Psychological Research (2021) 85:2364–2397
https://doi.org/10.1007/s00426-020-01394-1

ORIGINAL ARTICLE

How do task characteristics affect learning and performance? The roles of simultaneous, interactive, and continuous tasks

David J. Frank[1,2] · **Brooke N. Macnamara**[1,2]

Received: 4 June 2019 / Accepted: 15 July 2020 / Published online: 3 August 2020

Abstract

Why are some tasks more difficult to learn than others? Hoffman et al. (Accelerated expertise: training for high proficiency in a complex world. Psychology Press, New York, 2014) hypothesized that certain task characteristics—termed "dimensions of difficulty"—hindered learning and performance. Previously, we tested two dimensions: consistent vs. variably mapped and static vs. dynamic. Here, we test three more dimensions of difficulty: sequential vs. simultaneous, discrete vs. continuous, and separable vs. interactive. In each study, we manipulate a single task feature (dimension of difficulty) while holding all others constant. Tasks with continuous (rather than discrete) features slowed participants' performance but did not impair learning. Learning and performance were unimpaired in tasks with interactive (rather than largely separable) processes. By contrast, we found strong evidence that simultaneous tasks (i.e., those that demand multitasking) inhibit learning, slow performance, and increase task errors. Importantly, this occurred in the absence of perceptual and mechanical bottlenecks present in most other studies of multitasking. We also are the first to examine simultaneity on learning a new task while controlling for other dimensions of difficulty. We discuss the potential impact of these results on current theory and application to real-world domains.

Introduction

When first learning a new task, engagement is typically effortful. Over time, many tasks become less demanding and task processing becomes more automatic (Fitts & Posner, 1967). However, some tasks remain difficult even following extensive practice (Ackerman & Cianciolo, 2002; Ackerman & Woltz, 1994). Why do some tasks become relatively easy with practice whereas others remain challenging and demanding?

We propose that task characteristics play an important role. Few studies have attempted to determine the various underlying task features that moderate the effect of practice on performance (Macnamara & Frank, 2018). One of the only task characteristics that has been investigated as a moderator of practice on performance is the consistent vs. variably mapped feature (e.g., Ackerman, 1987; Ackerman & Cianciolo, 2002; Ackerman & Woltz, 1994).

According to the performance–ability relations theory (Ackerman, 1987), when tasks have consistent features, practice has a large effect on performance because this consistency allows automatic processes in cognition to develop (see also Shiffrin & Schneider, 1977). Once automatic processes are in place, variability in cognitive resources is less associated with variance in task performance. In contrast, when tasks have inconsistent (variably mapped) features, cognitive ability has a large effect on performance because controlled processing is required to perform the task even after extensive practice.

To investigate the performance–ability relations theory, Ackerman and Woltz (1994) used the noun-pair lookup task. In this task, a participant verifies whether a pair of words in the center of the screen (e.g., "Ivy-Bird") are matched in a table containing multiple word pairs. When word pairs were consistently mapped—e.g., when "Ivy" was always paired with "Bird"—Ackerman and Woltz (1994) found that (a) performance improved dramatically with practice, and (b) performance time was initially correlated with fluid

David J. Frank and Brooke N. Macnamara contributed equally to this work.

✉ David J. Frank
David.Frank@TAMUC.edu

1 Department of Psychology and Special Education, Texas A&M University Commerce, Binnion Hall 218, P.O. Box 3011, Commerce, TX 75429, USA

2 Case Western Reserve University, Cleveland, OH, USA

intelligence and processing speed, but the relationships decreased to non-significant levels with practice. In contrast, when the task was variably mapped—when the word pairs were rearranged so that "Ivy" was not always paired with "Bird"—Ackerman and Woltz (1994) found that (a) performance improved minimally with practice and (b) performance time remained correlated with fluid intelligence and processing speed throughout the task, despite practice. In other words, under consistently mapped conditions, performance improved rapidly and became relatively automatic, whereas under variably mapped conditions, performance improved minimally and remained effortful.

When the concept of consistent and variable mapping was extended to a complex air traffic control task, Ackerman and Cianciolo (2002), no longer found this pattern of results: they found that even under consistently mapped conditions the task remained effortful. Why might this be the case? It might be that the consistent vs. variably mapped dimension is not the only dimension at play.

Indeed, in 2014, Hoffman et al. proposed eight "dimensions of difficulty"—dimensions whose task features on one end of the continuum were theorized to increase learning difficulty via additional recruitment of cognitive resources (see Table 1). For example, Hoffman, Ward, Feltovich, DiBello, Fiore and Andrews, (2014) theorized that tasks with dynamic features were more difficult than tasks with static features, and that tasks with simultaneous processes were more difficult than tasks with sequential processes.[1]

Ackerman and Woltz's (1994) noun-pair task manipulated consistent vs. variably mapped characteristics, and all other task characteristics happened to fall on the theoretically "easy" side of each of Hoffman et al. (2014) dimensions, i.e., it was also static, with sequential and separable processes, and discrete features, etc. However, the air traffic control task used by Ackerman and Cianciolo (2002), even when consistently mapped, fell on the more difficult side of many of Hoffman et al's. (2014) dimensions. It was dynamic (rather than static): the aircrafts were constantly moving. It was simultaneous (rather than sequential): it involved concurrently managing takeoffs and landings. It was interactive (rather than separable): where you decided to land one aircraft had implications for all other takeoffs and landings. It was continuous (rather than discrete): there was a full range of altitudes and airspeeds. Thus, any one of these dimensions could be responsible for the continued cognitive difficulty of Ackerman and Cianciolo's (2002) consistently mapped air traffic control task.

These dimensions, along with Ackerman's (1987) consistent vs. variably mapped dimension, are relevant to all real-world tasks. That is, every task falls somewhere on the continua of consistent to variably mapped, static to dynamic, sequential to simultaneous, etc. Using a complex task so as to better mimic real-world tasks, we set out to examine the impact of five task features on learning and performance:

1 Ackerman's (1987) consistent vs. variable mapping dimension,
2 Hoffman et al.'s (2014) static vs. dynamic dimension,
3 Hoffman et al.'s (2014) sequential vs. simultaneous dimension,
4 Hoffman et al.'s (2014) separable vs. interactive dimension, and
5 Hoffman et al.'s (2014) discrete vs. continuous dimension.

The first two experiments, conducted 2016–2017, examined Ackerman's (1987) consistent vs. variable mapped dimension and Hoffman et al.'s (2014) static vs. dynamic dimension. We (Macnamara & Frank, 2018) found that variable mapping hindered learning and performance: For participants in the variably mapped conditions, performance was lower, improvements with practice were smaller, and efficiency was reduced compared to participants in the consistent conditions. We also found that dynamic conditions hindered learning and performance: For participants in the dynamic conditions, performance was substantially lower and improvements were smaller with practice compared to participants in the static conditions. Until now, this was the only study to experimentally isolate and test any of Hoffman et al. dimensions of difficulty.

In the present set of studies, conducted 2017–2019, we tested Hoffman et al.'s sequential vs. simultaneous dimension (Experiment 1), separable vs. interactive dimension (Experiment 2), and discrete vs. continuous dimension (Experiment 3). We use the same task paradigm used in Macnamara and Frank (2018), in which we compare performance, learning, and efficiency on two versions of the task. We also manipulate cognitive load within each task version. In each experiment, we hypothesize that the theoretically more difficult task characteristic (simultaneous, interactive, continuous) will impair learning and performance relative to the easy task characteristic directly or via interactions with cognitive load and that the more difficult task characteristic will decrease task efficiency.

[1] Hoffman et al. (2014) based these dimension on research by Feltovich, Spiro, and Coulson (1989, 1993); see also Dawson-Saunders, Feltovich, Coulson, and Steward, (1990), who interviewed medical school instructors and asked them to identify and describe which tasks medical student learned with the least and most difficulty.

Table 1 Hoffman et al.'s (2014) dimensions of difficulty

Dimension	Easier Characteristic	More Difficult Characteristic
Static vs. dynamic	Static: Important aspects of the task can be captured in "snapshots" Example: Scrabble, where players can pause to make decisions without relevant information changing	Dynamic: Important aspects of the task are in flux Example: Pacman (Namco, 1980), where the ghosts continue to move even if the player pauses
Sequential vs. simultaneous	Sequential: Processes occur one at a time Example: Baseball, where fielding and batting occur sequentially and a player is never engaged in both at once	Simultaneous: Processes occur concurrently Example: Piloting an aircraft, where one must control the speed, altitude, direction, and monitor gauges concurrently
Separable vs. interactive	Separable: Processes occur independently or with weak interactions Example: Games like Whac-A-Mole and Skeeball, where the player responds to stimuli independently of one another	Interactive: Processes are strongly interdependent Example: Chess, where each move can impact future moves
Discrete vs. continuous	Discrete: Task attributes are characterized by a small number of categories Example: Checkers, where there are discrete categories for each piece (black and red, crowned and uncrowned) and discrete positions on the board	Continuous: Task attributes are characterized by a continuum of features or a large number of categorical distinctions Example: Golf, where ball distance, ground slopes, and strike velocity fall on continua
Linear vs. nonlinear	Linear: Relationships among features are proportional and can be conveyed with a single line of explanation Example: Baseball pitching, where the faster a pitcher moves his arm, the faster the pitch	Nonlinear: Relationships among features are nonproportional and require multiple lines of explanation Example: Tennis, where the relationship between the angle that a ball should be hit and distance from the net is nonlinear
Single representations vs. multiple representations	Single representations: Task components have one or very few interpretations or uses Example: Solitaire, where each card can only serve as a single numeric value-suit combination	Multiple representations: Task components have multiple interpretations and uses depending on context Example: Sports video games, where the same button can have different functions depending on context (such as offense or defense)
Homogenous vs. heterogeneous	Homogenous: Components and conceptual representations are uniform across a system (i.e., there is a single explanation) Example: First-person shooter video games, where a player's health decreases because he has taken damage from an enemy and that is the only reason health would decrease	Heterogenous: Components and conceptual representations are diverse Example: Medical diagnostics, where there are often multiple potential explanations for a single problem. E.g., shortness of breath could be due to asthma, heart failure, a panic attack, or carbon monoxide poisoning

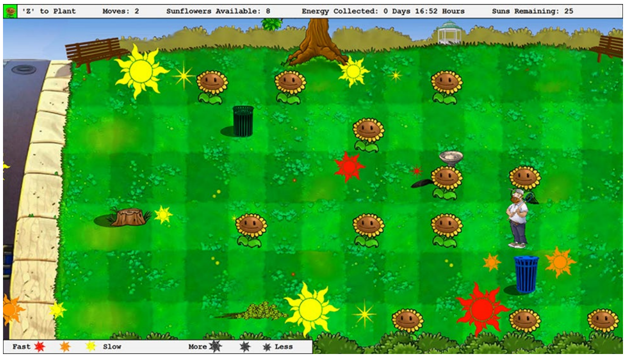

Fig. 1 Screenshot of an energy collection mission (baseline version). The small "starbursts" are sunlight energy (lumens) that travel from a sun to a sunflower on its right

Fig. 2 Screenshot of a zombie fighting mission (baseline version). The circles are the pea shots

General experimental paradigm

In each experiment, we compared a baseline version of the task to a manipulated version of the task. In the baseline version, all characteristics are kept on the theoretically easy side of each dimension. Each manipulated version is identical to the baseline version except for a single characteristic manipulated to be on the theoretically more difficult side of that dimension. Thus, any differences we find within an experiment should be due to the impact of the one manipulated task characteristic. Likewise, because the paradigm and baseline version is consistent across experiments, any differences we find across experiments should be due to the relative importance of the manipulated dimension. This consistency in paradigm controls for confounds that typically arise when changing task paradigms between experiments.

