	Task identity (2)	.09	.58	.13	.28	.68	.07	.09	.61	.57	1		
14	Physical activity (2)	.01	.09	.08	.21	.10	.11	.09	.04	.09	.03	1	
15	Group setting (3)	.27	.33	.14	.25	.28	.05	.13	.24	.29	.21	.40	1

Note: The numbering in the leftmost column corresponds to the numbering in Table 1 in the main text. Numbers between brackets refer to the number of categories of that moderator. Derived from $\chi2$, Cramer's V is an effect size measure for pairs of nominal variables. V can range between 0 (no association) and 1 (perfect association). V is sometimes referred to as $\varphi_{C.}$

Following common rules of thumb (Cohen, 1988, *Statistical power for the behavioral sciences*, 2nd Ed.), the interpretation of V depends on the lowest number of categories in the variable pair.

- If the variable with the lowest number of categories has 2 categories, V is considered large when V ≥ .50.
- If the variable with the lowest number of categories has 3 categories, V is considered large when V ≥ .35.

Shaded cells in Table S1 contain effect sizes that can be considered large, following these rules of thumb.

To interpret the 10 large associations in Table S1, we computed odds ratios (ORs). We report these by going through Table S1 column by column, from left to right:

- Tasks that had high skill–task fit, were more likely to also have high task variety (OR = 15.6), high task significance (OR = 95.5), and high task identity (17.3)
- Naturally, there was a perfect association between continent and country.
- Tasks from Canada were more likely to have continuous performance feedback, when compared to tasks form the USA (OR = 11.9) and Germany (OR = 15.4).
- Tasks that had high skill variety, were more likely to also have high control (OR = 29.9), high task significance (OR = 28.4), and high task identity (OR = 27.8).
- Tasks that had high control, were more likely to also have high task significance (OR = 10.2).
- Tasks that had high task significance, were more likely to also have high task identity (OR = 15.2).

Table S2. Association strength (η^2) between the four continuous moderators (columns) with the categorical moderators (rows).

		Work experience	Age	Gender	Duration of task
1	Education	.00	.32	.02	.02
3	Skill–task fit	.02	.07	.00	.06
4a	Continent	.16	.02	.00	.02
4b	Country	.01	.12	.02	.06
5	Skill variety	.00	.05	.01	.03
6	Monitoring feedback	.06	.02	.04	.01
7	Performance feedback	.00	.00	.01	.00
8	Control	.02	.01	.02	.01
9	Task significance	-	.09	.02	.04
10	Task identity	.01	.02	.00	.01
14	Physical activity	.10	.02	.00	.01
15	Group setting	.14	.00	.03	.17

Note: η^2 is an effect size measure that can be used to quantify the strength of association between continuous and categorical variables. Shaded cells contain effect sizes that can be considered large, $\eta^2 \geq .14$, following common rules of thumb (Cohen, 1988). We were unable to compute η^2 for the link between task significance and work experience, because where we could code work experience (k = 58), task significance was always high.

To interpret the 4 large associations in Table S1, we interpreted the means of the continuous moderators separately for all levels of the relevant categorical moderators. We report these by going through Table S2 column by column, from left to right:

On average, participants in studies conducted in Europe (M = 13.3y, SD = 8.8y) and Asia (M = 10.5y, SD = 11.8y) had more work experience than participants in studies conducted in North-America (M = 4.1y, SD = 3.9y).

On average, participants who took part in studies individually (M = 11.3y, SD = 9.9y) or with observers present (M = 15.0y, SD = 10.1y), had more work experience than participants who took part in studies that required them to engage with others (M = 5.9y, SD = 6.5y).

On average, participants without university education (M = 35.8y, SD = 12.2y) were older than participants with university education (M = 24.4y, SD = 4.0y).

On average, tasks that were administered with observers present were relatively short (M = 14.7m, SD = 16.6m). Tasks that required participants to engage with others were relatively long (M = 84.3m), though there was a lot of variation (SD = 123.0). Tasks that were conducted individually fell in between (M = 38.7m, SD = 16.6m).

Table S3. Association strength (Pearson r) between the four continuous moderators.

	Work experience	**Age**	**Gender**	**Duration of task**
Work experience	1			
Age	.89	1		
Gender	-.29	-.20	1	
Duration of task	-.16	.03	.24	1

Note: Work experience and age were coded in years; gender was coded as the proportion of females in the sample; duration of task was coded in minutes.

Following common rules of thumb (Cohen, 1988), we note the strong correlation between age and work experience. Naturally, participants with more work experience also tended to be older.

Detailed results of moderator analyses

Table S4. Moderator analyses for learning history moderators.

	Moderator	Term	β	95%CI	Q_M	Q_E	I^2
1	Education	Intercept	32.6 ***	[23.0, 42.3]	341.7 ***	35486.8 ***	97.8
		Effort	0.8 **	[0.3, 1.3]			
		Education = University	3.0	[-6.6, 12.7]			
		Effort × Education	0.1	[-0.3, 0.6]			
2	Work experience	Intercept	34.4 ***	[24.4, 44.4]	296.1 ***	8436.5 ***	98.8
		Effort	1.0 **	[0.5, 1.5]			
		Work experience	-0.3	[-1.5, 0.9]			
		Effort × Work experience	0.0	[-0.1, 0.1]			
3	Skill-task fit	Intercept	35.5 ***	[32.8, 38.2]	554.9 ***	33284.2 ***	97.2
		Effort	0.8 ***	[0.7, 1.0]			
		Skill–task fit = High	-0.7	[-5.8, 4.3]			
		Effort × Skill–task fit	0.0	[-0.3, 0.3]			
4a	Continent	Intercept	34.3 ***	[30.2, 38.4]	657.7 ***	25830.8 ***	97.2
		Effort	1.1 ***	[0.8, 1.3]			
		Continent = Asia	-1.7	[-7.2, 3.9]			
		Continent = Europe	3.4	[-2.0, 8.9]			
		Effort × Continent = Asia	-0.5 **	[-0.8, -0.2]			
		Effort × Continent = Europe	-0.2	[-0.5, 0.1]			
4b	Country	Intercept	36.2 ***	[32.3, 40.1]	285.0 ***	1324.1 ***	91.1
		Effort	1.1 ***	[0.9, 1.3]			
		Country = Canada	-7.9	[-17.4, 1.6]			
		Country = Germany	7.9 *	[0.4, 15.3]			
		Effort × Country = Canada	-0.4	[-1.1, 0.3]			
		Effort × Country = Germany	-0.4	[-0.9, 0.2]			

Note: Each moderator is tested in a separate model. In all models, the dependent variable was negative affect. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table S5. Moderator analyses for task design moderators.

