Reference: Ackerman, R. (in press). It's Time to Opt Out: Metacognitive Analysis of Time Regulation Under Uncertainty. Journal of Experimental Psychology: Learning, Memory, and Cognition.

It's Time to Opt Out:
Metacognitive Analysis of Time Regulation Under Uncertainty

Rakefet Ackerman

Technion—Israel Institute of Technology

Corresponding author:

Rakefet Ackerman

Faculty of Data and Decision Sciences

Technion—Israel Institute of Technology

Technion City, Haifa, Israel, 3200003

Email: ackerman@technion.ac.il

Author Note

The complete material sets, datasets, and SPSS code examples are publicly available on the Open Science Framework: [https://osf.io/khmsj/?view_only=de9ca5b10174446e80b8825545a7e240].

Keywords: Metacognition; Confidence; Effort Regulation; Thinking Efficiency; Opting Out

Acknowledgments: I thank Simon Jackson for early discussions and Meira Ben-Gad for editorial assistance.

Abstract

When performing cognitive tasks like solving problems in an exam or tackling challenging thinking tasks at work, knowing when to stop struggling with a difficult problem is often crucial for effective global performance. That is, opting out quickly when success is unlikely allows conserving time for other tasks with a higher chance of success. While research has examined how we find answers, giving-up efficiency has been largely ignored. This study delineates the metacognitive process by which individuals manage their mental effort when given a legitimate choice to quit. The present research extends prior metacognitive models, with one and two stopping rules, that did not take the temporal dynamics of opting out into account. The 3-Stopping Rule Model (3SRM) addresses this gap in understanding time regulation under uncertainty by adding an opting-out confidence criterion to the known confidence criterion and a time limit for providing answers. Three experiments ($N = 596$) used problem-solving and general knowledge tasks that differ in opt-out rates and patterns of the confidence stopping criterion for submitting answers. The set of opt-out measures revealed that while opting out frequency was sensitive to manipulations, tasks, and individual characteristics, opt-out confidence remained stable and independent of time across incentives, tasks, opt-out wording (don't know, skip, help), and individual-level characteristics. The 3SRM advances prior metacognitive frameworks by revealing that opting out is not merely a by-product of failed problem-solving, but a systematically governed control decision with its own stable confidence threshold by which people counteract the speed–accuracy tradeoff.

1. Introduction

Allocating thinking time is a core regulatory process when engaging in sequences of thinking tasks in contexts that require efficiency, including many forms of work and learning. For example, consider a physician seeing consecutive patients in the clinic, a garage mechanic who accepts clients by order of arrival, or a student facing an online test in an environment that does not support going back to questions previously seen. All must handle task items that vary in the challenge they present while aiming to work efficiently, without knowing the difficulty of items not yet encountered.

The metacognitive approach provides concepts and methods for probing these regulatory processes. Indeed, a vast body of metacognitive research has done much to uncover the regulatory decisions that underlie activities like learning, answering challenging questions, or solving problems (Bjork et al., 2013; Fiedler et al., 2019). Yet almost all of this research has used procedures in which participants are expected to provide substantive responses to all task items. In these designs, choosing not to answer (e.g., typing "no idea" as the answer) is considered as violating expectations that one will respond in good faith. However, enforcing substantive responses stands at odds with many common real-life scenarios, in which people are free to choose whether to provide an answer or opt out (Scoboria & Fisico, 2013). The present study deals with the temporal dynamics of opting out—deciding when to stop investing effort while acknowledging failure to find an appropriate answer.