General paradigm: plants vs. zombies task

We use a task that is visually similar to the commercial game by the same name but contains a novel set of rules and relationships (videos of the task can be downloaded from https://osf.io/f4auv/). In this task, participants engage in two missions per round, (1) to collect energy, and (2) to fight zombies. Participants complete five rounds of the task.

Energy collection missions

In the energy collection missions, the goal is to collect as much energy for the neighborhood as possible. To collect energy, participants strategically move an avatar to plant sunflowers that collect energy (lumens) from moving suns (see Fig. 1). Participants make two actions before the game takes an action. Participants must make at least three moves (using the arrow keys) between sunflower planting actions (using the "z" key).

Suns Suns move across the screen from left to right and come in one of three sizes and one of three colors: red, orange, and yellow. The larger the sun, the more energy its lumens and the redder the sun, the faster it moves across the screen. The size and color mappings are visually displayed in the lower left part of the screen throughout the energy collecting missions. There are 27 total suns (three of each size-color combination) in each energy collection mission with a maximum of 16 suns on the screen at a time. The position and time at which each of the suns emerges vary across rounds. Arrangement order is counterbalanced across participants. Difficulty is equivalent across all arrangements. When the final sun disappears from the right side of the screen the mission ends.

Planting The participants plant sunflowers to the right of a sun to collect energy from it. Suns destroy sunflowers as they pass over them, thus resulting in frequent replanting. This, combined with participants only being able to plant every fourth move, requires them to strategically prioritize their planting to maximize energy collection from the "best" suns available (typically those currently farther to the left, larger, and slower). See Macnamara and Frank (2018) for more details.

Zombie fighting missions

In the zombie fighting missions, the goal is to prevent zombies from escaping by killing them before they travel across the screen. To fight zombies, participants strategically plant peashooters to attack approaching zombies (see Fig. 2). Participants make two actions before the game takes an action. Participants must make at least three moves (using the arrow keys) between peashooter planting actions (using the "z"

Table 2 Baseline version characteristics

Characteristic	Description
Consistent	The rules for the stimuli are the same throughout the task.
Static	The task is “turn-taking” such that the game does not advance until the participant has executed actions.
Sequential	Each mission occurs consecutively.
Discrete	Stimuli are categorical.
Separable	Each process involved in a mission—moving the avatar and planting plants—is largely independent.
Linear	Size to sun energy/zombie toughness mappings are mathematically linear, as are color to speed mappings.
Single representation	Each stimulus has one meaning per mission.
Mechanistic	The task is constructed of individual parts. Causal agents are direct.
Homogenous	Task components (key mapping, score displays, planting counters, movement) are identical or nearly identical across missions and rounds.

key). Participants can only have twenty peashooters on the screen at once. See Fig. 2.

Zombies Zombies come in one of three sizes and three (clothing) colors. The larger the zombie, the more hits are required from the peashooters to kill it and the redder the zombie’s clothing, the faster it moves across the screen. Killing larger zombies also earns more points.

There are 27 total zombies (three of each color-size combination) in each zombie fighting mission with a maximum of 16 zombies on the screen at a time. The position and time at which each of the zombies emerges vary across rounds. Arrangement order is counterbalanced across participants. Difficulty is equivalent across all arrangements. When the final zombie is killed or escapes by reaching the left side of the screen, the mission ends.

Planting The participants plant peashooters to the left of a zombie to attack it. Zombies destroy peashooters as they walk over them, thus resulting in frequent replanting. This, combined with participants only being able to plant every fourth move, requires them to strategically prioritize their planting to maximize attacks on the most difficult-to-kill zombies (the larger and faster ones). Participants can also uproot planted peashooters to return the plant to their resources. However, this comes at a time cost—participants must wait another three moves before they can plant after uprooting. Thus, participants must plant strategically to maximize their score. See Macnamara and Frank (2018) for more details.

Baseline version task characteristics

In the baseline version of the task, all characteristics fall on the easier side of every dimension of difficulty (see Table 2). For additional details, see Macnamara and Frank (2018) and https://osf.io/yzxer/.

Experiment 1: sequential vs. simultaneous

Hoffman et al. (2014) suggest that simultaneous tasks are more likely to hinder learning and performance than sequential tasks. This is the one hypothesized dimension of difficulty with a long history of investigation in terms of impacts on performance, but not learning. That is, a number of studies have examined how performing two tasks simultaneously (multitasking; e.g., Drews, Pasupathi, & Strayer, 2008) or frequently switching between two task goals (task switching; e.g., Monsell, 2003) influences performance. The general finding is that performance is slower and/or less accurate on one or both tasks (Monsell, 2003; Rubinstein, Meyer, & Evans, 2001). While task switching and multitasking have been extensively studied, both paradigms have significant limitations for examining the impact of simultaneity on learning.

Task switching

A large literature on task switching has shown that alternating between two tasks results in a switch cost (see Vandierendonck, Liefooghe, & Verbruggen, 2010 for a meta-analysis and review). When switching from one task goal to another (a switch trial), there is a cost in time and attention as the previous task goal is suppressed and the new task goal is activated. Mixing costs also occur. Mixing costs represent the cost of simply maintaining two different task goals/procedures in memory. Mixing costs are evident in longer response times even when the same goal is repeated on successive trials (non-switch trials) as compared to when a task is purely sequential without any switch trials.

Limitations of task-switching paradigms

Simplicity The first limitation of task-switching paradigms is that most task-switching paradigms are relatively simple. Learning occurs only in the first few trials. This simplicity severely limits insights task-switching paradigms can provide on how simultaneity affects task learning.

Mandatory task switching The second limitation of most task-switching paradigms is that the paradigm dictates, via a cue, when a given task should be performed. Studies do not typically allow for *strategic* task switching—allowing participants to decide when to switch and how to divide attentional resources. Strategic task switching would be most applicable to complex, real-world tasks. For example, researchers might switch from writing up one study to planning another depending on the amount of time they currently have to work, how cognitively demanding each of those tasks are, and if someone (e.g., a collaborator) is waiting on them to make progress on one of these tasks more than another. Task switching paradigms, which do not allow for strategic switch decisions, are therefore limited in the insight they can provide on how simultaneity affects learning to task switch strategically in complex tasks.

Confounds The third limitation of task-switching paradigms is that response mappings are often confounded with other dimensions of difficulty, namely variable mapping and multiple representations: In a typical task-switching paradigm, the participant sees a stimulus and makes one of two judgements. For example, they see a number and must either make an odd/even judgment or judge whether the number is greater than or less than five, depending on which task is cued (e.g., Rubinstein et al., 2001). If "odd" and "less than five" responses are mapped to the "z" key and "even" and "greater than five" are mapped to the "x" key, then the appropriate response to "9" differs depending on the task condition while the appropriate response to "3" is the same in both task conditions.

As we (Macnamara & Frank, 2018; see also Ackerman & Woltz, 1994) recently demonstrated, variable mapping alone negatively influences learning and performance. Likewise, using the same key to represent different responses is theorized to hinder performance (Hoffman et al., 2014). Thus, in many task-switching paradigms, we cannot know the degree to which performance decrements are due to variable mapping and/or multiple representations versus switching.

Multitasking Research suggests that multitasking itself may be a misnomer as people often cannot engage in two tasks simultaneously, but rather switch between the two (Borger, 1963; Creamer, 1963). However, multitasking paradigms often differ substantially from task-switching paradigms. Multitasking paradigms tend to make use of more complex tasks that are closer to real-life situations than task-switching paradigms, such as distracted (e.g., by texting) driving, air traffic control simulation, and dichotomous listening. However, a closer look at these studies reveals a number of perceptual bottlenecks, mechanical bottlenecks, and confounds that could be contributing to the lower performance found in multitasking compared to the single-task paradigms used across multitasking studies.

Limitations of multitasking paradigms

Perceptual bottlenecks In many multitasking paradigms, participants are asked to monitor multiple stimuli at once (e.g., Hambrick, Oswald, Darowski, Rench, & Brou, 2010; Strayer & Drews, 2007; Wood & Cowan, 1995). In the case of visual stimuli, simultaneously monitoring objects far apart in the field of view may be impossible due to the limited size of the fovea (Henderson & Ferreira, 1990). For example, it is impossible to simultaneously monitor texting accuracy and driving conditions due to the inability to fixate on the road and one's phone at the same time. These physical perceptual limitations create a ceiling where someone cannot possibly perform as accurately on multiple tasks as they could when focusing on a single task. Thus, in many multitasking studies, we cannot know the degree to which multitasking performance declines due to perceptual limits versus cognitive processing difficulty.

Mechanical bottlenecks Similar to perceptual bottlenecks, many multitasking paradigms include mechanical bottlenecks where the task involves incompatible responses that must be carried out simultaneously or in quick succession (e.g., Bratfisch & Hagman, 2003; Strayer & Drews, 2007). For example, one cannot steer with two hands while simultaneously texting, nor can one verbally redirect two aircraft simultaneously in an air traffic control simulation task. Thus, in many multitasking studies, we cannot know the degree to which multitasking performance declines due to human mechanical limits versus cognitive processing difficulty.

Confounds Many multitasking paradigms involve a dynamic environment (e.g., Bühner, König, Pick, & Krumm, 2006; Caird, Simmons, Wiley, Johnston, & Horrey, 2018; Drews et al., 2008; Strayer & Drews, 2007). Take, for example, texting and driving: While texting is a relatively static task (one can usually pause as frequently as one wishes while typing), driving is a highly dynamic task (the vehicle is moving, other vehicles are moving, street lights change color, etc.). Likewise, tasks that involve verbal language processing (e.g., dichotic listening tasks: Conway, Cowan, & Bunting, 2001; Wood & Cowan, 1995; using a laptop during a lecture: Hembrooke & Gay, 2003) are dynamic because the information continues to change and must be continuously attended for comprehension.

As we (Macnamara & Frank, 2018) recently demonstrated, dynamic task environments alone produced substantially lower task performance and hindered learning, even when no multitasking was involved. Thus, in many multitasking studies, we cannot know the degree to which multitasking performance declines due to the dynamic characteristic of the task versus the simultaneous characteristic of the task.

Benefits of the plants vs. zombies task

The Plants vs. Zombies task allows for the examination of how maintaining simultaneous goals and attempting to carry them both out concurrently impacts learning and performance. This paradigm eliminates the bottlenecks and confounds found in task switching tasks and multitasking tasks (described in more detail under the heading “simultaneous version”). If we find that simultaneity hinders performance after eliminating known perceptual and mechanical bottlenecks and confounds, we can rest assured that accumulated knowledge attained from task switching and multitasking research are not due to these extraneous factors. If we do not find an effect of simultaneity, this suggests an extraneous factor or interaction might be responsible for previous effects.

In addition to effects on performance, the present experiment provides a unique opportunity to examine simultaneity on learning. Task switching paradigms are generally too simple to examine learning with practice. Multitasking studies typically examine situations where at least one of the tasks is relatively simple or well-learned (e.g., texting and driving; Caird et al., 2018; Drews et al., 2008; Strayer & Drews, 2007). We examine simultaneity on learning in a complex, novel task, where participants have the opportunity to increase their performance with practice.

Methods

This study and all analyses, stopping rules, and exclusion criteria were formally pre-registered via the Open Science Framework (https://osf.io/a4dr3/register/5730e99a9ad5a102c5745a8a).

Participants

Our stopping rule was to collect 120 participants (divided roughly evenly among conditions). Data collection was terminated at the end of the week following the 120th participant. A total of 135 students enrolled in psychology classes at Case Western Reserve University participated in exchange for partial course credit or extra credit. This study was approved by the Institutional Review Board at Case Western Reserve University.