	Moderator	Term	β	95%CI	Q_M	Q_E	I^2
5	Skill variety	Intercept	35.2 ***	[32.9, 37.5]	567.9 ***	22433.9 ***	97.2
		Effort	0.8 ***	[0.7, 1.0]			
		Skill variety = High	0.0	[-4.5, 4.4]			
		Effort × Skill variety	0.1	[-0.2, 0.3]			
6	Monitoring feedback	Intercept	34.8 ***	[30.3, 39.3]	553.0 ***	24414.0 ***	97.2
		Effort	0.9 ***	[0.7, 1.1]			
		Monitoring feedback = Yes	0.6	[-4.2, 5.4]			
		Effort × Monitoring feedback	-0.1	[-0.3, 0.2]			
7	Performance feedback	Intercept	35.2 ***	[32.9, 37.6]	552.5 ***	38745.8 ***	97.3
		Effort	0.9 ***	[0.7, 1.0]			
		Performance feedback = Throughout	0.6	[-3.1, 4.4]			
		Effort × Performance feedback	-0.1	[-0.6, 0.4]			
8	Control	Intercept	34 ***	[32.0, 36.0]	578.7 ***	18501.6 ***	97.1
		Effort	0.8 ***	[0.7, 0.9]			
		Control = High	3.4	[-1.3, 8.0]			
		Effort × Control	0.1	[-0.1, 0.3]			
9	Task significance	Intercept	34.8 ***	[32.2, 37.4]	548.3 ***	23351.9 ***	97.1
		Effort	0.9 ***	[0.7, 1.0]			
		Task significance = High/medium	0.8	[-3.6, 5.3]			
		Effort × Task significance	0.0	[-0.3, 0.2]			
10	Task identity	Intercept	34.4 ***	[32.0, 36.9]	581.7 ***	21226.0 ***	97.2
		Effort	0.8 ***	[0.6, 0.9]			
		Task identity = High	1.8	[-3.1, 6.7]			
		Effort × Task identity	0.1	[-0.1, 0.4]			

Note: Each moderator was tested in a separate model. In all models, the dependent variable was negative affect. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table S6. Moderator analyses for task design moderators.

Term	β	95%CI	Q_M	Q_E	I^2
Intercept	33.5 ***	[29.4, 37.7]	601.9 ***	13342.3 ***	97.1
Effort	0.9 ***	[0.6, 1.1]			
Skill variety	-2.4	[-7.0, 2.1]			
Monitoring feedback	1.4	[-4.3, 7.1]			
Performance feedback	1.1	[-3.1, 5.2]			
Control	4.0	[-0.8, 8.8]			
Task significance	-1.0	[-6.2, 4.2]			
Task identity	2.1	[-5.4, 9.6]			
Effort × Skill variety	0.0	[-0.4, 0.4]			
Effort × Monitoring feedback	-0.1	[-0.3, 0.2]			
Effort × Performance feedback	-0.1	[-0.5, 0.4]			
Effort × Control	0.1	[-0.3, 0.5]			
Effort × Task significance	-0.3	[-0.6, 0.1]			
Effort × Task identity	0.3	[-0.3, 0.8]			

Note: All moderators were tested together in the same model. The dependent variable was negative affect. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table S7. Moderator analyses for exploratory moderators.

	Moderator	Term	β	95%CI	Q_M	Q_E	I^2
11	Age	Intercept	35.6 ***	[32.8, 38.4]	356.2 ***	38291.5 ***	98.0
		Effort	0.8 ***	[0.6, 0.9]			
		Age	-0.2	[-0.5, 0.1]			
		Effort × Age	0.0	[0.0, 0.0]			
12	Gender	Intercept	34.2 ***	[32.2, 36.2]	513.2 ***	2384.1 ***	95.2
		Effort	0.8 ***	[0.7, 0.9]			
		Gender	-5.0	[-15.2, 5.3]			
		Effort × Gender	-0.3	[-0.7, 0.1]			
13	Duration	Intercept	35.8 ***	[32.6, 39.0]	286.4 ***	1258.0 ***	92.5
		Effort	0.8 ***	[0.6, 1.0]			
		Duration	0.0	[-0.1, 0.1]			
		Effort × Duration	0.0	[0.0, 0.0]			
14	Physical activity	Intercept	36.3 ***	[33.7, 38.9]	635.0 ***	36818.5 ***	97.5
		Effort	0.8 ***	[0.7, 0.9]			
		Physical activity = Light activity	-2.3	[-6.8, 2.2]			
		Effort × Physical activity	0.2	[0.0, 0.4]			
15	Group setting	Intercept	36.9 ***	[33.4, 40.4]	551.4 ***	37516 ***	97.2
		Effort	0.8 ***	[0.7, 0.9]			
		Group setting = Together with others	-2.8	[-8.7, 3.1]			
		Group setting = Observers present	-2.8	[-6.8, 1.2]			
		Effort × Group Setting: Together w. others	0.0	[-0.5, 0.5]			
		Effort × Group Setting: Observers present	0.1	[-0.1, 0.4]			

Note: Each moderator was tested in a separate model. In all models, the dependent variable was negative affect. * p < .05; ** p < .01; *** p < .001.

Detailed results of robustness analysis

Table S8. Robustness analysis.

	Term	β	95%CI	Q_M	Q_E	I^2
	Intercept	35.8 ***	[28.5, 43.2]	597.2 ***	10030.0 ***	97.3
	Effort	0.9 ***	[0.8, 1.0]			
3	Skill–task fit = High	-1.4	[-11.3, 8.4]			
4a	Continent = Asia	-2.7	[-9.9, 4.6]			
4a	Continent = Europe	3.8	[-1.9, 9.6]			
5	Skill variety = High	-2.8	[-10.3, 4.7]			
6	Monitoring feedback = Yes	0.1	[-5.6, 5.8]			
7	Performance feedback = Throughout	0.6	[-3.8, 4.9]			
8	Control = High	5.2 *	[0.2, 10.3]			
9	Task significance = High/medium	-0.1	[-7.5, 7.2]			
10	Task identity = High	1.9	[-5.2, 9.0]			
14	Physical activity = Light activity	-3.1	[-9.7, 3.5]			
15	Group setting = Together with others	-1.7	[-8.7, 5.4]			
15	Group setting = Observers present	-2.8	[-7.1, 1.4]			

Note: All moderators are tested together in the same model. The dependent variable was negative affect. * $p < .05$; ** $p < .01$; *** $p < .001$.

Simulations to examine if individual-level response biases can account for the results

Background

One may suspect that our main finding—i.e., the association between effort and negative affect—does not reflect a true association between effort and negative affect, but that it can be explained from individual-level response biases. For example, people who are more likely to agree with questionnaire statements (i.e., people higher in acquiescence bias) will likely score higher both on effort *and* on negative affect. As a result of such variation in response tendencies, within individual samples, the correlation between effort and negative affect could be inflated or even fully spurious. However, our meta-analysis was done *across samples*, i.e., on summary statistics that were computed on the sample level. It is an open question whether within-samples correlations (which may be due to response biases) can explain our main finding. We conducted simulations to explore this question.

Method

We present four sets of simulations. In each set, we assumed a different true correlation between effort and negative affect (within individual samples). In the first set, we assumed $\rho = .30$; in the second, $\rho = .50$; in the third, $\rho = .70$; in the fourth, $\rho = .90$.