1.1. Effective Opting Out

In both learning and problem-solving contexts using multi-item tasks, a robust finding is that the largest portion of time is devoted to the most challenging items (e.g., Ackerman, Yom-Tov, et al., 2020; Koriat et al., 2006; Metcalfe & Kornell, 2005; Son & Sethi, 2010). That is, the items with the least chance of success are most responsible for inefficiencies in time investment. Facing the prospect of accumulating wasted time, people may apply a partial remedy by choosing to quickly quit very challenging items in order to invest their time more effectively in easier items in the stream (Glucksberg & McCloskey, 1981; Metcalfe & Kornell, 2005; Payne & Duggan, 2011). For example, while lingering over a difficult question may be helpful during skill acquisition (Cohen et al., 2024; Masís et al., 2023), students facing a time-restricted test would do better to invest more time in questions they can expect to answer correctly (Bae et al., 2021). Quick opting out could thus be an efficiency improvement strategy readily available in people's toolboxes (see Sidi & Ackerman, 2024, for a review). The core idea is that when performing challenging thinking tasks, people can counteract the speed–accuracy tradeoff—i.e., the fact that responding faster risks increasing the error rate—by treating time as a resource to be managed (Donkin et al., 2014; Reed, 1973).

Efficient opting out depends on both the *rate* at which one opts out and the *time* invested before opting out. Regarding the opt-out rate, one might expect that when people can legitimately save time by opting out, they will do so, rather than persist with task items they are struggling to solve, in line with the *cognitive miserliness* principle (Stanovich, 2018). Law et al. (2022) demonstrated across several visual processing tasks with two substantive answer options that some people tend to opt out (by clicking "Uncertain") more than others (see also Reynolds et al., 2021). Law et al. (2025) found that some people use opting out more effectively than others as a personality trait. In three tasks used by Law et al. (2022), mean success rates were 71%–87% and opt-out rates ranged between only 4% and 14%. In Law et al. (2025), mean success rates across three tasks were 43%–63% and opt-out rates 14%–18%. Notably, there was nevertheless enough variability between people to support the individual differences analysis. In a study in which participants could submit their solution or opt out by clicking "I don't know" (DK), Ackerman (2014) documented both higher and lower opt-out rates. She found that with a non-misleading open-ended verbal task, the Compound Remote Associate task (CRA; word triplets where a common word generates a phrase with each of the three words), participants opted out in about 25% of responses despite success rates in submitted solutions being only 55%. With a misleading math and logic task (the Cognitive Reflection Test), in which confidence typically remains high even for wrong responses, participants opted out in only 7% of their responses, and their success rate was only about 35%. These examples demonstrate that people do not overuse opting out, but, if anything, underuse it.

The second challenge for improving efficiency when people are offered opt-out response options is the need to implement opt-out decisions quickly, rather than waste thinking time and end up with no answer that may be correct. In one of the earliest studies into this topic, Glucksberg and McCloskey (1981) showed that people sometimes do opt out quickly when answering knowledge questions. They concluded that "a person who is asked a question first conducts a preliminary search of memory to locate any stored information that may be relevant. If nothing relevant is found, a rapid DK response is made. If, however, potentially relevant facts are retrieved, they are examined in detail to determine whether they specify an answer to the question. If the retrieved information fails to provide an answer, a slow DK response is made" (Abstract). In a metacognitive study, Singer and Tiede (2008) used the quick Feeling of Knowing judgment (phrased as "once-knew-it") to capture whether the participant expected to find the correct answer during a subsequent, separate answering phase (yes/no). Notably, time to "yes" and "no" responses was highly similar. Moreover, this Feeling of Knowing judgment positively predicted how long participants took to opt out. This finding implies that there was substantial variability in the time people invested before opting out, with quick and slow opt-out responses, in parallel to the time range seen when participants chose to submit their answers. In Law et al. (2025), who used three different problem-solving tasks, response times (RTs) for opting out were mostly parallel to those for submitted responses. For example, means of per-participant RTs for Raven matrices were 23.7s (SD = 19.3) for opt-out responses and 25.0s (SD = 13.8) for submitted responses.

Yet another piece of evidence regarding the time invested before opting out comes from Ackerman (2014). In that paper, RTs were reported only for submitted responses. For the present study, I analyzed the RTs for opt-out responses from Ackerman's (2014) data set alongside the same participants' submitted responses (N = 44 who used both). While mean RTs per participant were longer for opted-out (44.2s) than submitted responses (25.6s), the lengthiest RTs per participant were shorter for the former than the latter (69.6s vs. 74.9s). Together, these studies indicate that when an explicit opt-out option is available, people do opt out relatively early, cutting some of the thinking time that would be wasted if they were to opt out only later (i.e., following similar time investment as responses submitted after deliberation). However, it is not yet clear just how effectively opting out is used.