Table 3 Baseline vs. simultaneous task version characteristics

Baseline version characteristics	Simultaneous version characteristics
Consistent	Consistent
Static	Static
Sequential	**Simultaneous**
Separable	Separable
Discrete	Discrete
Linear	Linear
Single representation	Single representation
Mechanistic	Mechanistic
Homogenous	Homogenous

Fig. 3 Screenshot of the simultaneous version where participants both collect energy and fight zombies. The colored box indicates which avatar is currently selected

Plants vs. zombies task

The program was created and administered via E-Prime 2 (Schneider, Eschman, & Zuccolotto, 2002) on computers at a resolution of 1920 × 1080.

Baseline version The baseline version, described above, is sequential. Participants complete the energy collection missions and the zombie fighting missions consecutively in each round.

Simultaneous version The simultaneous version is identical to the baseline version (see Table 3, non-bolded characteristics), except that participants complete both missions simultaneously (see Table 3, bold characteristics). Two avatars, one for planting sunflowers to collect energy and one for planting peashooters to fight zombies are present on the screen (see Fig. 3). Participants toggle between the avatars by pressing the “X” key. Suns move from left to right and zombies move from right to left as they do on all task versions.

The plants vs. zombies task eliminates perceptual bottlenecks often found in multitasking paradigms.

Participants can take the time to fixate and re-fixate relevant task information whenever needed. The plants vs. zombies task eliminates mechanical bottlenecks often found in multitasking paradigms. Participants toggle between the avatars: one for energy collection and one for zombie fighting. To switch between the two avatars, the participant presses the "X" key. This does not count as a move and the switch is nearly instantaneous.

The plants vs. zombies task eliminates confounds often found in task switching and multitasking paradigms. The task is consistently mapped and has single representations rather than being variably mapped with multiple representations (as opposed to many task-switching paradigms). The task is static (turn taking) rather than dynamic (as opposed to many multitasking paradigms).

Unlike in the baseline version where participants have two actions before the game stimuli move, participants in the simultaneous version have four actions before the game stimuli move. Because both suns and zombies move at the same time, participants have the same number of total actions per round as participants in other task versions (though not divided by two missions). If participants switched avatars evenly between turns (i.e., every two actions), they would perform identically as they would in the baseline version of the task.

However, despite an equal number of actions, participants in the simultaneous version have a slight potential advantage because they can strategically focus efforts when one mission needs more attention than the other at any given time. Given the ability to obtain identical, or slightly higher, scores than in the baseline version, any decrease in learning and performance should be purely due to increased cognitive demands.

Cognitive load manipulation All participants wore noise-cancelling headphones. Half the participants in each task version were assigned to perform the task under cognitive load. Participants in these conditions heard beeps in their headphones at semi-random intervals no fewer than two seconds apart and no greater than six seconds apart. Participants were instructed to mentally rehearse and update letters of the alphabet each time they heard a beep beginning with the letter "A" at the first beep, "B" at the second beep and so on. If they reached "Z," they were instructed to begin the alphabet again. At the end of each mission, participants typed the last letter they had mentally rehearsed during that mission.

Procedures

After providing informed consent, participants read the plants vs. zombies task instructions and completed five rounds of either the baseline version or simultaneous version. Immediately after the task, participants rated their task experience in terms of task interest, fun, tediousness, difficulty, frustration, engagement, and fatigue. Participants in the load conditions also indicated how they responded on the letter counting task if they forgot which letter they were on (see Supplemental Materials). Participants completed all activities at their own pace. Sessions lasted between 45 and 180 min with most participants taking approximately 90 min.

Design and analyses

The experiment was a 2 (task version: baseline, simultaneous) × 2 (cognitive load: load, no load) × 2 (mission: energy collection, zombie fighting) × 5 (Round: 1–5) mixed design with task version and cognitive load as between-subject factors. Participants were assigned to the between-subject conditions via counterbalancing.

Changes in performance are often non-linear (Fitts, 1964). To account for this, we used SAS Proc MIXED (Littell, Milliken, Stroup, & Wolfinger, 2000) to analyze learning curves across task conditions. Specifically, we analyzed performance scores and completion times, each with the following variables as predictors: round (1–5), round squared (1–5), task version (baseline, simultaneous), cognitive load (no load, load), mission (energy collection, zombie fighting) and all possible interaction terms except for round X round squared. The baseline no load group (i.e., those in the theoretically easiest condition) served as the reference group. For regression weights and all fixed effect results, see Appendices 1 and 2.

Results

We predicted that the simultaneous version would be more cognitive demanding and ultimately lower performance, hinder learning, and make participants less efficient either directly or by interacting with cognitive load.

Performance scores

We first removed outliers in accordance with our pre-registration. We removed any round where a participant in the simultaneous version was not generally multitasking but instead spent 85% or more of their moves on a single avatar. If a participant had more than one such round, that participant was removed from analyses entirely. This resulted in four participants and ten additional observations (five energy collection scores and their corresponding zombie fighting scores) being removed (approximately 1% of the data from the simultaneous version). Next, we removed any outliers that were three standard deviations below the mean for their task version on that particular round. This resulted in eight zombie fighting scores (one baseline no load, five baseline load, one simultaneous no load, one simultaneous load) and six energy collection scores (four baseline no load, two baseline load) being removed. Both criteria combined removed approximately 6% of the overall data.

We next standardized performance scores based on round 1 baseline no load group performance. Thus, each score is the number of standard deviations above or below the mean score for the first round for the theoretically easiest condition.

Main effects of round, $F(1, 1055) = 35.55$, $p < 0.001$, and round squared, $F(1, 1055) = 9.34$, $p = 0.002$, indicated better performance over rounds with a negative exponential curve following typical learning curves (Fitts, 1964).

A main effect of task version, $F(1, 117) = 7.83$, $p = 0.006$, indicated worse performance on the simultaneous version than on the baseline version.

A marginally significant main effect of the mission, $F(1, 117) = 4.60$, $p = 0.034$, occurred from participants performing slightly better in the zombie fighting missions. We also observed task version X mission, $F(1, 117) = 8.34$, $p = 0.005$, and cognitive load X mission, $F(1, 117) = 5.85$, $p = 0.017$, interactions. To follow up on these interactions, we conducted separate models for each of the two missions: energy collection and zombie fighting. Figure 4 shows standardized mean performance data over rounds (upper panel) and separated by mission (lower panels).

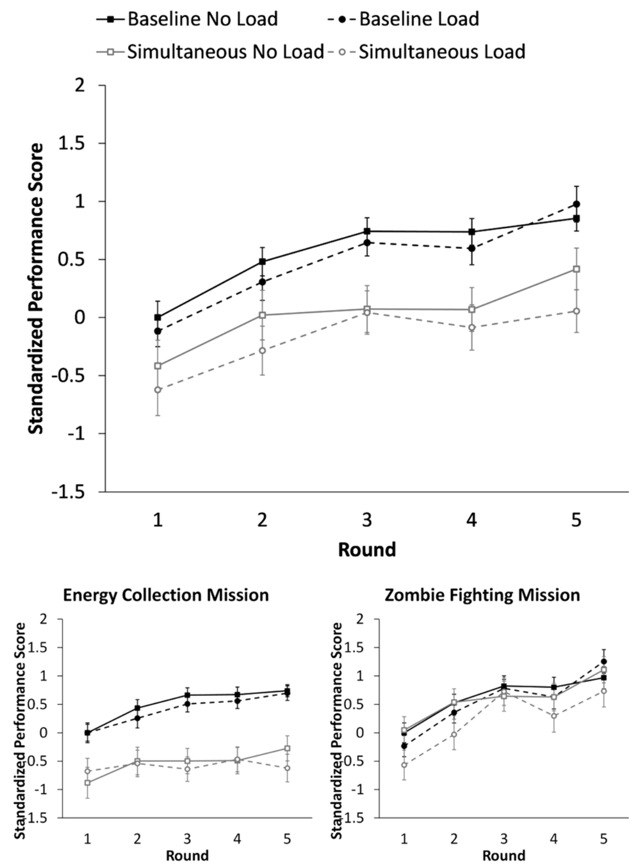

Fig. 4 Experiment 1 standardized mean performance on the plants vs. zombies task. Performance data are standardized around round 1 performance in the baseline no load group, which has a *z*-score of 0. A *z*-score of 1/−1 indicates performance one standard deviation above/below the round 1 baseline no load group mean. Error bars represent standard errors of the mean

Energy collection scores

Main effects of round, $F(1, 1067) = 43.15$, $p < 0.001$, and round squared, $F(1, 1067) = 15.88$, $p < 0.001$, indicated better performance over rounds with a negative exponential curve following typical learning curves (Fitts, 1964). There was a negative cognitive load X round interaction, $F(1, 1067) = 5.92$, $p = 0.015$, which resulted from reduced improvements with practice under cognitive load compared to no cognitive load.

We observed a main effect of task version, $F(1, 117) = 19.58$, $p < 0.001$, resulting from poorer performance in the simultaneous version compared to the baseline version. A task version X round interaction, $F(1, 1067) = 7.83$, $p = 0.005$, resulted from reduced improvements with practice in the simultaneous version compared to the baseline version. Thus, in the simultaneous version, learning and performance were hindered.

Participants in the simultaneous version devoted fewer moves ($M = 352$, SD = 95.08) to energy collection than those in the baseline version ($M = 385$, SD = 40.39), $t(70.41) = -3.67$, $p < 0.001$, $d = -0.45$. Despite the difference, participants were still devoting hundreds of moves to the energy collection missions. However, while participants in the baseline version improved with practice, participants' in the simultaneous version showed little to no improvement with practice.

Zombie fighting scores

Main effects of round, $F(1, 1067) = 60.51$, $p < 0.001$, and round squared, $F(1, 1067) = 13.05$, $p < 0.001$, indicated better performance over rounds with a negative exponential curve following typical learning curves (Fitts, 1964).

Contrary to energy collection scores, for zombie fighting scores, we did not observe a main effect of task version, $F(1, 117) = 0.69$, $p = 0.407$, but did observe a main effect of cognitive load, $F(1, 117) = 6.25$, $p = 0.014$. Thus, performance on the zombie fighting missions was slightly lower under cognitive load but otherwise improved similarly across task versions.

However, participants in the simultaneous version devoted substantially more moves ($M = 579$, SD = 144.26) to zombie fighting than participants in the baseline version

($M = 426$, SD = 61.19), $t(72.92) = 8.70$, $p < 0.001$, $d = 1.67$. Despite devoting substantially more moves to these missions, participants in the simultaneous version did not outperform their baseline counterparts on zombie fighting missions.

Efficiency

As with performance scores, we removed the completion time outliers in accordance with our pre-registration. The criteria that participants could not have spent greater than 85% of their moves on a single avatar removed 1% of the data in the simultaneous version (the same data removed from performance scores). In addition, nine observations (three baseline no load, two baseline load, four simultaneous no load) were removed because they were more than three standard deviations longer than the block mean for their task version. In total, these criteria removed approximately 2% of the overall data.

Performance times are in minutes. Participants in the simultaneous version had to press the "x" key to switch between avatars (which adds a relatively small amount of time). Additionally, participants in the simultaneous version end up with more moves at the end of each round after the last zombie is killed, but before the last sun has left the screen. The time taken for these additional moves is accounted for (removed) for all efficiency analyses so as not to inflate the effects.

Main effects of round, $F(1, 469) = 168.51$, $p < 0.001$, and round squared, $F(1, 469) = 56.70$, $p < 0.001$, indicate faster performance over rounds with a positive exponential curve.

A main effect of task version resulted from slower completion times for the simultaneous version compared to the baseline version, $F(1, 117) = 97.23$, $p < 0.001$ (Fig. 5). Because completion times cannot be separated by the mission in the simultaneous version, we drop that predictor for this analysis only. Figure 5 shows mean completion times over rounds in minutes.