The core part of our simulation script (https://osf.io/mktbr/) worked as follows:

1) We first simulated individual datasets. In several ways, these simulated datasets mirrored the datasets that we included in our meta-analysis. That is, they had similar N ($N_{mean} = 26$, $N_{sd} = 16$, $N_{min} = 10$). Also, the sample means and standard deviations for effort and negative affect were similar to those from the real datasets. However, as described above, we assumed different correlations between effort and negative affect in the populations from which the simulated samples were drawn (i.e., different ρs). We constructed these datasets using the rnorm_multi() function from the *faux* package in R (DeBruine, 2021, https:/doi.org/10.5281/zenodo.2669586).
2) From these datasets, we constructed a meta-analysis dataset. This meta-analysis dataset included means and standard deviations for effort and negative affect, computed for each of the simulated samples. As in our meta-analysis, this meta-analysis dataset contained summary data from 357 individual samples.
3) We ran a meta-analysis on the dataset constructed at step #2, using the same procedures as described in the main text.
4) We stored the β-value for the effect of effort on negative affect.

We ran the core part of our script 4 x 1000 times, i.e., 1000 times for each correlation. We report the distributions of β-values and compare these distributions to the β of 0.85 that we found in our meta-analysis.

If within-samples correlations did not affect the outcome of the meta-analysis at all, βs should be distributed around 0 (as we did not manufacture any across-samples association in our simulations). By contrast, if our main finding can be fully explained by within-samples correlations, βs should be in same range as the β we found in our meta-analysis.

Results

Figure S1. Simulation results.

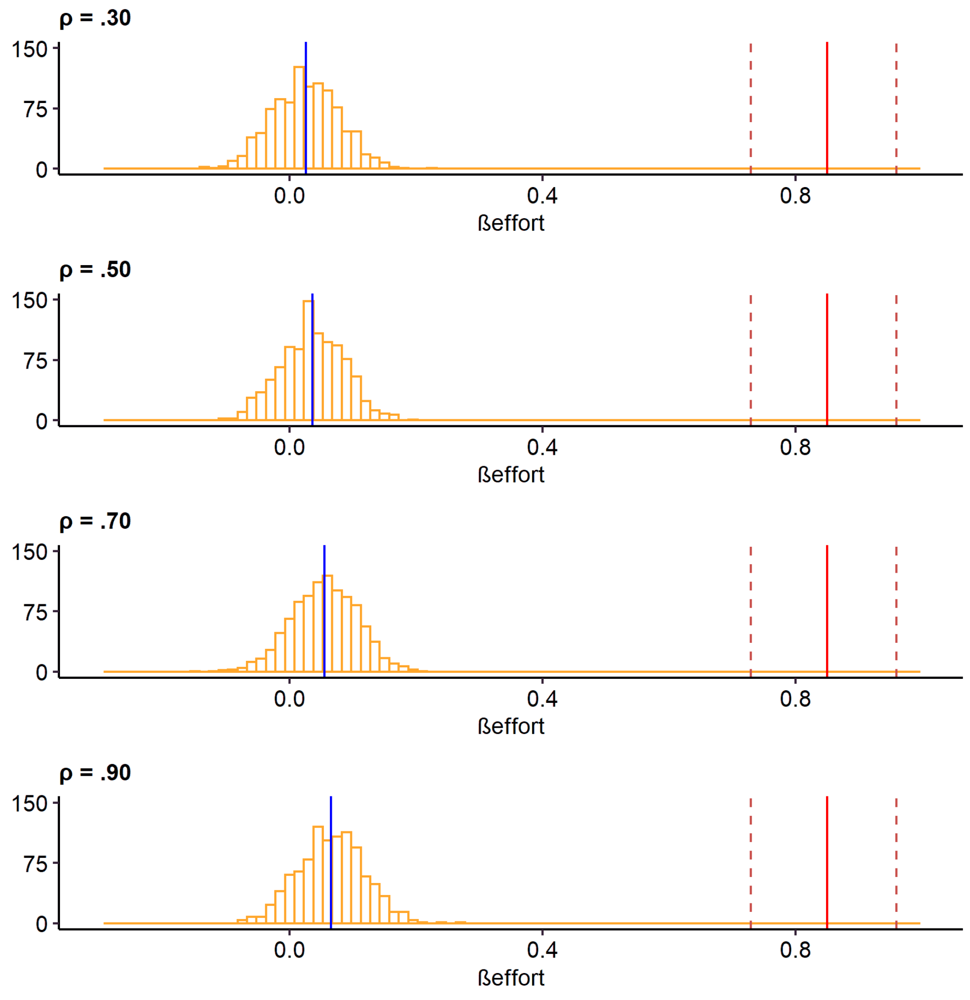


Figure S1 shows results from four sets of simulations, each with a different correlation (within individual samples). Histograms reflect the distribution of simulated β-values for the meta-analytic effect of effort on negative affect. Red vertical lines mark the β-value of 0.85 that we found in our main meta-analysis (with the 95% confidence intervals in dashed lines).

Discussion

Inspection of Figure S1 reveals that the simulated meta-analyses yielded β-values that were, on average, slightly above 0. So, if within-samples correlations were present in the original datasets (which, in turn, could be due to response biases), these may have slightly inflated the overall effect size that we found. We say 'slightly' because even when within-samples correlations were unrealistically strong (ρ = .90), the distribution of β-values centered only at around 0.07. Importantly, inspection of Figure S1 further reveals that the β-value we found in our meta-analysis was well outside the range of our simulations, suggesting that within-samples correlations (and, thus, potential response biases) cannot account for our findings.

List of papers included in the meta-analysis

Abd Rahman, N. I., Md Dawal, S. Z., & Yusoff, N. (2020). Driving mental workload and performance of ageing drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *69*, 265–285. https://doi.org/10.1016/j.trf.2020.01.019

Alagi, H., Navarro, S. E., Hergenhan, J., Music, S., & Hein, B. (2020). Teleoperation with tactile feedback based on a capacitive proximity sensor array. *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1–6. https://doi.org/10.1109/I2MTC43012.2020.9128701

Armougum, A., Gaston-Bellegarde, A., Marle, C. J.-L., & Piolino, P. (2020). Expertise reversal effect: Cost of generating new schemas. *Computers in Human Behavior*, *111*, 106406. https://doi.org/10.1016/j.chb.2020.106406

Aromaa, S., Väätänen, A., Aaltonen, I., Goriachev, V., Helin, K., & Karjalainen, J. (2020). Awareness of the real-world environment when using augmented reality head-mounted display. *Applied Ergonomics*, *88*, 103145. https://doi.org/10.1016/j.apergo.2020.103145

Arsintescu, L., Chachad, R., Gregory, K. B., Mulligan, J. B., & Flynn-Evans, E. E. (2020). The relationship between workload, performance and fatigue in a short-haul airline. *Chronobiology International*, *37*(9–10), 1492–1494. https://doi.org/10.1080/07420528.2020.1804924