In the present study, I use the metacognitive approach to examine the association between confidence and the time people invest before submitting or opting out. For this purpose, I extend existing stopping rule models by introducing an opt-out stopping rule.

1.2. Metacognitive Stopping Rules

According to the classic meta-memory framework (Nelson & Narens, 1990), people monitor their chance of success before, during, and after memorizing using Ease of Learning or ongoing and final Judgments of Learning. While attempting knowledge tasks, people use the Feeling of Knowing judgment before retrieval and the Confidence judgment after it. All these metacognitive judgments guide metacognitive control decisions regarding allocation of study/retrieval time and choice of relevant strategies.

Later theoretical advancements have brought up opting out as a strategic metacognitive control tool, allowing people to improve the proportion of correct answers among those they choose to provide (see Guzel & Higham, 2013, for a review). Koriat and Goldsmith (1996) used the term *output-bound accuracy* for the percentage of correct responses out of total number of answers participants chose to provide, and *input-bound accuracy* for the percentage of correct responses out of out of total number of questions asked (these are also known as free-report quantity and forced-report quantity, respectively; Guzel & Higham, 2013). Participants in that study achieved higher output-bound than input-bound accuracy by crossing out the answers they were less confident in. The actual application of confidence for opting out is termed *control sensitivity* (Koriat & Goldsmith, 1996), and is measured by within-participant correlation between confidence and the decision whether to submit or opt out. High control sensitivity is found when confidence in submitted responses is almost always higher than confidence in responses where one chooses to opt out (see Goldsmith, 2016, for a review; e.g., Hanczakowski et al., 2013; Undorf et al., 2021). These measures are used in the present study. Notably, none of these studies considered temporal dynamics or efficiency.

According to the more recent meta-reasoning framework (Ackerman & Thompson, 2017), when solving problems (or other thinking challenges, like syllogistic reasoning or creative tasks), people initially judge whether the problem is solvable (in terms of either objective solvability or whether they are personally equipped to solve the problem; Ackerman & Beller, 2017). If the answer is "no," they opt out. If the answer is "yes," they attempt to produce a first solution, to which a Feeling of Rightness or initial confidence judgment is attached. That Feeling of Rightness (or initial confidence) guides the decision to provide the initial solution as final, or deliberate more (Thompson et al., 2011). If one engages in deliberation, intermediate confidence judgments accompany intermediate solution candidates as they fall under consideration. The ultimate chosen response may be to provide a substantive solution or opt out (Lauterman & Ackerman, 2019, 2024). Thus, multiple points in the process involve control decisions whether to submit, continue thinking, or opt out.

Several metacognitive stopping rule models describe effort allocation. The classic framework, the *Discrepancy Reduction Model* (Nelson & Narens, 1990), is based on a *single pre-defined stopping rule*, representing the learner's or thinker's self-determined target level (or goal) for judgments of learning or confidence (e.g., Ackerman & Goldsmith, 2011; Dunlosky & Rawson, 2012). People invest learning or thinking time in order to improve their perceived chance of success, reflected by a cumulatively increasing metacognitive judgment until they reach this pre-defined target. The target is set by integrating task conditions (e.g., time pressure: I cannot learn everything in the given time) and motivation to succeed (e.g., a test score of 80% is good enough for my current purpose). See Fiedler et al. (2019) for a review.