Discussion

We found strong evidence that performing two tasks simultaneously impaired learning, hindered overall performance, and decreased task efficiency. These results are attributable to cognitive difficulty rather than perceptual or mechanical bottlenecks. It is important to note that the increased perceptual load in the simultaneous version (visual clutter) likely slows visual search and could lead to increased errors (e.g., a participant overlooks a sun or zombie in their search). However, this should have been at least partially mitigated by the fact that participants had no time constraints and could spend as much time as they wanted on visual search. Thus, while our task eliminates the perceptual bottleneck created by many dynamic simultaneous tasks, part of the cognitive difficulty of the task may still be due partly to perceptual load. Importantly, these results were obtained while controlling for other known (dynamic and variable mapping) and theoretical (interactive, continuous, non-linear, multiple representations, organic, and heterogeneous) dimensions of difficulty.

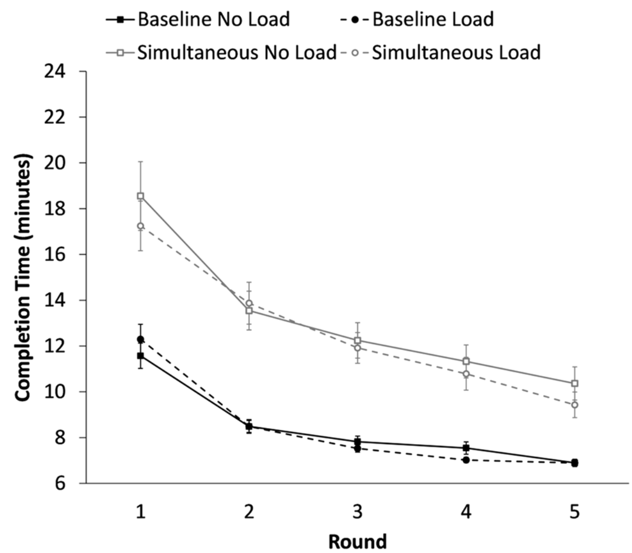

Fig. 5 Experiment 1 mean task completion times in minutes for the energy collection missions and zombie fighting missions combined. Error bars represent standard errors of the mean

Interestingly, participants in the simultaneous version were biased in the number of moves they devoted to each mission. They devoted slightly fewer moves toward the energy collection missions than their baseline counterparts (9% fewer moves); despite this relatively small amount, their performance on the energy collection missions was substantially worse than those in the baseline version. Conversely, participants in the simultaneous version devoted substantially more moves toward the zombie fighting missions (36% more moves), yet they did not outperform their baseline counterparts on these missions.

There are a number of reasons participants may have devoted more resources to the zombie fighting missions. An escaping zombie may be viewed as a loss, in contrast to the suns which pass regardless of how much energy was collected from them. People typically show greater sensitivity to losses than gains (Kahneman & Tversky, 1979). Although participants were told to maximize their performance on both missions, stopping zombies from escaping may feel more critical than collecting energy. Finally, zombie fighting might be more intuitive, fun, or engaging than energy collection.

Although the missions of zombie fighting and energy collection appear similar, the ideal strategies for each differ

Psychological Research (2021) 85:2364–2397

drastically. Zombie fighting is a combination of proactive and reactive maneuvers. The ideal strategy involves keeping a minimum number of peashooters in each row, then adding more as the situation demands. The strategy is then improved upon by learning just how many peashooters each zombie requires (depending on speed and toughness) and learning to strategically place them so that too many peashooters are not tied up in a single row after a zombie attack (which then requires one to waste valuable time and resources to uproot the extra peashooters). By contrast, energy collection is largely reactive. The ideal strategy involves prioritizing the largest, slowest suns as they appear, and replenishing sunflowers to continue collecting energy from these "best" suns. Although the strategy sounds simpler, it requires constant vigilance and is less similar to typical tower defense games. Thus, the strategy may require more attention to both discover and execute well relative to the more proactive zombie fighting strategy.

Participants in the simultaneous version performed both worse and more slowly than participants in the baseline version. Therefore, the results cannot be accounted for by a speed-accuracy tradeoff. While participants in the simultaneous version needed to press a key to toggle between avatars, the time taken to press the switch key cannot solely account for the substantially slower completion times relative to participants in the baseline version.

This experiment was the first to examine how learning and performance are impacted by simultaneous goal management, while controlling for perceptual and mechanical bottlenecks and other dimensions of difficulty, on a complex task. Managing two goals simultaneously, even in the absence of bottlenecks and confounds, impaired learning, performance, and efficiency.

Experiment 2: separable vs. interactive

Separable tasks are those where one response has little or no impact on future decisions and where performance on one aspect of the task does not influence performance on another. Many simple tasks are separable and lack the potential for interactivity because there is only one response per trial and because trials are not contingent upon one another. For example, there is only one response per trial on the noun-pair look-up task, thus one does not need to determine which stimuli need to be dealt with first or consider how one's response will impact the next response.

In contrast, many complex tasks have interactions that are difficult to eliminate. In Ackerman and Cianciolo's (2002) air traffic control task, one aircraft's route has implications for other aircrafts' routes. Likewise, assigning an aircraft to land or take off temporarily ties up a runway that other aircraft cannot use. To remove these interactions would completely alter the task: It is nearly impossible to fail at an air traffic control task if aircraft are allowed to pass through one another, eliminating the need to coordinate decisions.

In this experiment, we test whether interactivity in a complex task, isolated from other dimensions of difficulty (i.e., in a static environment, where missions are sequential, etc.), impacts learning and performance. To the best of our knowledge, this is the first experiment designed to test whether interactivity hinders learning and performance as hypothesized by Hoffman et al. (2014).

Methods

This study and all analyses, stopping rules, and exclusions criteria were pre-registered via the Open Science Framework (https://osf.io/ygwq5/register/5730e99a9ad5a102c5745a8a).

Participants

Our stopping rule was to collect 120 participants (roughly evenly divided among conditions). Data collection was terminated at the end of the week following the 120th participant. A total of 122 participants—60 students enrolled in psychology classes at Case Western Reserve and 62 students enrolled in psychology classes at University or Texas A&M University–Commerce—participated in exchange for partial course credit or extra credit. This study was approved by Institutional Review Boards at Case Western Reserve University and Texas A&M University–Commerce.

Materials

The program was administered via E-Prime 2 (Schneider et al., 2002) on computers at a resolution of 1920 × 1080.

Plants vs. zombies task: baseline version

The baseline version was identical to the baseline version used in Experiment 1. As described previously, the baseline version of the task is separable. Participants must decide which stimuli to prioritize, but interactions are otherwise kept to a minimum and performance on one part of the task does not influence performance on subsequent attempts. Although a participant must move efficiently to maximize their planting, whether they move up, down, left, or right does not influence how often they can plant. Once participants in the baseline version plant a sunflower or peashooter they must move three more times before they can plant again. A countdown timer shows the number of moves before they can plant again. The time counts down from three, advancing with each move.

Table 4 Baseline vs. interactive task version characteristics

Baseline version characteristics	Interactive version characteristics
Consistent	Consistent
Static	Static
Sequential	Sequential
Separable	**Interactive**
Discrete	Discrete
Linear	Linear
Single representation	Single representation
Mechanistic	Mechanistic
Homogenous	Homogenous

Plants vs. zombies task: interactive version

The Interactive Version is identical to the baseline version (see Table 4, non-bolded characteristics), except that it is interactive rather than separable (see Table 4, bold characteristics). To increase the interactivity between the two main processes the participant engages in—moving into position and planting—we made the availability of the plants contingent on the direction the participant moves. In the interactive version, the timer counts down from six, and advances by three whenever the player moves left or right, and advances by one whenever the player moves up or down. Thus, participants must coordinate their movement and planting strategies to maximize performance.

Cognitive load

The cognitive load manipulation was identical to that of Experiment 1.

Procedures

The procedures were identical to Experiment 1, except participants were assigned to either the baseline version or the interactive version.

Design and analyses

The design and analyses were identical to Experiment 1, except that the interactive version of the task was compared to the baseline version. For regression weights and all fixed effect results, see Appendices 3 and 4.

Results

We predicted that the Interactive Version would require more effortful decision making and ultimately lower performance, hinder learning, and decrease efficiency, either directly or by interacting with cognitive load.

Performance scores

As in Experiment 1, we first removed outliers in accordance with our pre-registration. This resulted in seven zombie fighting scores (one baseline no load, two baseline load, four interactive no load) and one energy collection score (interactive load) being removed. In total, these criteria removed < 1% of the overall data.

We next standardized performance scores based on round 1 baseline no load group performance. Thus, each score is the number of standard deviations above or below the mean score for the first round for the theoretically easiest condition.

We observed main effects of both round, $F(1, 1078) = 55.51$, $p < 0.001$, and round squared, $F(1, 1078) = 18.49$, $p < 0.001$, indicating better performance over rounds with a negative exponential curve following typical learning curves (Fitts, 1964).

We did not observe a main effect of task version, $F(1, 118) = 0.68$, $p = 0.411$. A main effect of mission resulted from higher scores on the zombie fighting missions, $F(1, 118) = 16.65$, $p < 0.001$. A marginally significant task version X cognitive load X mission interaction was observed, $F(1, 118) = 4.00$, $p = 0.048$, due to the baseline no load group performing slightly better than the other groups, but limited to the energy collection missions. Figure 6 shows standardized mean performance data over rounds (upper panel) and for each mission separately (lower panels).

Efficiency

As with performance times, we removed outliers in accordance with our pre-registration. This resulted in 14 zombie fighting times (two baseline no load, four baseline load, seven interactive no load, one interactive load) and 15 energy collection times (two baseline no load, two baseline load, three interactive no load, eight interactive load) being removed. In total, these criteria removed < 1% of the overall data.

Performance times are in minutes. The main effects of round, $F(1, 1050) = 675.71$, $p < 0.001$, and round squared, $F(1, 1050) = 296.27$, $p < 0.001$, indicate faster performance over rounds with a positive exponential curve. There was a main effect of mission, $F(1, 117) = 14.00$, $p < 0.001$, a mission X round interaction, $F(1, 1050) = 63.12$, $p < 0.001$, and a mission X round squared interaction, $F(1, 1050) = 39.53$, $p < 0.001$. These results were from steeper improvement with more abrupt plateaus for the energy collection mission. A main effect of cognitive load emerged, indicating slower

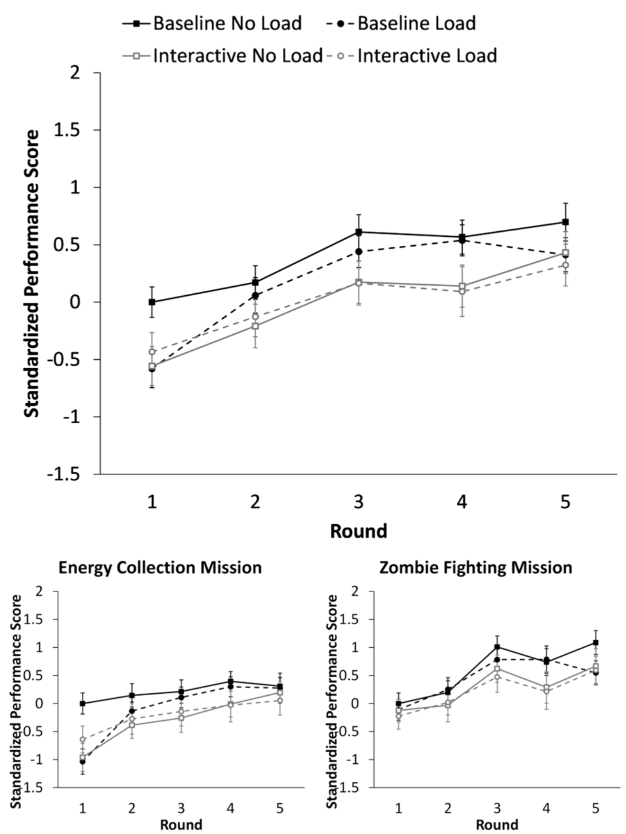

Fig. 6 Experiment 2 standardized mean performance on the plants vs. zombies task (upper panel) and by mission (lower panels). Performance data are standardized around round 1 performance in the baseline no load group, which has a score of 0. A score of 1/− 1 indicates performance one standard deviation above/below the round 1 baseline no load group mean. Error bars represent standard errors of the mean