Ataya, A., Kim, W., Elsharkawy, A., & Kim, S. (2020). Gaze-head input: Examining potential interaction with immediate experience sampling in an autonomous vehicle. *Applied Sciences*, *10*(24), 9011. https://doi.org/10.3390/app10249011

Badesa, F. J., Diez, J. A., Catalan, J. M., Trigili, E., Cordella, F., Nann, M., Crea, S., Soekadar, S. R., Zollo, L., Vitiello, N., & Garcia-Aracil, N. (2019). Physiological responses during hybrid BCNI control of an upper-limb exoskeleton. *Sensors*, *19*(22), 4931. https://doi.org/10.3390/s19224931

Beege, M., Nebel, S., Schneider, S., & Rey, G. D. (2021). The effect of signaling in dependence on the extraneous cognitive load in learning environments. *Cognitive Processing*, *22*(2), 209–225. https://doi.org/10.1007/s10339-020-01002-5

Berg, M., Bayazit, D., Mathew, R., Rotter-Aboyoun, A., Pavlick, E., & Tellex, S. (2020). Grounding language to landmarks in arbitrary outdoor environments. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 208–215. https://doi.org/10.1109/ICRA40945.2020.9197068

Bier, L., Emele, M., Gut, K., Kulenovic, J., Rzany, D., Peter, M., & Abendroth, B. (2019). Preventing the risks of monotony related fatigue while driving through gamification. *European Transport Research Review*, *11*(1), 44. https://doi.org/10.1186/s12544-019-0382-4

Blundell, J., Scott, S., Harris, D., Huddlestone, J., & Richards, D. (2020a). With flying colours: Pilot performance with colour-coded head-up flight symbology. *Displays*, *61*, 101932. https://doi.org/10.1016/j.displa.2019.101932

Blundell, J., Scott, S., Harris, D., Huddlestone, J., & Richards, D. (2020b). Workload benefits of colour coded head-up flight symbology during high workload flight. *Displays*, *65*, 101973. https://doi.org/10.1016/j.displa.2020.101973

Bogaard, T., Wielemaker, J., Hollink, L., Hardman, L., & van Ossenbruggen, J. (2020). Understanding user behavior in digital libraries using the MAGUS session visualization tool. In M. Hall, T. Merčun, T. Risse, & F. Duchateau (Eds.), *Digital Libraries for Open Knowledge* (Vol. 12246, pp. 171–184). Springer International Publishing. https://doi.org/10.1007/978-3-030-54956-5_13

Braarud, P. Ø. (2020). An efficient screening technique for acceptable mental workload based on the NASA Task Load Index—

Development and application to control room validation. *International Journal of Industrial Ergonomics*, *76*, 102904. https://doi.org/10.1016/j.ergon.2019.102904

Caggianese, G., Capece, N., Erra, U., Gallo, L., & Rinaldi, M. (2020). Freehand-steering locomotion techniques for immersive virtual environments: A comparative evaluation. *International Journal of Human–Computer Interaction*, *36*(18), 1734–1755. https://doi.org/10.1080/10447318.2020.1785151

Cai, S., Ke, P., Narumi, T., & Zhu, K. (2020). ThermAirGlove: A pneumatic glove for thermal perception and material identification in virtual reality. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 248–257. https://doi.org/10.1109/VR46266.2020.00044

Calvi, A., D'Amico, F., Ciampoli, L. B., & Ferrante, C. (2020). Evaluation of driving performance after a transition from automated to manual control: A driving simulator study. *Transportation Research Procedia*, *45*, 755–762. https://doi.org/10.1016/j.trpro.2020.02.101

Caria, M., Todde, G., Sara, G., Piras, M., & Pazzona, A. (2020). Performance and usability of smartglasses for augmented reality in precision livestock farming operations. *Applied Sciences*, *10*(7), 2318. https://doi.org/10.3390/app10072318

Cecotti, H., Day-Scott, Z., Huisinga, L., & Gordo-Pelaez, L. (2020). Virtual reality for immersive learning in art history. *2020 6th International Conference of the Immersive Learning Research Network (ILRN)*, 16–23. https://doi.org/10.23919/iLRN47897.2020.9155108

Chacko, S. M., Granado, A., & Kapila, V. (2020). An augmented reality framework for robotic tool-path teaching. *Procedia CIRP*, *93*, 1218–1223. https://doi.org/10.1016/j.procir.2020.03.143

Chan, W. P., Sakr, M., Quintero, C. P., Croft, E., & Van der Loos, H. F. M. (2020). Towards a multimodal system combining augmented reality and electromyography for robot trajectory programming and execution. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 419–424. https://doi.org/10.1109/RO-MAN47096.2020.9223526

Chen, Y., Katsuragawa, K., & Lank, E. (2020). Understanding viewport- and world-based pointing with everyday smart devices in immersive augmented reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376592

Ciumedean, C.-B., Patras, C., Cibulskis, M., Váradi, N., & Nilsson, N. C. (2020). Mission impossible spaces: Using challenge-based distractors to reduce noticeability of self-overlapping virtual architecture. *Symposium on Spatial User Interaction*, 1–4. https://doi.org/10.1145/3385959.3418453

Czarnek, G., Strojny, P., Strojny, A., & Richter, M. (2020). Assessing engagement during rescue operation simulated in virtual reality: A psychophysiological study. *International Journal of Human–Computer Interaction*, *36*(5), 464–476. https://doi.org/10.1080/10447318.2019.1655905

Datta, R. R., Schönhage, S., Dratsch, T., Toader, J., Müller, D. T., Wahba, R., Kleinert, R., Thomas, M., Dieplinger, G., Stippel, D. L., Bruns, C. J., & Fuchs, H. F. (2021). Learning curve of surgical novices using the single-port platform SymphonX: Minimizing OR trauma to only one 15-mm incision. *Surgical Endoscopy*, *35*(9), 5338–5351. https://doi.org/10.1007/s00464-020-07998-3

de Melo, C. M., Kim, K., Norouzi, N., Bruder, G., & Welch, G. (2020). Reducing cognitive load and improving warfighter problem solving with intelligent virtual assistants. *Frontiers in Psychology*, *11*, 554706. https://doi.org/10.3389/fpsyg.2020.554706

DeLucia, P. R., & Greenlee, E. T. (2020). Tactile vigilance is stressful and demanding. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872082096529. https://doi.org/10.1177/0018720820965294

Desolda, G., Matera, M., & Lanzilotti, R. (2020). Metamorphic data sources: A user-centric paradigm to consume linked data in interactive workspaces. *Future Generation Computer Systems*, *102*, 992–1015. https://doi.org/10.1016/j.future.2019.09.032

Dickinson, P., Gerling, K., Wilson, L., & Parke, A. (2020). Virtual reality as a platform for research in gambling behaviour. *Computers in Human Behavior*, *107*, 106293. https://doi.org/10.1016/j.chb.2020.106293

Ding, Y., Cao, Y., Duffy, V. G., Wang, Y., & Zhang, X. (2020). Measurement and identification of mental workload during simulated computer tasks with multimodal methods and machine learning. *Ergonomics*, *63*(7), 896–908. https://doi.org/10.1080/00140139.2020.1759699