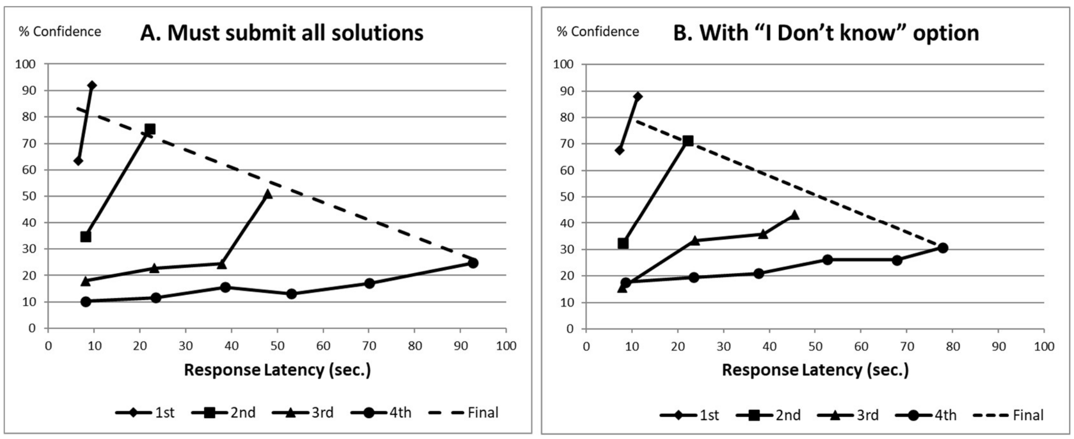
It is widely agreed in the metacognitive literature that people do indeed accumulate confidence as they invest more time in learning, thinking, or trying to retrieve information (Ackerman, 2014; Koriat et al., 2002; Metcalfe & Wiebe, 1987; Thompson et al., 2013). Similarly, in decision-making the *Drift-Diffusion Model* (see Ratcliff et al., 2016, for a review) posits that the decision to stop deliberating between the presented options depends on the level of accumulated evidence matching the target strength. Bridging between decision-making research and metacognitive research, recent investigations see confidence as reflecting the subjective quality of the accumulated evidence (Calder-Travis et al., 2023; Lee et al., 2023). Notably, diffusion models in various variations have been often considered using two-alternative forced choice tasks, many of them perceptual (e.g., identifying the direction of moving dot displays, Bottemanne & Dreher, 2019), and their briefer go/no-go variants (Gomez et al., 2007), involving little knowledge and inference beyond the given details, which differ from the focus of reasoning and meta-reasoning research on more complex tasks (Funke, 2010).

By all these models, across domains, adherence to a single stopping rule should produce a zero time–confidence correlation, because regardless of the time invested, all final confidence should be around the participant's pre-set target level (e.g., a pre-set target of 80% as above would yield mostly confidence ratings around 80%, regardless of the time invested before reaching there). In fact, vast empirical evidence does not support this prediction, as negative time–judgment correlations—high judgments for quick responses and lower for slower ones—are prevalent across tasks and contexts (Bago & De Neys, 2017; Mazor et al., 2023; Undorf & Erdfelder, 2011; Walker et al., 2019). These findings underly the large body of literature on collapsing boundaries and urgency signals in simple two-alternative forced choice perceptual tasks (e.g., brightness discrimination. See Hawkins et al., 2015, for a review). The common metacognitive explanation for these negative correlations is based on fluency, or ease of processing (Koriat et al., 2006; Oppenheimer, 2008)—inferring that the task was easy if the response came to mind quickly, and that the task was challenging, with a low chance of success, if it took a long time to come up with a response. Notably, though, this fluency-based inference process should produce very strong associations between time and metacognitive judgments, as people read just one parameter, their effort, and infer their chance to succeed based on it. However, these negative correlations, while significant, tend to be moderate rather than strong (e.g., -.42, Koriat et al., 2006), hinting that more factors affect the time–confidence association.

Furthermore, when the task is very challenging, to the extent that one is unable to progress as time goes by, both the Discrepancy Reduction Model and classic diffusion models predict thinking forever when one's pre-set confidence target—the sole condition for stopping—cannot be met (Ackerman & Morsanyi, 2023). Of course, people do not actually think forever, but eventually give up—either by opting out, or by providing a solution which they acknowledge has a low chance of being correct (Ackerman, 2014; Bae et al., 2021; Payne & Duggan, 2011).