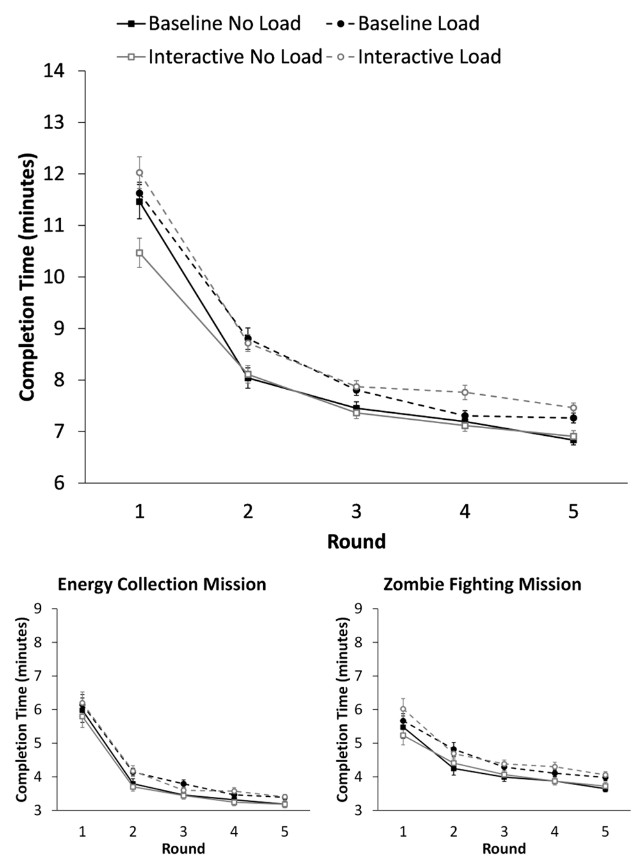

Fig. 7 Experiment 2 mean task completion times in minutes (upper panel) and for the energy collection missions and zombie-fighting missions separately (lower panels). Error bars represent standard errors of the mean

completion times for participants under cognitive load, $F(1, 117) = 9.11$, $p = 0.003$.

There was no effect of task version, $F(1, 117) = 0.14$, $p = 0.708$, nor were there any interactions with task version. Participants generally performed the interactive version as fast as the baseline version.

Figure 7 shows mean completion times over rounds (upper panel) and for each mission separately (lower panels) in minutes. These results suggest that interactive stimuli do not hinder efficiency.

Discussion

We found some evidence that performance scores were poorer in the interactive version, though this was limited to the energy collection missions and may be an anomaly. Performance was poorer and slower under cognitive load. While this was expected, we did not find evidence for cognitive load slowing completion times in the other experiments.

All task versions include a small degree of interactivity, in that participants must prioritize which suns/zombies are most important given the other stimuli (sunflowers/peashooters and suns/zombies) on the screen. Increasing the level of interactivity did not produce robust changes in learning or performance. Interactivity could have been manipulated in a number of ways. For example, the behavior of task elements (peashooters, sunflowers, suns, or zombies) could have differed depending on the presence or absence of the player's avatar in their row, or based on their position on the screen. Alternatively, performance on one phase of the task could have influenced the resource availability on the following phase of the task. However, any introduction of new task element behaviors or shifts in resources inherently alter the optimal strategy for the task. The optimal strategy for the interactive version would then be inherently more complex than for the baseline version. Our goal was to keep the optimal strategy as similar as possible between the task versions.

There are several possible reasons our manipulation of interactivity did not hinder learning. Any amount of interactivity may impair learning and performance whereas the relative amount of interactivity may be less important. Alternatively, our manipulation of interactivity may not have been robust: participants may not have attended to the effects of moving, thereby moving vertically and horizontally roughly equal amounts of time, washing out the impact on planting availability. Participants in the interactive version may also have improved their scores by strategically moving horizontally to get more plants to offset any poor decisions made due to increased cognitive difficulty. Finally, Hoffman et al.'s (2014) separable versus interactive dimension, by itself, may not be sufficient to increase task difficulty or hinder learning.

Experiment 3: discrete vs. continuous

People often categorize continuous variables. For example, people categorize wealth into lower, middle, and upper class and birth years into generations (e.g., baby boomers, generation X, millennials). The reason people categorize continuous stimuli is presumably that doing so eases the cognitive burden of nuanced decision making and eases interpretation. For example, psychologists frequently categorize continuous variables (e.g., personality: Webster & Kruglanski, 1994; cognitive functioning: Farias et al., 2008; expertise: Ericsson, Krampe, & Tesch-Römer, 1993; Güss, Edelstein, Babibanga, & Bartow, 2017; working memory capacity: Conway & Engle, 1994; Engle & Kane, 2004; and mental health and behavior risk factors: Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010). Although many tasks discussed previously involved continuous stimuli, and Hoffman et al. (2014) theorize that continua increase task difficulty, no experiment to our knowledge has specifically tested whether processing continuous stimuli affects learning, performance, or efficiency.

Methods

This study and all analyses, stopping rules, and exclusions criteria were pre-registered via the Open Science Framework (https://osf.io/u2gcn/register/5730e99a9ad5a102c5745a8a).

Participants

Our stopping rule was to collect 120 participants (roughly evenly divided among conditions). Data collection was terminated at the end of the week following the 120th participant.

A total of 138 participants—68 students enrolled in psychology classes at Case Western Reserve University and 70 students enrolled in psychology classes at Texas A&M University–Commerce—participated in exchange for partial course credit or extra credit. This study was approved by Institutional Review Boards at Case Western Reserve University and Texas A&M University–Commerce.

Table 5 Baseline vs. continuous task version characteristics

Baseline version characteristics	Continuous version characteristics
Consistent	Consistent
Static	Static
Sequential	Sequential
Separable	Separable
Discrete	**Continuous**
Linear	Linear
Single representation	Single representation
Mechanistic	Mechanistic
Homogenous	Homogenous

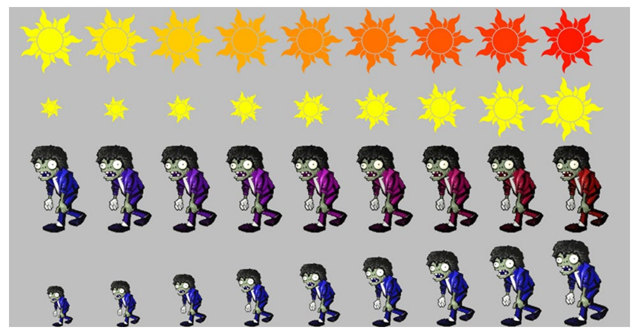

Fig. 8 Continuous version. The first and third rows represent the color continua, shown here in the large size. The second and fourth rows represent the size continua, shown here in yellow (suns) and blue (zombies), respectively. The baseline version uses the leftmost, center, and rightmost colors and sizes

Materials

Plants vs. zombies task: baseline version

The baseline version was identical to the baseline versions used in Experiments 1 and 2. As described previously, the baseline version of the task is sequential. Participants complete the energy collection mission followed by the zombie fighting mission in each round.

Plants vs. zombies task: continuous version

The continuous version is identical to the baseline version (see Table 5, non-bolded characteristics), except that it is continuous rather than discrete (see Table 5, bold characteristics). Instead of the three different sizes and three different colors used in the baseline version, the continuous version uses nine different sizes and nine different colors for a total of 81 possible size-color combinations, of which 27 appear in each round. The additional sizes and colors were linearly interpolated between the baseline versions' characteristics to create a continuum. See Fig. 8.

Cognitive load

The cognitive load manipulation was identical to that of Experiments 1 and 2.

Procedures

The procedure was identical to Experiments 1 and 2, except participants were assigned to either the baseline version or the continuous version.

Design and analyses

The design and analyses were identical to Experiments 1 and 2, except that the continuous version was compared to the baseline version. For regression weights and all fixed effect results, see Appendices 5 and 6.

Results

We predicted that the continuous version would require more nuanced decisions making and ultimately lower performance, hinder learning, and decrease efficiency either directly or by interacting with cognitive load.

Performance scores

We first removed outliers in accordance with our pre-registration. This resulted in eight zombie fighting scores (three baseline no load, one baseline load, four continuous no load) and eight energy collection scores (six baseline no load, one continuous no load, one continuous load) being removed. In total, these criteria removed approximately 1% of the overall data. We next standardized performance scores based on round 1 baseline no load group performance. Thus, each score is the number of standard deviations above or below the mean score for the first round for the theoretically easiest condition.

We observed main effects of both round, $F(1, 1207) = 45.46$, $p < 0.001$, and round squared, $F(1, 1207) = 9.25$, $p = 0.002$, indicating better performance over rounds with a negative exponential curve following typical learning curves (Fitts, 1964).

We did not observe a main effect of task version, $F(1, 133) = 0.45$, $p = 0.502$, or cognitive load, $F(1, 133) < 0.01$, $p = 0.965$. A marginally significant main effect of mission resulted from slightly higher scores on the zombie fighting missions, $F(1, 133) = 4.80$, $p = 0.030$.

Figure 9 shows standardized mean performance data over rounds (upper panel) and for each mission separately (lower panels). These results suggest that continuous stimuli do not hinder learning or performance.

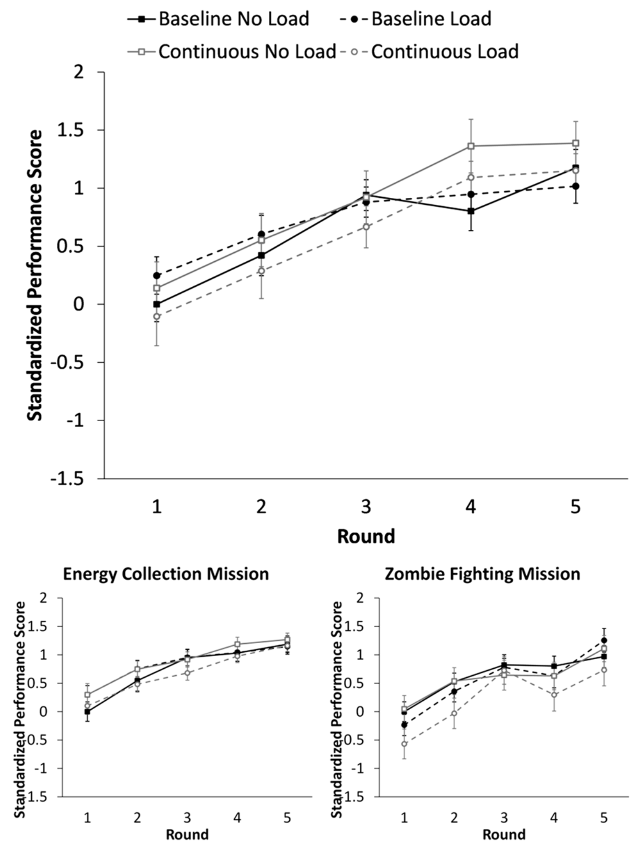

Fig. 9 Experiment 3 standardized mean performance on the plants vs. zombies task (upper panel) and separated by mission (lower panels). Performance data are standardized around round 1 performance in the baseline no load group, which has a score of 0. A score of 1/− 1 indicates performance one standard deviation above/below the round 1 baseline no load group mean. Error bars represent standard errors of the mean

Efficiency

As with performance scores, we removed outliers in accordance with our pre-registration. This resulted in seven zombie fighting times (one baseline no load, two baseline load, three continuous no load, one continuous load) and nine energy collection times (three baseline load, three continuous no load, three continuous load) being removed. In total, these criteria removed approximately 1% of the overall data.