Do, S., & Lee, B. (2020). Improving reliability of virtual collision responses: A cue integration technique. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376819

Dodwell, A., & Trick, L. M. (2020). The effects of secondary tasks that involve listening and speaking on young adult drivers with traits associated with autism spectrum disorders: A pilot study with driving simulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *69*, 120–134. https://doi.org/10.1016/j.trf.2019.12.011

Ezcurdia, I., Arregui, A., Ardaiz, O., Ortiz, A., & Marzo, A. (2020). Content adaptation and depth perception in an affordable multi-view display. *Applied Sciences*, *10*(20), 7357. https://doi.org/10.3390/app10207357

Filho, J. A. W., Stuerzlinger, W., & Nedel, L. (2020). Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 514–524. https://doi.org/10.1109/TVCG.2019.2934415

Ganni, S., Li, M., Botden, S. M. B. I., Nayak, S. R., Ganni, B. R., Rutkowski, A.-F., Goossens, R. H. M., & Jakimowicz, J. (2020). Virtual operating room simulation setup (vorss) for procedural training in minimally invasive surgery – a pilot study. *Indian Journal of Surgery*, *82*(5), 810–816. https://doi.org/10.1007/s12262-020-02131-z

Gerull, W., Zihni, A., & Awad, M. (2020). Operative performance outcomes of a simulator-based robotic surgical skills curriculum. *Surgical Endoscopy*, *34*(10), 4543–4548. https://doi.org/10.1007/s00464-019-07243-6

Ghaleb, A. M., Khalaf, T. M., Ramadan, M. Z., Ragab, A. E., & Badwelan, A. (2020). Effect of cycling on a stationary bike while performing assembly tasks on human physiology and performance parameters. *International Journal of Environmental Research and Public Health*, *17*(5), 1761. https://doi.org/10.3390/ijerph17051761

Gilardi, F., De Falco, F., Casasanta, D., Andellini, M., Gazzellini, S., Petrarca, M., Morocutti, A., Lettori, D., Ritrovato, M., Castelli, E., Raponi, M., Magnavita, N., & Zaffina, S. (2020). Robotic technology in pediatric neurorehabilitation: A pilot study of human factors in an italian pediatric hospital. *International Journal of Environmental Research and Public Health*, *17*(10), 3503. https://doi.org/10.3390/ijerph17103503

Glimne, S., Brautaset, R., & Österman, C. (2020). Visual fatigue during control room work in process industries. *Work*, *65*(4), 903–914. https://doi.org/10.3233/WOR-203141

Gold, B., & Windscheid, J. (2020). Observing 360-degree classroom videos – Effects of video type on presence, emotions, workload, classroom observations, and ratings of teaching quality. *Computers & Education*, *156*, 103960. https://doi.org/10.1016/j.compedu.2020.103960

Golmohammadi, R., Darvishi, E., Faradmal, J., Poorolajal, J., & Aliabadi, M. (2020). Attention and short-term memory during occupational noise exposure considering task difficulty. *Applied Acoustics*, *158*, 107065. https://doi.org/10.1016/j.apacoust.2019.107065

Greenhalgh, M., Blaauw, E., Deepak, N., St. Laurent, C. O. L. M., Cooper, R., Bendixen, R., Koontz, A. M., & Cooper, R. A. (2020). Usability and task load comparison between a robotic assisted transfer device and a mechanical floor lift during caregiver assisted transfers on a care recipient. *Disability and Rehabilitation: Assistive Technology*, 1–7. https://doi.org/10.1080/17483107.2020.1818137

Greenlee, E. T., Lui, T. G., & Maw, E. L. (2021). Is physiobehavioral monitoring nonintrusive? An examination of transcranial doppler sonography in a vigilance task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *63*(7), 1256–1270. https://doi.org/10.1177/0018720820920118

Gross, I. T., Whitfill, T., Redmond, B., Couturier, K., Bhatnagar, A., Joseph, M., Joseph, D., Ray, J., Wagner, M., & Auerbach, M. (2020). Comparison of two telemedicine delivery modes for neonatal resuscitation support: A simulation-based randomized trial. *Neonatology*, *117*(2), 159–166. https://doi.org/10.1159/000504853

Guo, W., & Kim, J. H. (2020). How augmented reality influences student workload in engineering education. In C. Stephanidis, D. Harris, W.-C. Li, D. D. Schmorrow, C. M. Fidopiastis, P. Zaphiris, A. Ioannou, X. Fang, R. A. Sottilare, & J. Schwarz (Eds.), *HCI International 2020 – Late Breaking Papers: Cognition, Learning and Games* (Vol. 12425, pp. 388–396). Springer International Publishing. https://doi.org/10.1007/978-3-030-60128-7_29

Hafiz, P., & Bardram, J. E. (2020). The ubiquitous cognitive assessment tool for smartwatches: Design, implementation, and evaluation study. *JMIR MHealth and UHealth*, *8*(6), e17506. https://doi.org/10.2196/17506

Hangli, G., Hamada, T., Sumitomo, T., & Koshizuka, N. (2020). Intellevator: An intelligent elevator system proactive in traffic control for time-efficiency improvement. *IEEE Access*, *8*, 35535–35545. https://doi.org/10.1109/ACCESS.2020.2975020

Hoppe, A. H., Marek, F., Camp, F. van de, & Stiefelhagen, R. (2020). Extending movable surfaces with touch interaction using the VirtualTablet: An extended view. *Advances in Science, Technology and Engineering Systems Journal*, *5*(2), 328–337. https://doi.org/10.25046/aj050243

Hsia, L., & Hwang, G. (2020). From reflective thinking to learning engagement awareness: A reflective thinking promoting approach to improve students' dance performance, self-efficacy and task load in flipped learning. *British Journal of Educational Technology*, *51*(6), 2461–2477. https://doi.org/10.1111/bjet.12911

Ishibashi, T., Nakao, Y., & Sugano, Y. (2020). Investigating audio data visualization for interactive sound recognition. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 67–77. https://doi.org/10.1145/3377325.3377483

Jafari, M., Zaeri, F., Jafari, A. H., Payandeh Najafabadi, A. T., Al-Qaisi, S., & Hassanzadeh-Rangi, N. (2020). Assessment and monitoring of mental workload in subway train operations using physiological, subjective, and performance measures. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *30*(3), 165–175. https://doi.org/10.1002/hfm.20831

Jain, M., Tripathi, R., Bhansali, I., & Kumar, P. (2019). Automatic generation and evaluation of usable and secure audio reCAPTCHA. *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 355–366. https://doi.org/10.1145/3308561.3353777

Keunecke, J. G., Gall, C., Birkholz, T., Moritz, A., Eiche, C., & Prottengeier, J. (2019). Workload and influencing factors in non-emergency medical transfers: A multiple linear regression analysis of a cross-sectional questionnaire study. *BMC Health Services Research*, *19*(1), 812. https://doi.org/10.1186/s12913-019-4638-4