In attempt to reconcile the contradiction between the target-directed effort regulation implied by the classic models, which predicts zero time–judgment correlations, and the observed persistent negative time–confidence associations, Ackerman (2014) suggested a model with two stopping rules, the *Diminishing Criterion Model* (DCM; Ackerman, 2014). By the DCM, the first stopping rule is a diminishing confidence criterion: as people think longer, they become willing to provide solutions despite their confidence being lower than what they would have required for solutions produced quickly. The second stopping rule is a *time limit* beyond which people are not willing to continue thinking about an item. Notably, the DCM was the first among the stopping rule effort regulation models to take temporal dynamics into account, and time is central in both stopping rules encompassed in it. In the context of the drift-diffusion model (Ratcliff et al., 2016), the time-linked stopping criterion was also introduced by Hawkins and Heathcote (2021), in addition to accumulating evidence (or confidence in the evidence, in metacognitive terms, Lee et al., 2023).

Figure 1. Examples *of intermediate confidence ratings while solving Compound Remote Associates, adapted from Ackerman (2014, Figure 4 and Figure 5).*



Note. Panel A: Submitting all solutions. Panel B: Solutions submitted when participants were allowed to respond "I don't know" without solving first.

Figure 1, adapted from Ackerman (2014), illustrates the regulatory process regarding quick (easy) versus slow (hard) tasks. Participants in the study provided confidence ratings every 15 seconds while solving problems of varying difficulty, with a final confidence rating when they stopped. Notably, one group submitted all responses (Panel A), while another could choose to respond "I don't know" instead of submitting responses (Panel B shows submitted responses).

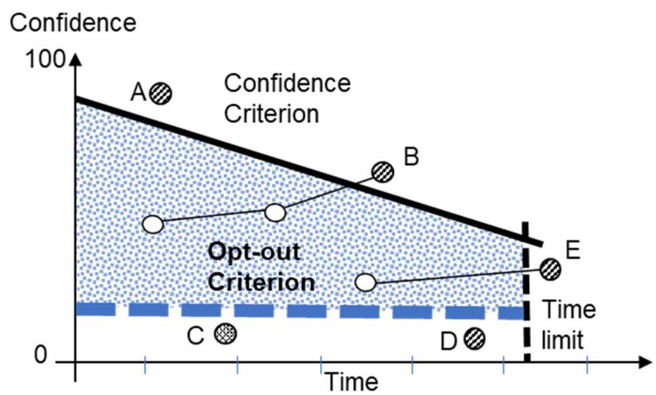
The solid lines in Figure 1 were generated by dividing each individual's response times for submitted responses (in Panel A, all items; in Panel B, only responses voluntarily submitted) into four quartiles. Notably, participants' initial judgments, provided after gleaning some impression of the problem and before the first 15 seconds elapsed, were highly predictive of final confidence and final thinking times. That is, when initial confidence was lower, the final response was submitted after more time and accompanied by lower confidence relative to items where initial confidence was higher. Moreover, as the figure shows, confidence consistently ends up higher than its starting point, exposing the confidence accumulation mentioned above. In cases where initial confidence is particularly low, intermediate confidence stays similarly low throughout the solving process but does not drop below its starting point.

The dashed lines in Figure 1 show the classic negative time–confidence association for both conditions. According to the DCM (Ackerman, 2014), these negative correlations arise from adherence to strategic stopping rules: some problems are addressed quickly with high confidence, while solving others takes longer, continuing until either the diminishing confidence criterion is met or the allotted time runs out. If people adhered to a pre-defined fixed confidence criterion, as suggested by the single-criterion models described above, then in cases where opting out is available they should provide only responses with confidence at this pre-defined criterion, rather than choosing to submit responses with lower confidence. However, permitting opting out (Panel B) did not expose a pre-set fixed criterion, but continued to produce a negative time–confidence correlation, suggesting adherence to a diminishing confidence criterion. Moreover, when allowed to opt out, participants waived their slowest responses, in line with adherence to a time limit. The DCM has been robustly supported across populations, tasks, and manipulated conditions, as well as with real-life multi-step tasks, like web searching and working with computerized applications (Ackerman et al., 2023; Ackerman et al., 2019; Ackerman & Levontin, 2024; Ackerman, Yom-Tov, et al., 2020).