Performance times are in minutes. The main effects of round, $F(1, 1186) = 714.56$, $p < 0.001$, and round squared, $F(1, 1186) = 300.28$, $p < 0.001$, indicate faster performance over rounds with a positive exponential curve. There was a main effect of mission, $F(1, 131) = 104.81$, $p < 0.001$, due to participants completing the energy collection missions faster than the zombie fighting missions. Mission also interacted with round, $F(1, 1186) = 154.88$, $p < 0.001$, and round squared, $F(1, 1186) = 96.38$, $p < 0.001$, due to steeper learning curves with more abrupt plateaus for the energy collection missions. Mission also interacted with cognitive load: mission X cognitive load, $F(1, 131) = 5.21$, $p = 0.024$.

A marginally significant main effect of task version, $F(1, 131) = 4.71$, $p = 0.032$, resulting from slower completion times for participants assigned to the continuous version compared to those assigned to the baseline version. We also observed significant three-way interactions: task version X cognitive load X round, $F(1, 1186) = 7.03$, $p = 0.008$, and task version X cognitive load X round squared, $F(1, 1186) = 5.96$, $p = 0.015$. These interactions resulted from participants in the continuous load not experiencing the same improvements in efficiency that their baseline counterparts experienced until later in the task.

The interactions between task version and cognitive load, along with the main effect of task version, are consistent with the hypothesis that making decisions based on continuous stimuli increases demands on cognitive resources. That is, tying up cognitive resources via a secondary task slowed performance on the continuous version, but not the baseline version. Figure 10 shows mean completion times over rounds (upper panel) and separated by mission (lower panels) in minutes.

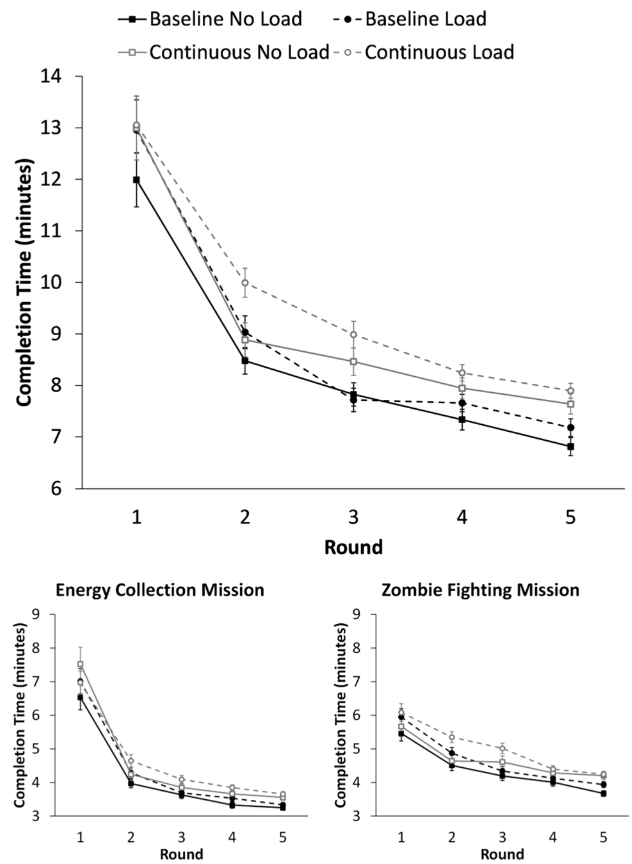

Fig. 10 Experiment 3 mean task completion times in minutes (upper panel) and for the energy collection missions and the zombie fighting missions separately (lower panels). Error bars represent standard errors of the mean

Discussion

Performance accuracy remained high and improved similarly regardless of whether stimuli were discrete or continuous. This is particularly surprising given that our manipulation was rather robust; we increased the number of possible stimulus categories by a factor of nine. It is possible that had we increased this number further such that the shifts in color and size were even more subtle that we would have found an effect. However, given the lack of an effect increasing by a factor of nine, this seems unlikely. It is also possible that once a task has any continuous features—suns and zombies move continuously across the screen in the plants vs. zombies task—increasing the number of features with continuous properties does not impact performance. Another possibility is that people are rather proficient at handling continua of color and size, but might struggle with other continua not tested here, such as time, or money. Finally, it is possible that people can handle two continua, but would struggle if combining the effects of three or more continuous features (e.g., air speed, altitude, and heading).

Despite equivalent scores across task versions, this came at a cost to efficiency for those in the continuous version. Although the effect of this slowing was only about half a minute per round, this amounts to just shy of half a standard deviation ($d = 0.47$) between task versions. Overall these results suggest that, despite it seeming intuitively important, Hoffman et al. (2014) discrete versus continuous dimension primarily affects performance speed, but not learning or performance accuracy.

General discussion

In two previous experiments (Macnamara & Frank, 2018), we investigated Ackerman's (1987) consistent vs. variably mapped dimension and Hoffman et al.'s (2014) static vs. dynamic dimension. We found that variable mapping decreased performance and also lowered efficiency. We also found that dynamic stimuli drastically decreased performance and hindered learning relative to static stimuli, particularly in the presence of cognitive load.

In the present set of studies, we conducted three experiments investigating Hoffman et al.'s (2014) sequential vs. simultaneous dimension, separable vs. interactive dimension, and discrete vs. continuous dimension. Importantly, we investigated these dimensions by isolating a single task characteristic and controlling for all other dimensions of difficulty.

Table 6 The role of examined task characteristics on learning, performance, and efficiency

Task characteristic	Learning and performance	Efficiency
Variably mapped	Negative impact	Negative impact
Dynamic	Large negative impact	–
Simultaneous	Large negative impact	Large negative impact
Interactive	No impact	No impact
Continuous	No impact	Small negative impact

Variably mapped and dynamic characteristics were examined in Macnamara and Frank (2018). Simultaneous, interactive, and continuous characteristics were examined in the present set of studies. The variably mapped characteristic was proposed by Ackerman (1987). All other task characteristics in this table were proposed by Hoffman et al. (2014)

– = efficiency could not be tested in the dynamic version of our paradigm independent of performance scores

In Experiment 1, we found strong evidence that simultaneity increases task difficulty and hinders learning and performance. Performance was both poorer and less efficient under simultaneous task conditions. Importantly, we established these effects controlling for the limitations frequently observed in task switching and multitasking paradigms—we eliminated confounds, perceptual bottlenecks, and mechanical bottlenecks, and were able to examine effects on learning.

Participants in the simultaneous version demonstrated a bias towards focusing resources on one mission (fighting zombies) over the other (energy collecting). Participants in the simultaneous version devoted slightly fewer resources to the energy collection missions than their baseline version counterparts, but performed *substantially* worse than those in the baseline version on these missions. Participants in the simultaneous version devoted substantially more resources to zombie fighting missions than their baseline version counterparts but did not outperform them.

Participants in the simultaneous version also demonstrated little to no improvement with practice and performed the task considerably slower than participants in the baseline version, an amount that cannot be explained by needing to press the toggle key. Thus, simultaneous task environments appear to severely impact learning, performance, and efficiency. See Table 6.

In Experiment 2, we found that people demonstrated a surprising ability to manage interactive task rules. We found only minimal evidence that interactive task features are sufficient to hinder learning and performance. These results suggest that further nuances and adjustments to Hoffman et al.'s (2014) theory may be warranted.

In Experiment 3, continuous stimuli had no effects on learning or performance accuracy. However, continuous stimuli slowed performance overall. These results suggest that we may underestimate people's ability to process continuous stimuli and make nuanced decisions, provided they have additional time to do so. Again, these results suggest that further nuances and adjustments to Hoffman et al.'s (2014) theory are warranted.

Taken together, these results begin to paint a picture of which characteristics are important to consider when examining learning and performance, and which do not appear to be important. To this end, knowing both which features produce significant learning, performance, and efficiency decrements, and which do not, is crucial for making potential adjustments to learning theories. See Table 6.

Future directions

Although task features influenced overall performance outcomes to varying degrees, it is also possible that they resulted in additional changes to performer behavior. That is, even if the final scores were similar, the strategies imposed to reach them may have differed across task versions. The current tasks did not record the full set of participant responses (e.g., timing of individual key presses) or screen capture for observation and analysis of their interactions with the task. Examining such qualitative changes in performance behavior may be an important avenue for future research.

Another important future direction is to examine the effects of more than one dimension of difficulty. Are tasks that are both simultaneous and dynamic twice as difficult as those that are only simultaneous or only dynamic? While managing stimuli that fall on a continuum does not hinder learning in isolation, does managing continuous stimuli become difficult in variably mapped task environments? Now that these task characteristics are better understood in their component parts, we can seek to understand their influence on behaviors when presented in combinations in real-world tasks.

Another important future direction is to test individual differences in cognitive abilities as predictors of performance under different task characteristics. Both Ackerman and colleagues (Ackerman & Cianciolo, 2002; Ackerman & Woltz, 1994) and Hoffman et al. (2014) suggested that certain task features increase task difficulty through a reliance on basic cognitive resources. An important avenue for future research is to test which cognitive abilities (e.g., working memory, updating, processing speed, reasoning) best predict performance on a task with a given set of features. This would provide insight into which cognitive processes are being recruited under different conditions. For example, working memory capacity might be a more important predictor of learning and performance on simultaneous tasks, where two goals must be maintained and updated throughout the task, than in sequential tasks where less information is processed at once.

Identifying task characteristics that moderate the predictive power of cognitive abilities and the predictive power of practice on learning and performance can provide insights into the cognitive processes underlying skill acquisition. It also has the potential to change how we design training paradigms, determine selection criteria, and classify jobs, depending on the task characteristics. Knowing which task characteristics recruit certain cognitive abilities, and which task characteristics impact the benefits of practice, would better enable employers to best select applicants and plan training protocols.

Conclusion

Task characteristics can play an important role in explaining variance in performance. Specifically, understanding which task characteristics impact learning, performance, and efficiency are crucial for building a more nuanced model of skill acquisition and expertise. By incorporating task characteristics into a theoretical framework, we can better understand variance in complex human performance.

Compliance with ethical standards

Conflict of interest No conflicts of interest are reported and all research was in compliance with the ethical standards of the American Psychological Association.

Open Practices statement Data for Experiments 1–3 are available on the Open Science Framework (https://osf.io/r4j48/, https://osf.io/29uyn/, and https://osf.io/ejqs3/, respectively). All were preregistered on the Open Science Framework (https://osf.io/a4dr3, https://osf.io/ygwq5, and https://osf.io/u2gcn respectively).

Appendix 1

Regression coefficients and fixed effects for Experiment 1 performance scores.