Khairat, S., Coleman, C., Ottmar, P., Bice, T., Koppel, R., & Carson, S. S. (2019). Physicians' gender and their use of electronic health records: Findings from a mixed-methods usability study. *Journal of the American Medical Informatics Association*, *26*(12), 1505–1514. https://doi.org/10.1093/jamia/ocz126

Khalaf, A. S., Alharthi, S. A., Alshehri, A., Dolgov, I., & Toups, Z. O. (2020). A comparative study of hand-gesture recognition devices for games. In M. Kurosu (Ed.), *Human-Computer Interaction. Multimodal and Natural Interaction* (Vol. 12182, pp. 57–76). Springer International Publishing. https://doi.org/10.1007/978-3-030-49062-1_4

Krekhov, A., Cmentowski, S., Waschk, A., & Kruger, J. (2019). Deadeye visualization revisited: Investigation of preattentiveness and applicability in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 1–1. https://doi.org/10.1109/TVCG.2019.2934370

Kruger, J. M., & Bodemer, D. (2020). Different types of interaction with augmented reality learning material. *2020 6th International Conference of the Immersive Learning Research Network (ILRN)*, 78–85. https://doi.org/10.23919/iLRN47897.2020.9155148

Kurzhals, K., Göbel, F., Angerbauer, K., Sedlmair, M., & Raubal, M. (2020). A view on the viewer: Gaze-adaptive captions for videos. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376266

Large, D. R., Harrington, K., Burnett, G., & Georgiou, O. (2019). Feel the noise: Mid-air ultrasound haptics as a novel human-vehicle interaction paradigm. *Applied Ergonomics*, *81*, 102909. https://doi.org/10.1016/j.apergo.2019.102909

Lee, J.-I., Asente, P., Kim, B., Kim, Y., & Stuerzlinger, W. (2020). Evaluating automatic parameter control methods for locomotion in multiscale virtual environments. *26th ACM Symposium on Virtual Reality Software and Technology*, 1–10. https://doi.org/10.1145/3385956.3418961

Lee, L.-H., Zhu, Y., Yau, Y.-P., Braud, T., Su, X., & Hui, P. (2020). One-thumb text acquisition on force-assisted miniature interfaces for mobile headsets. *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 1–10. https://doi.org/10.1109/PerCom45495.2020.9127378

Léger, É., Reyes, J., Drouin, S., Popa, T., Hall, J. A., Collins, D. L., & Kersten-Oertel, M. (2020). MARIN: An open-source mobile augmented reality interactive neuronavigation system. *International Journal of Computer Assisted Radiology and Surgery*, *15*(6), 1013–1021. https://doi.org/10.1007/s11548-020-02155-6

Lewis, B., & Vogel, D. (2020). Longer delays in rehearsal-based interfaces increase expert use. *ACM Transactions on Computer-Human Interaction*, *27*(6), 45:1-45:41. https://doi.org/10.1145/3418196

Li, M., Ganni, S., Ponten, J., Albayrak, A., Rutkowski, A.-F., & Jakimowicz, J. (2020). Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 566–572. https://doi.org/10.1109/VR46266.2020.1581301697128

Li, S., Zhang, T., Zhang, W., Liu, N., & Lyu, G. (2020). Effects of speech-based intervention with positive comments on reduction of driver's anger state and perceived workload, and improvement of driving performance. *Applied Ergonomics*, *86*, 103098. https://doi.org/10.1016/j.apergo.2020.103098

Li, W.-C., Horn, A., Sun, Z., Zhang, J., & Braithwaite, G. (2020). Augmented visualization cues on primary flight display facilitating pilot's monitoring performance. *International Journal of Human-Computer Studies*, *135*, 102377.

https://doi.org/10.1016/j.ijhcs.2019.102377

Li, W.-C., Yan, Z., Zhang, J., Braithwaite, G., Court, S., Lone, M., & Thapa, B. (2020). Evaluating pilot's perceived workload on interacting with augmented reality device in flight operations. In D. Harris & W.-C. Li (Eds.), *Engineering Psychology and Cognitive Ergonomics. Cognition and Design* (Vol. 12187, pp. 332–340). Springer International Publishing. https://doi.org/10.1007/978-3-030-49183-3_26

Lowndes, B. R., Forsyth, K. L., Blocker, R. C., Dean, P. G., Truty, M. J., Heller, S. F., Blackmon, S., Hallbeck, M. S., & Nelson, H. (2020). NASA-TLX assessment of surgeon workload variation across specialties. *Annals of Surgery*, *271*(4), 686–692. https://doi.org/10.1097/SLA.0000000000003058

Lu, F., Davari, S., Lisle, L., Li, Y., & Bowman, D. A. (2020). Glanceable AR: Evaluating information access methods for head-worn augmented reality. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 930–939. https://doi.org/10.1109/VR46266.2020.00113

MacCormick, D., & Zaman, L. (2020). Echo: Analyzing gameplay sessions by reconstructing them from recorded data. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 281–293. https://doi.org/10.1145/3410404.3414254

Mandlekar, A., Booher, J., Spero, M., Tung, A., Gupta, A., Zhu, Y., Garg, A., Savarese, S., & Fei-Fei, L. (2019). Scaling robot supervision to hundreds of hours with roboturk: Robotic manipulation dataset through human reasoning and dexterity. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1048–1055. https://doi.org/10.1109/IROS40897.2019.8968114

Markov-Vetter, D., Luboschik, M., Islam, A. T., Gauger, P., & Staadt, O. (2020). The effect of spatial reference on visual attention and workload during viewpoint guidance in augmented reality. *Symposium on Spatial User Interaction*, 1–10. https://doi.org/10.1145/3385959.3418449

Matsuda, Y., & Komuro, T. (2020). Dynamic layout optimization for multi-user interaction with a large display. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 401–409. https://doi.org/10.1145/3377325.3377481

Mazur, L. M., Adams, R., Mosaly, P. R., Stiegler, M. P., Nuamah, J., Adapa, K., Chera, B., & Marks, L. B. (2020). Impact of simulation-based training on radiation therapists' workload, situation awareness, and performance. *Advances in Radiation Oncology*, *5*(6), 1106–1114. https://doi.org/10.1016/j.adro.2020.09.008

Medathati, N. V. K., Desai, R., & Hillis, J. (2020). Towards inferring cognitive state changes from pupil size variations in real world conditions. *ACM Symposium on Eye Tracking Research and Applications*, 1–10. https://doi.org/10.1145/3379155.3391319

Meena, Y. K., Cecotti, H., Wong-Lin, K., & Prasad, G. (2019). Design and evaluation of a time adaptive multimodal virtual keyboard. *Journal on Multimodal User Interfaces*, *13*(4), 343–361. https://doi.org/10.1007/s12193-019-00293-z

Mendes, H. C. M., Costa, C. I. A. B., da Silva, N. A., Leite, F. P., Esteves, A., & Lopes, D. S. (2020). PIÑATA: Pinpoint insertion of intravenous needles via augmented reality training assistance. *Computerized Medical Imaging and Graphics*, *82*, 101731. https://doi.org/10.1016/j.compmedimag.2020.101731