The above literature review reveals that despite opting out being prevalent in real life, it has only rarely been the focus of research, although sometimes it has been used as a way to highlight characteristics of submitted responses (e.g., Ackerman, 2014; Koriat & Goldsmith, 1996). In particular, the metacognitive models associating monitoring with control over time and over the decision to opt out seem to miss a substantial phenomenon: according to the single-criterion models (Nelson & Narens, 1990; Ratcliff et al., 2016) opting out is not expected at all (Ackerman & Morsanyi, 2023), while by the DCM (Ackerman, 2014) opting out is expected only when reaching one's time limit without first meeting the confidence criterion.

The present study puts forward the *3-Stopping-Rule Model* (3SRM) for explaining the timeline of opting out, by adding to the DCM an *opt-out criterion*—the confidence level below which people waive further deliberation. See Figure 2. The main hypothesis guiding the present study is that people continue thinking as long as the combination of time and confidence does not satisfy any stopping criterion (shaded area in Figure 2). Both the confidence and opt-out stopping criteria may be met either quickly or slowly. When people come up with an answer at a confidence level exceeding their current confidence criterion (points A and B in Figure 2), or when they reach their time limit (point E in Figure 2), they submit their best answer candidate. The 3SRM's novelty is that under conditions in which opting out is legitimate, when confidence falls below the opt-out criterion, people conclude that there is no hope for improvement and terminate the solving process by opting out, regardless of the time invested till this point (points C and D in Figure 2).

Figure 2. The 3-Stopping-Rule Model (3SRM).



1.3. Measures and Predictions

The present study uses three classic direct measures: *success rate* (%), *mean response time* (RT), and *mean confidence* (%). It also uses the measures introduced above adapted from prior research on opting out: *output-bound accuracy* (success rate among submitted responses) and *control sensitivity* (gamma correlation between confidence and the decision to submit or opt out). Next, based on the DCM, I measure the *confidence criterion slope* (within-participant RT–confidence correlation for submitted responses). Given the little effect of permission to opt out on submitting behavior in Ackerman (2014; Figure 2 above), I also expected no such effect here.

Opting-out behavior was measured using three indicators: the *opt-out rate* (% of responses chosen to be excluded from scoring), the *opt-out criterion level* (mean confidence in opted-out responses), and the *opt-out criterion slope* (within-participant RT–confidence correlation for opted-out responses). Regarding the opt-out rate, at the task level, prior research documented differences between tasks. For instance, Ackerman (2014) found differences in DK rates between two open-ended (free answer entry) problem-solving tasks: 25% for the non-misleading Compound Remote Associate task and 7% for the Cognitive Reflection Test, a misleading task. Importantly, confidence was much higher in the latter task, with less opting out, hinting at potential strategic regulation. At the individual level, opt-out rate has been recently found to be an individual characteristic associated positively with cognitive ability, rational decision-making, and academic performance (4-18% DK responses, Law et al., 2022; Law et al., 2025).

The theoretical concept of opt-out criterion level can be interpreted in light of the classic *Signal Detection Theory* (SDT), which associated a criterion (or response bias) with motivation to succeed based on the payoff structure (see Wixted, 2020, for a review). While the classic context in which SDT has been applied is basic perceptual discrimination or stimuli identification tasks, Koriat and Goldsmith (1996) associated the opt-out rate and confidence adjustments to the incentive structure with the criterion-setting process under SDT: the decision to opt out is determined by comparing confidence in each answer to a criterion below which people prefer opting out over the risk of error (see Higham, 2007; Maniscalco et al., 2024, for similar conceptualizations of SDT using confidence in other contexts).

Koriat and Goldsmith (1996) empirically examined the strategic association between opt-out rates and confidence levels using an incentive structure that demotivated submitting wrong answers to a greater or lesser degree (+1/-10 vs. +1/-1 points for correct/wrong responses, respectively). They found that demotivating errors (+1/-10) rationally increased the opt-out rate (from 39% to 45%). At the same time there was also a rise in the maximum confidence level at which people still opted out. The present study includes conceptual replications of these findings, including an even stronger demotivating manipulation (Experiment 2) than that used by Koriat and Goldsmith (1996). *Accumulated points* were similarly used in the present study as a measure of variations associated with the incentive structures. Here, the *opt-out criterion level* was calculated simply as the mean confidence accompanying opted-out responses; it was expected to be sensitive to the incentive structure, as found by Koriat and Goldsmith (1996).