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Intercept	0.02	0.18	117	0.00	0.855	Mean score estimate for round 1 (coded as round 0) of energy mission in the baseline no load group. This score is fixed at zero due to standardization
Version	−0.85	0.26	117	7.83	0.006	Difference in round 1 score between the simultaneous no load and baseline no load groups (for energy mission)
Load	−0.02	0.26	117	1.28	0.260	Difference in round 1 score between the baseline load and baseline no load groups (for energy mission)
Mission	0.01	0.20	117	4.60	0.034	Difference in round 1 score between zombie and energy missions in the baseline no load group
Round	0.44	0.17	1055	35.55	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
$Round^2$	−0.07	0.04	1055	9.34	0.002	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	0.19	0.38	117	0.20	0.653	Difference in round 1 score between the simultaneous load and simultaneous no load groups (for energy mission)
Version X Mission	0.94	0.29	117	8.34	0.005	Difference in round 1 score between zombie and energy missions for the simultaneous no load group
Load X Mission	−0.18	0.28	117	5.85	0.017	Difference in round 1 score between zombie and energy missions for the baseline load group
Round X Version	−0.21	0.24	1055	0.87	0.351	Difference in the linear effect of round between the simultaneous no load and baseline no load groups (for energy missions)
Round X Load	−0.15	0.23	1055	0.01	0.942	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Mission	0.07	0.24	1055	3.24	0.072	Difference in the linear effect of round between zombie and energy missions for the baseline no load group
$Round^2$ X Version	0.03	0.06	1055	0.18	0.672	Difference in the exponential effect of round between the simultaneous no load and baseline no load groups (for energy missions)
$Round^2$ X Load	0.04	0.06	1055	0.01	0.931	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Round^2 X Mission	−0.01	0.06	1055	0.44	0.508	Difference in the exponential effect of round between zombie and energy missions for the baseline no load group
Version X Load X Mission	−0.66	0.42	117	2.48	0.118	Difference in round 1 score between zombie and energy missions for the simultaneous load group
Round X Version X Load	−0.01	0.35	1055	0.74	0.391	Difference in the linear effect of round between the simultaneous load and simultaneous no load groups (for energy missions)
Round X Version X Mission	−0.03	0.34	1055	0.65	0.420	Difference in the linear effect of round between zombie and energy missions in the simultaneous no load group
Round X Load X Mission	0.10	0.33	1055	1.75	0.186	Difference in the linear effect of round between zombie and energy missions in the baseline load group
Round^2 X Version X Load	−0.02	0.09	1055	1.33	0.249	Difference in the exponential effect of round between the simultaneous load and simultaneous no load groups (for energy missions)
Round^2 X Version X Mission	0.02	0.08	1055	0.23	0.631	Difference in the exponential effect of round between zombie and energy missions in the simultaneous no load group
Round^2 X Load X Mission	0.00	0.08	1055	0.61	0.434	Difference in the exponential effect of round between zombie and energy missions in the baseline load group
Round X Version X Load X Mission	0.46	0.50	1055	0.83	0.362	Difference in the linear effect of round between zombie and energy missions in the simultaneous load group

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Round^2 X Version X Load X Mission	− 0.10	0.12	1055	0.63	0.428	Difference in the exponential effect of round between zombie and energy missions in the simultaneous load group

Version task version (baseline, simultaneous), *Load* cognitive load (load, no load), *Round* linear improvement over rounds; *Round*^{2} nonlinear (quadratic) improvement over rounds, *Energy* energy collection, *Zombie* zombie fighting

Appendix 2

Regression coefficients and fixed effects for Experiment 1 completion times.

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Intercept	11.28	0.58	117	383.77	< 0.001	Mean completion time estimate for round 1 (coded as round 0) of energy mission in the baseline no load group
Version	6.86	0.83	117	97.23	< 0.001	Difference in round 1 completion time between the simultaneous no load and baseline no load groups (for energy mission)
Load	0.74	0.81	117	0.04	0.834	Difference in round 1 completion time between the baseline load and baseline no load groups (for energy mission)
Round	− 2.53	0.50	469	168.51	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4
Round^2	0.38	0.12	469	56.70	< 0.001	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	− 1.74	1.21	117	2.05	0.155	Difference in round 1 completion time between the simultaneous load and simultaneous no load groups (for energy mission)
Round X Version	− 1.72	0.72	469	2.47	0.116	Difference in the linear effect of round between the simultaneous no load and baseline no load groups (for energy missions)

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Round X Load	−0.93	0.70	469	0.00	0.955	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
$Round^2$ X Version	0.22	0.17	469	0.02	0.894	Difference in the exponential effect of round between the simultaneous no load and baseline no load groups (for energy missions)
$Round^2$ X Load	0.18	0.17	469	0.03	0.868	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Version X Load	1.79	1.05	469	2.92	0.088	Difference in the linear effect of round between the simultaneous load and simultaneous no load groups (for energy missions)
$Round^2$ X Version X Load	−0.41	0.25	469	2.59	0.108	Difference in the exponential effect of round between the simultaneous load and simultaneous no load groups (for energy missions)

Version task version (baseline, simultaneous), *Load* cognitive load (load, no load), *Round* linear improvement over rounds, $Round^2$ non-linear (quadratic) improvement over rounds

Appendix 3

Regression coefficients and fixed effects for Experiment 2 performance scores.

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
Intercept	0.00	0.18	118	0.00	0.986	Mean score estimate for round 1 (coded as round 0) of energy mission in the baseline no load group. This score is fixed at zero due to standardization
Version	−0.65	0.26	118	0.68	0.411	Difference in round 1 score between the interactive no load and baseline no load groups (for energy mission)

Effect	Estimate	SE	DF	F value	p value	Description of estimate
Load	−0.71	0.25	118	0.23	0.632	Difference in round 1 score between the baseline load and baseline no load groups (for energy mission)
Mission	−0.02	0.18	118	16.65	< 0.001	Difference in round 1 score between zombie and energy missions in the baseline no load group
Round	0.24	0.15	1078	55.51	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4
Round^2	−0.04	0.04	1078	18.49	< 0.001	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	0.95	0.36	118	3.71	0.056	Difference in round 1 score between the interactive load and interactive no load groups (for energy mission)
Version X Mission	0.44	0.25	118	0.28	0.600	Difference in round 1 score between zombie and energy missions for the interactive no load group
Load X Mission	0.67	0.25	118	3.25	0.074	Difference in round 1 score between zombie and energy missions for the baseline load group
Round X Version	0.13	0.21	1078	2.26	0.133	Difference in the linear effect of round between the interactive no load and baseline no load groups (for energy missions)
Round X Load	0.47	0.21	1078	0.31	0.576	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Mission	0.21	0.21	1078	0.02	0.877	Difference in the linear effect of round between zombie and energy missions for the baseline no load group
Round^2 X Version	0.00	0.05	1078	2.24	0.135	Difference in the exponential effect of round between the interactive no load and baseline no load groups (for energy missions)

Effect	Estimate	SE	DF	F value	p value	Description of estimate
$Round^2$ X Load	−0.08	0.05	1078	0.56	0.454	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)
$Round^2$ X Mission	−0.02	0.05	1078	0.01	0.918	Difference in the exponential effect of round between zombie and energy missions for the baseline no load group
Version X Load X Mission	−0.70	0.35	118	4.00	0.048	Difference in round 1 score between zombie and energy missions for the interactive load group
Round X Version X Load	−0.56	0.30	1078	3.17	0.075	Difference in the linear effect of round between the interactive load and interactive no load groups (for energy missions)
Round X Version X Mission	−0.21	0.30	1078	0.01	0.936	Difference in the linear effect of round between zombie and energy missions in the interactive no load group
Round X Load X Mission	−0.44	0.30	1078	1.49	0.222	Difference in the linear effect of round between zombie and energy missions in the baseline load group
$Round^2$ X Version X Load	0.08	0.07	1078	1.61	0.205	Difference in the exponential effect of round between the interactive load and interactive no load groups (for energy missions)
$Round^2$ X Version X Mission	0.01	0.07	1078	0.00	0.963	Difference in the exponential effect of round between zombie and energy missions in the interactive no load group
$Round^2$ X Load X Mission	0.05	0.07	1078	0.56	0.455	Difference in the exponential effect of round between zombie and energy missions in the baseline load group
Round X Version X Load X Mission	0.38	0.42	1078	0.83	0.363	Difference in the linear effect of round between zombie and energy missions in the interactive load group

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
$Round^2$ X Version X Load X Mission	− 0.03	0.10	1078	0.09	0.762	Difference in the exponential effect of round between zombie and energy missions in the interactive load group

Version task version (baseline, interactive), *Load* cognitive load (load, no load), *Round* linear improvement over rounds, *Round*^{2} non-linear (quadratic) improvement over rounds,*Energy* energy collection, *Zombie* zombie fighting

Appendix 4

Regression coefficients and fixed effects for Experiment 2 completion times.

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Intercept	5.81	0.17	117	1153.28	< 0.001	Mean completion time for round 1 (coded as round 0) of energy mission in the baseline no load group
Version	− 0.23	0.24	117	0.14	0.708	Difference in round 1 completion time between the interactive no load and baseline no load groups (for energy mission)
Load	0.17	0.24	117	9.11	0.003	Difference in round 1 completion time between the baseline load and baseline no load groups (for energy mission)
Mission	− 0.44	0.18	117	14.00	< 0.001	Difference in round 1 completion time between zombie and energy missions in the baseline no load group
Round	− 1.86	0.15	1050	675.71	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4
$Round^2$	0.31	0.04	1050	296.27	< 0.001	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	0.28	0.34	117	1.69	0.197	Difference in round 1 completion time between the interactive load and interactive no load groups (for energy mission)

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Version X Mission	−0.03	0.26	117	0.12	0.733	Difference in round 1 completion time between zombie and energy missions for the interactive no load group
Load X Mission	0.14	0.25	117	1.70	0.195	Difference in round 1 completion time between zombie and energy missions for the baseline load group
Round X Version	0.14	0.22	1050	0.21	0.646	Difference in the linear effect of round between the interactive no load and baseline no load groups (for energy missions)
Round X Load	0.11	0.21	1050	0.47	0.495	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Mission	0.84	0.21	1050	63.12	< 0.001	Difference in the linear effect of round between zombie and energy missions for the baseline no load group
Round^2 X Version	−0.02	0.05	1050	0.06	0.802	Difference in the exponential effect of round between the interactive no load and baseline no load groups (for energy missions)
Round^2 X Load	−0.03	0.05	1050	0.10	0.747	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)
Round^2 X Mission	−0.16	0.05	1050	39.53	< 0.001	Difference in the exponential effect of round between zombie and energy missions for the baseline no load group
Version X Load X Mission	0.18	0.36	117	0.24	0.623	Difference in round 1 completion time between zombie and energy missions for the interactive load group

Effect	Estimate	SE	*DF*	*F* value	*p* value	Description of estimate
Round X Version X Load	−0.21	0.30	1050	2.57	0.109	Difference in the linear effect of round between the interactive load and interactive no load groups (for energy missions)
Round X Version X Mission	0.16	0.30	1050	0.03	0.871	Difference in the linear effect of round between zombie and energy missions in the interactive no load group
Round X Load X Mission	−0.04	0.30	1050	0.60	0.440	Difference in the linear effect of round between zombie and energy missions in the baseline load group
$Round^2$ X Version X Load	0.04	0.07	1050	1.75	0.187	Difference in the exponential effect of round between the interactive load and interactive no load groups (for energy missions)
$Round^2$ X Version X Mission	−0.04	0.07	1050	0.02	0.893	Difference in the exponential effect of round between zombie and energy missions in the interactive no load group
$Round^2$ X Load X Mission	0.00	0.07	1050	0.45	0.503	Difference in the exponential effect of round between zombie and energy missions in the baseline load group
Round X Version X Load X Mission	−0.25	0.42	1050	0.37	0.545	Difference in the linear effect of round between zombie and energy missions in the interactive load group
$Round^2$ X Version X Load X Mission	0.06	0.10	1050	0.34	0.559	Difference in the exponential effect of round between zombie and energy missions in the interactive load group

Version task version (baseline, interactive), *Load* cognitive load (load, no load), *Round* linear improvement over rounds, $Round^2$ non-linear (quadratic) improvement over rounds, *Energy* energy collection, *Zombie* zombie fighting

Appendix 5

Regression coefficients and fixed effects for Experiment 3 performance scores.