Mendes, V., Bruyere, F., Escoffre, J. M., Binet, A., Lardy, H., Marret, H., Marchal, F., & Hebert, T. (2020). Experience implication in subjective surgical ergonomics comparison between laparoscopic and robot-assisted surgeries. *Journal of Robotic Surgery*, *14*(1), 115–121. https://doi.org/10.1007/s11701-019-00933-2

Milleville-Pennel, I., & Marquez, S. (2020). Comparison between elderly and young drivers' performances on a driving simulator and self-assessment of their driving attitudes and mastery. *Accident Analysis & Prevention*, *135*, 105317. https://doi.org/10.1016/j.aap.2019.105317

Mills, B., Dykstra, P., Hansen, S., Miles, A., Rankin, T., Hopper, L., Brook, L., & Bartlett, D. (2020). Virtual reality triage training can provide comparable simulation efficacy for paramedicine students compared to live simulation-based scenarios. *Prehospital Emergency Care*, *24*(4), 525–536. https://doi.org/10.1080/10903127.2019.1676345

Mingardi, M., Pluchino, P., Bacchin, D., Rossato, C., & Gamberini, L. (2020). Assessment of implicit and explicit measures of mental workload in working situations: Implications for industry 4.0. *Applied Sciences*, *10*(18), 6416. https://doi.org/10.3390/app10186416

Mohanavelu, K., Poonguzhali, S., Ravi, D., Singh, P. K., Mahajabin, M., K., R., Singh, U. K., & Jayaraman, S. (2020). Cognitive workload analysis of fighter aircraft pilots in flight simulator environment. *Defence Science Journal*, *70*(2), 131–139. https://doi.org/10.14429/dsj.70.14539

Muñoz, A., Martí, A., Mahiques, X., Gracia, L., Solanes, J. E., & Tornero, J. (2020). Camera 3D positioning mixed reality-based interface to improve worker safety, ergonomics and productivity. *CIRP Journal of Manufacturing Science and Technology*, *28*, 24–37. https://doi.org/10.1016/j.cirpj.2020.01.004

Nadj, M., Maedche, A., & Schieder, C. (2020). The effect of interactive analytical dashboard features on situation awareness and task performance. *Decision Support Systems*, *135*, 113322. https://doi.org/10.1016/j.dss.2020.113322

Navratil, G., Konturek, P., & Giannopoulos, I. (2020). Interacting with 3d models – 3D-CAD vs. Holographic models. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *VI-4/W1-2020*, 129–134. https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-129-2020

Nino, L., Marchak, F., & Claudio, D. (2020). Physical and mental workload interactions in a sterile processing department. *International Journal of Industrial Ergonomics*, *76*, 102902. https://doi.org/10.1016/j.ergon.2019.102902

Nuamah, J. K., Seong, Y., Jiang, S., Park, E., & Mountjoy, D. (2020). Evaluating effectiveness of information visualizations using cognitive fit theory: A neuroergonomics approach. *Applied Ergonomics*, *88*, 103173. https://doi.org/10.1016/j.apergo.2020.103173

Panganiban, A. R., Matthews, G., & Long, M. D. (2020). Transparency in autonomous teammates: Intention to support as teaming information. *Journal of Cognitive Engineering and Decision Making*, *14*(2), 174–190. https://doi.org/10.1177/1555343419881563

Parent, M., Albuquerque, I., Tiwari, A., Cassani, R., Gagnon, J.-F., Lafond, D., Tremblay, S., & Falk, T. H. (2020). PASS: A multimodal database of physical activity and stress for mobile passive body/ brain-computer interface research. *Frontiers in Neuroscience*, *14*, 542934. https://doi.org/10.3389/fnins.2020.542934

Ponathil, A., Ozkan, F., Welch, B., Bertrand, J., & Chalil Madathil, K. (2020). Family health history collected by virtual conversational agents: An empirical study to investigate the efficacy of this approach. *Journal of Genetic Counseling*, *29*(6), 1081–1092. https://doi.org/10.1002/jgc4.1239

Ramirez Gomez, A., & Gellersen, H. (2020). More than looking: Using eye movements behind the eyelids as a new game mechanic. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 362–373. https://doi.org/10.1145/3410404.3414240

Reiser, J. E., Wascher, E., & Arnau, S. (2019). Recording mobile EEG in an outdoor environment reveals cognitive-motor interference dependent on movement complexity. *Scientific Reports*, *9*(1), 13086. https://doi.org/10.1038/s41598-

019-49503-4

Reiser, J. E., Wascher, E., Rinkenauer, G., & Arnau, S. (2021). Cognitive-motor interference in the wild: Assessing the effects of movement complexity on task switching using mobile EEG. *European Journal of Neuroscience*, *54*(12), 8175–8195. https://doi.org/10.1111/ejn.14959

Romanowski, A., Chaniecki, Z., Koralczyk, A., Woźniak, M., Nowak, A., Kucharski, P., Jaworski, T., Malaya, M., Rózga, P., & Grudzień, K. (2020). Interactive timeline approach for contextual spatio-temporal ECT data investigation. *Sensors*, *20*(17), 4793. https://doi.org/10.3390/s20174793

Saito, Y., & Raksincharoensak, P. (2019). Effect of risk-predictive haptic guidance in one-pedal driving mode. *Cognition, Technology & Work*, *21*(4), 671–684. https://doi.org/10.1007/s10111-019-00558-3

Schwarz, S., Regal, G., Kempf, M., & Schatz, R. (2020). Learning success in immersive virtual reality training environments: Practical evidence from automotive assembly. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1–11. https://doi.org/10.1145/3419249.3420182

Shuggi, I. M., Oh, H., Wu, H., Ayoub, M. J., Moreno, A., Shaw, E. P., Shewokis, P. A., & Gentili, R. J. (2019). Motor performance, mental workload and self-efficacy dynamics during learning of reaching movements throughout multiple practice sessions. *Neuroscience*, *423*, 232–248. https://doi.org/10.1016/j.neuroscience.2019.07.001

Solanes, J. E., Muñoz, A., Gracia, L., Martí, A., Girbés-Juan, V., & Tornero, J. (2020). Teleoperation of industrial robot manipulators based on augmented reality. *The International Journal of Advanced Manufacturing Technology*, *111*(3–4), 1077–1097. https://doi.org/10.1007/s00170-020-05997-1

Souders, D. J., Charness, N., Roque, N. A., & Pham, H. (2020). Aging: Older adults' driving behavior using longitudinal and lateral warning systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *62*(2), 229–248. https://doi.org/10.1177/0018720819864510

Spinelli, R., Magagnotti, N., & Labelle, E. R. (2020). The effect of new silvicultural trends on mental workload of harvester operators. *Croatian Journal of Forest Engineering*, *41*(2), 177–190. https://doi.org/10.5552/crojfe.2020.747

Szychowska, M., & Wiens, S. (2020). Visual load does not decrease the auditory steady-state response to 40-Hz amplitude-modulated tones. *Psychophysiology*, *57*(12). https://doi.org/10.1111/psyp.13689

Takizawa, R., Verhulst, A., Seaborn, K., Fukuoka, M., Hiyama, A., Kitazaki, M., Inami, M., & Sugimoto, M. (2019). Exploring perspective dependency in a shared body with virtual supernumerary robotic arms. *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 25–257. https://doi.org/10.1109/AIVR46125.2019.00014

Tokadli, G., Dorneich, M. C., Matessa, M., & Eda, S. (2019). The evaluation of a playbook interface for human-autonomy teaming in single pilot operations. *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*, 1–7. https://doi.org/10.1109/DASC43569.2019.9081620

Turner, C. J., Chaparro, B. S., & He, J. (2021). Typing on a smartwatch while mobile: A comparison of input methods. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *63*(6), 974–986. https://doi.org/10.1177/0018720819891291

Van Cutsem, J., De Pauw, K., Vandervaeren, C., Marcora, S., Meeusen, R., & Roelands, B. (2019). Mental fatigue impairs visuomotor response time in badminton players and controls. *Psychology of Sport and Exercise*, *45*, 101579. https://doi.org/10.1016/j.psychsport.2019.101579

Varas-Diaz, G., Kannan, L., & Bhatt, T. (2020). Effect of mental fatigue on postural sway in healthy older adults and stroke populations. *Brain Sciences*, *10*(6), 388. https://doi.org/10.3390/brainsci10060388

Verschueren, J., Tassignon, B., Proost, M., Teugels, A., Van Cutsem, J., Roelands, B., Verhagen, E., & Meeusen, R. (2020). Does mental fatigue negatively affect outcomes of functional performance tests? *Medicine & Science in Sports & Exercise*, *52*(9), 2002–2010. https://doi.org/10.1249/MSS.0000000000002323

Vurgun, N., Vongsurbchart, T., Myszka, A., Richter, P., & Rogula, T. (2021). Medical student experience with robot-assisted surgery after limited laparoscopy exposure. *Journal of Robotic Surgery*, *15*(3), 443–450. https://doi.org/10.1007/s11701-020-01129-9

Wang, I., Buchweitz, L., Smith, J., Bornholdt, L.-S., Grund, J., Ruiz, J., & Korn, O. (2020). Wow, you are terrible at this! An intercultural study on virtual agents giving mixed feedback. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (pp. 1–8). Association for Computing Machinery. https://doi.org/10.1145/3383652.3423887

Wang, Y., Hu, Y., & Chen, Y. (2021). An experimental investigation of menu selection for immersive virtual environments: Fixed versus handheld menus. *Virtual Reality*, *25*(2), 409–419. https://doi.org/10.1007/s10055-020-00464-4

Wang, Y., Wang, K., Huang, Y., Wu, D., Wu, J., & He, J. (2020). An assessment method and toolkit to evaluate keyboard design on smartphones. *Journal of Visualized Experiments*, *164*, 61796. https://doi.org/10.3791/61796

Warner, D. O., Nolan, M., Garcia-Marcinkiewicz, A., Schultz, C., Warner, M. A., Schroeder, D. R., & Cook, D. A. (2020). Adaptive instruction and learner interactivity in online learning: A randomized trial. *Advances in Health Sciences Education*, *25*(1), 95–109. https://doi.org/10.1007/s10459-019-09907-3

Widyanti, A., & Firdaus, M. (2020). Assessment of mental workload of flight attendants based on flight duration: An effort to provide safe working condition. *Aviation*, *23*(3), 97–103. https://doi.org/10.3846/aviation.2019.11847

Wikström, V., Martikainen, S., Falcon, M., Ruistola, J., & Saarikivi, K. (2020). Collaborative block design task for assessing pair performance in virtual reality and reality. *Heliyon*, *6*(9), e04823. https://doi.org/10.1016/j.heliyon.2020.e04823

Woźniak, M. P., Dominiak, J., Pieprzowski, M., Ładoński, P., Grudzień, K., Lischke, L., Romanowski, A., & Woźniak, P. W. (2020). Subtletee: Augmenting posture awareness for beginner golfers. *Proceedings of the ACM on Human-Computer Interaction*, *4*(ISS), 204:1-204:24. https://doi.org/10.1145/3427332

Xu, W., Liang, H.-N., Zhang, Z., & Baghaei, N. (2020). Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. *Games for Health Journal*, *9*(6), 405–414. https://doi.org/10.1089/g4h.2019.0102

Xu, X., Shi, H., Yi, X., Liu, W., Yan, Y., Shi, Y., Mariakakis, A., Mankoff, J., & Dey, A. K. (2020). EarBuddy: Enabling on-face interaction via wireless earbuds. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376836

Xu, X., Yu, C., Wang, Y., & Shi, Y. (2020). Recognizing unintentional touch on interactive tabletop. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(1), 1–24. https://doi.org/10.1145/3381011

Yang, Y., Karreman, J., & de Jong, M. (2020). Comparing the effects of paper and mobile augmented reality instructions to guide assembly tasks. *2020 IEEE International Professional Communication Conference (ProComm)*, 96–104. https://doi.org/10.1109/ProComm48883.2020.00021

Yi, H., Hong, J., Kim, H., & Lee, W. (2019). DexController: Designing a VR controller with grasp-recognition for enriching natural game experience. *25th ACM Symposium on Virtual Reality Software and Technology*, 1–11. https://doi.org/10.1145/3359996.3364263

Yu, D., Zhou, Q., Newn, J., Dingler, T., Velloso, E., & Goncalves, J. (2020). Fully-occluded target selection in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *26*(12), 3402–3413. https://doi.org/10.1109/TVCG.2020.3023606

Yuksel, B. F., Fazli, P., Mathur, U., Bisht, V., Kim, S. J., Lee, J. J., Jin, S. J., Siu, Y.-T., Miele, J. A., & Yoon, I. (2020). Human-in-the-loop machine learning to increase video accessibility for visually impaired and blind users. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 47–60. https://doi.org/10.1145/3357236.3395433

Yuksel, B. F., Kim, S. J., Jin, S. J., Lee, J. J., Fazli, P., Mathur, U., Bisht, V., Yoon, I., Siu, Y.-T., & Miele, J. A. (2020). Increasing video accessibility for visually impaired users with human-in-the-loop machine learning. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–9. https://doi.org/10.1145/3334480.3382821

Zargari Marandi, R., Fjelsted, C. A., Hrustanovic, I., Dan Olesen, R., & Gazerani, P. (2020). Eye movements in response to pain-related feelings in the presence of low and high cognitive loads. *Behavioral Sciences*, *10*(5), 92. https://doi.org/10.3390/bs10050092

Zeller, R., Williamson, A., & Friswell, R. (2020). The effect of sleep-need and time-on-task on driver fatigue. *Transportation Research Part F: Traffic Psychology and Behaviour*, *74*, 15–29. https://doi.org/10.1016/j.trf.2020.08.001