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
Intercept	−0.02	0.22	133	0.00	0.932	Mean score estimate for round 1 (coded as round 0) of energy mission in the baseline no load group. This score is fixed at zero due to standardization
Version	0.34	0.32	133	0.45	0.502	Difference in round 1 score between the continuous no load and baseline no load groups (for energy mission)
Load	0.34	0.30	133	0.00	0.965	Difference in round 1 score between the baseline load and baseline no load groups (for energy mission)
Mission	0.00	0.25	133	4.80	0.030	Difference in round 1 score between zombie and energy missions in the baseline no load group
Round	0.62	0.21	1207	45.46	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4
Round^2	−0.08	0.05	1207	9.25	0.002	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	−0.54	0.44	133	1.97	0.163	Difference in round 1 score between the continuous load and continuous no load groups (for energy mission)
Version X Mission	−0.41	0.36	133	2.24	0.137	Difference in round 1 score between zombie and energy missions for the continuous no load group
Load X Mission	−0.17	0.34	133	0.27	0.601	Difference in round 1 score between zombie and energy missions for the baseline load group
Round X Version	−0.20	0.30	1207	0.11	0.746	Difference in the linear effect of round between the continuous no load and baseline no load groups (for energy missions)

Effect	Estimate	SE	DF	F value	p value	Description of estimate
Round X Load	−0.19	0.29	1207	0.18	0.669	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Mission	−0.20	0.29	1207	0.35	0.556	Difference in the linear effect of round between zombie and energy missions for the baseline no load group
$Round^2$ X Version	0.04	0.07	1207	0.13	0.718	Difference in the exponential effect of round between the continuous no load and baseline no load groups (for energy missions)
$Round^2$ X Load	0.03	0.07	1207	0.02	0.877	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)
$Round^2$ X Mission	0.04	0.07	1207	0.01	0.912	Difference in the exponential effect of round between zombie and energy missions for the baseline no load group
Version X Load X Mission	0.07	0.50	133	0.02	0.887	Difference in round 1 score between zombie and energy missions for the continuous load group
Round X Version X Load	0.11	0.42	1207	0.07	0.787	Difference in the linear effect of round between the continuous load and continuous no load groups (for energy missions)
Round X Version X Mission	0.42	0.43	1207	1.79	0.181	Difference in the linear effect of round between zombie and energy missions in the continuous no load group
Round X Load X Mission	0.18	0.41	1207	0.27	0.607	Difference in the linear effect of round between zombie and energy missions in the baseline load group

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
Round^2 X Version X Load	0.00	0.10	1207	0.00	0.985	Difference in the exponential effect of round between the continuous load and continuous no load groups (for energy missions)
Round^2 X Version X Mission	−0.05	0.10	1207	0.54	0.465	Difference in the exponential effect of round between zombie and energy missions in the continuous no load group
Round^2 X Load X Mission	−0.04	0.10	1207	0.34	0.559	Difference in the exponential effect of round between zombie and energy missions in the baseline load group
Round X Version X Load X Mission	−0.06	0.59	1207	0.01	0.918	Difference in the linear effect of round between zombie and energy missions in the continuous load group
Round^2 X Version X Load X Mission	0.00	0.14	1207	0.00	0.989	Difference in the exponential effect of round between zombie and energy missions in the continuous load group

Version task version (baseline, continuous), *Load* cognitive load (load, no load), *Round* linear improvement over rounds, *Round*^{2} non-linear (quadratic) improvement over rounds, *Energy* energy collection, *Zombie* zombie fighting

Appendix 6

Regression coefficients and fixed effects for Experiment 3 completion times.

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
Intercept	6.30	0.18	131	1253.86	< 0.001	Mean completion time for round 1 (coded as round 0) of energy mission in the baseline no load group
Version	0.91	0.26	131	4.71	0.032	Difference in round 1 completion time between the continuous no load and baseline no load groups (for energy mission)

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
Load	0.48	0.25	131	2.88	0.092	Difference in round 1 completion time between the baseline load and baseline no load groups (for energy mission)
Mission	−0.92	0.20	131	104.81	< 0.001	Difference in round 1 completion time between zombie and energy missions in the baseline no load group
Round	−2.15	0.16	1186	714.56	< 0.001	Linear effect of round in the baseline no load group (for energy missions). Rounds coded 0–4
Round^2	0.36	0.04	1186	300.28	< 0.001	The exponential effect of round in the baseline no load group (for energy missions)
Version X Load	−0.91	0.36	131	2.57	0.111	Difference in round 1 completion time between the continuous load and continuous no load groups (for energy mission)
Version X Mission	−0.68	0.29	131	1.63	0.204	Difference in round 1 completion time between zombie and energy missions for the continuous no load group
Load X Mission	0.03	0.27	131	5.21	0.024	Difference in round 1 completion time between zombie and energy missions for the baseline load group
Round X Version	−0.58	0.24	1186	0.01	0.903	Difference in the linear effect of round between the continuous no load and baseline no load groups (for energy missions)
Round X Load	−0.23	0.23	1186	0.43	0.511	Difference in the linear effect of round between the baseline load and baseline no load groups (for energy missions)
Round X Mission	1.34	0.23	1186	154.88	< 0.001	Difference in the linear effect of round between zombie and energy missions for the baseline no load group

Effect	Estimate	SE	DF	*F* value	*p* value	Description of estimate
$Round^2$ X Version	0.11	0.06	1186	0.01	0.923	Difference in the exponential effect of round between the continuous no load and baseline no load groups (for energy missions)
$Round^2$ X Load	0.04	0.06	1186	1.18	0.278	Difference in the exponential effect of round between the baseline load and baseline no load groups (for energy missions)
$Round^2$ X Mission	−0.26	0.06	1186	96.38	< 0.001	Difference in the exponential effect of round between zombie and energy missions for the baseline no load group
Version X Load X Mission	0.85	0.40	131	4.47	0.036	Difference in round 1 completion time between zombie and energy missions for the continuous load group
Round X Version X Load	0.91	0.34	1186	7.03	0.008	Difference in the linear effect of round between the continuous load and continuous no load groups (for energy missions)
Round X Version X Mission	0.56	0.35	1186	1.45	0.229	Difference in the linear effect of round between zombie and energy missions in the continuous no load group
Round X Load X Mission	−0.01	0.32	1186	1.45	0.229	Difference in the linear effect of round between zombie and energy missions in the baseline load group
$Round^2$ X Version X Load	−0.18	0.08	1186	5.96	0.015	Difference in the exponential effect of round between the continuous load and continuous no load groups (for energy missions)

Effect	Estimate	SE	DF	F value	p value	Description of estimate
$Round^2$ X Version X Mission	−0.09	0.08	1186	0.89	0.346	Difference in the exponential effect of round between zombie and energy missions in the continuous no load group
$Round^2$ X Load X Mission	0.00	0.08	1186	0.60	0.439	Difference in the exponential effect of round between zombie and energy missions in the baseline load group
Round X Version X Load X Mission	−0.55	0.48	1186	1.36	0.244	Difference in the linear effect of round between zombie and energy missions in the continuous load group
$Round^2$ X Version X Load X Mission	0.08	0.11	1186	0.48	0.490	Difference in the exponential effect of round between zombie and energy missions in the continuous load group

Version task version (baseline, continuous), *Load* cognitive load (load, no load), *Round* linear improvement over rounds, $Round^2$ nonlinear (quadratic) improvement over rounds, *Energy* energy collection, *Zombie* zombie fighting

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*(1), 3–27. https://doi.org/10.1037/0033-2909.102.1.3.

Ackerman, P. L., & Cianciolo, A. T. (2002). Ability and task constraint determinants of complex task performance. *Journal of Experimental Psychology: Applied, 8*(3), 194–208. https://doi.org/10.1037/1076-898X.8.3.194.

Ackerman, P. L., & Woltz, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task constraints, individual differences, volition, and motivation. *Journal of Educational Psychology, 86*(4), 487–515. https://doi.org/10.1037//0022-0663.86.4.487.

Borger, R. (1963). The refractory period and serial choice-reactions. *Quarterly Journal of Experimental Psychology, 15*(1), 1–12. https://doi.org/10.1080/17470216308416546.

Bratfisch, O., & Hagman, E. (2003). Simultankapazität/Multi-Tasking (SIMKAP) Version 24.00: Handanweisung [Simultaneous capacity/Multi-tasking (SIMKAP) release 24.00: Manual]. Mödling, Austria: Schuhfried.

Bühner, M., König, C. J., Pick, M., & Krumm, S. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human Performance, 19*(3), 253–275. https://doi.org/10.1207/s15327043hup1903_4.

Caird, J. K., Simmons, S. M., Wiley, K., Johnston, K. A., & Horrey, W. J. (2018). Does talking on a cell phone, with a passenger, or dialing affect driving performance? An updated systematic review and meta-analysis of experimental studies. *Human Factors, 60*(1), 101–133. https://doi.org/10.1177/0018720817748145.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*(2), 331–335. https://doi.org/10.3758/BF03196169.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*(4), 354–373. https://doi.org/10.1037/0096-3445.123.4.354.

Creamer, L. R. (1963). Event uncertainty, psychological refractory period, and human data processing. *Journal of Experimental Psychology, 66*(2), 187–194. https://doi.org/10.1037/h0041443.

Dawson-Saunders, B., Feltovich, P. J., Coulson, R. L., & Steward, D. E. (1990). A survey of medical school teachers to identify basic biomedical concepts medical students should understand. *Academic Medicine, 65*, 448–454. https://doi.org/10.1097/00001888-199007000-00008.

Drews, F. A., Pasupathi, M., & Strayer, D. L. (2008). Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied, 14*(4), 392–400. https://doi.org/10.1037/a0013119.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *The Psychology of Learning and Motivation, 44*, 145–199. https://doi.org/10.1016/S0079-7421(03)44005-X.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363.

Farias, S. T., Mungas, D., Reed, B. R., Cahn-Weiner, D., Jagust, W., Baynes, K., et al. (2008). The measurement of everyday cognition (ECog): Scale development and psychometric properties. *Neuropsychology, 22*(4), 531–544. https://doi.org/10.1037/0894-4105.22.4.531.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1989). The nature of conceptual understanding in biomedicine: The deep structure of complex ideas and the development of misconceptions. In D. A. Evans & D. A. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling* (pp. 113–172). Cambridge, MA: The MIT Press.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1993). Learning, teaching, and testing for complex conceptual understanding. In

N. Frederiksen, R. Mislevy, & I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 181–217). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243–285). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-1-4832-3145-7.50016-9.

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Güss, C. D., Edelstein, H. D., Babibanga, A., & Bartow, S. (2017). Comparing business experts and novices in complex problem solving. *Journal of Intelligence, 5*, 1–18.

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology, 24*(8), 1149–1167. https://doi.org/10.1002/acp.1624.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education, 15*(1), 46–64. https://doi.org/10.1007/BF02940852.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 417–429. https://doi.org/10.1037/0278-7393.16.3.417.

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., & Andrews, D. H. (2014). *Accelerated expertise: Training for high proficiency in a complex world*. New York: Psychology Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292. https://doi.org/10.1142/9789814417358_0006.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (2000). *SAS system for mixed models* (4th ed.). Cary, NC: SAS Institute.

Macnamara, B. N., & Frank, D. J. (2018). How do task characteristics affect learning and performance? The roles of variably mapped and dynamic tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(5), 764–778.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7.

Namco. (1980). *Pacman*. Tokyo: Author.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 763–797. https://doi.org/10.1037/0096-1523.27.4.763.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190. https://doi.org/10.1037/0033-295X.84.2.127.

Strayer, D. L., & Drews, F. A. (2007). Cell-phone–induced driver distraction. *Current Directions in Psychological Science, 16*(3), 128–131. https://doi.org/10.1177/154193120404801902.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development, 81*(3), 737–756. https://doi.org/10.1111/j.1467-8624.2010.01431.x.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136*(4), 601–626. https://doi.org/10.1037/a0019791.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049–1062. https://doi.org/10.1037/0022-3514.67.6.1049.

Wood, N. L., & Cowan, N. (1995). The cocktail party phenomenon re-visited: Attention and memory in the classic selective listening procedure of Cherry (1953). *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 255–260. https://doi.org/10.1037/0096-3445.124.3.243.