The heart of examining the unique contribution of the 3SRM is the association between the time invested till opting out and the confidence level associated with those responses, reflecting the *opt-out criterion slope*. My starting prediction followed Koriat and Goldsmith (1996)—namely, that the opt-out criterion remains constant, at the same confidence level, throughout the thinking process. This is a parsimonious model with one parameter: the confidence level below which people opt out. However, in light of recent insights into the effects of thinking time—as seen in the diminishing confidence criterion—it is essential to consider the possibility that the opt-out criterion is tilted as well, at least under some conditions. On the one hand, under the *sunk cost effect* (Arkes & Blumer, 1985), people may become less willing to opt out over time. The rationale here is that initial intermediate (or low) confidence encourages an early exit, but with the passage of time people become loathe to waste prior thinking efforts by quitting late, leading them to continue deliberation even with confidence levels that would have led to opting out early on. On the other hand, we can expect an upward tilt of the opt-out criterion based on the core idea that underlies the *Region of Proximal Learning* model (Metcalfe & Kornell, 2005). By this model, people stop investing effort in learning when they do not feel they are making sufficient progress. Applying this reasoning beyond the learning context, to problem-solving or answering knowledge questions, a similar process might take effect as well: people may be initially reluctant to opt out, and so the confidence criterion for opting out starts off low. Later, as they invest time without seeing sufficiently rapid progress, answers for which confidence rises too slowly are waived. These potential effects on the opt-out criterion slope are empirically tested in the present study. Specifically, I examined the effects of task characteristics and motivational considerations on the tendency to opt out quickly vs. slowly.

1.4. Overview of the Study

The present study includes two new experiments (Experiment 1 and Experiment 2) and a re-analysis of a published study (Experiment 3) that systematically test opting-out behavior while attempting to generalize as much as possible. The experiments involve two populations and languages (Experiment 1), two tasks known to differ in their time–confidence slopes when submitting responses (across experiments), three opt-out wordings (Experiment 1 and Experiment 3), and three motivational schemes (Experiment 2). The specific conditions and manipulations are explained as part of the introduction to each experiment.

In prior research, opting out has been operationalized in various ways: allowing respondents to cross out already-provided answers (e.g., Koriat & Goldsmith, 1996; Strudwicke et al., 2023); permitting them to simply refrain from answering (Shapira & Pansky, 2019); or requiring them to choose explicit "I don't know," "Get help," or "Uncertain" response options (Hanczakowski et al., 2013; Law et al., 2022; Undorf et al., 2021). In the present study, to examine within-participant time–confidence associations, answers and associated confidence ratings were required for all responses, regardless of whether the responses were submitted or opted out. Hence, all experiments operationalized opting out through an explicit response.

All procedures were performed in compliance with relevant laws and institutional guidelines and were approved by the institutional ethics committee (Certificate no. 2022-024).

2. Experiment 1 – Baseline

Experiment 1 provides the baseline for the effect of allowing opting out on effort regulation. The aim was to provide the default pattern of opt-out behavior in terms of rate, confidence level, and RT–confidence association slopes.

To verify the robustness of the findings, two sub-experiments were conducted, both online. In both cases, the task was multiple-choice verbal analogies in the participants' mother tongue (e.g., "Chick–Hen is like Calf–_______" [Cow]). An advantage of a multiple-choice task is that the items do not differ in answer entry time, which might correlate with task difficulty (e.g., rare words tend to be longer than common ones, requiring more time to type the full word).

A pilot study ($N = 126$) examined the effects of this answering format on opt-out measures. One group reflected the typical real-world answering process, by being allowed to opt out (by responding "I don't know," DK) without a requirement to provide an answer first. The second group had to answer using the procedure essential for providing the required measures for the present study: they could opt out only after providing an answer and confidence in each response. This pilot revealed that, overall, participants opted out in about

13% (SD = 12.5) of the problems, with no difference between the groups, p = .13 (there was a slight but not significant tendency to opt out less in the group where answering first was not required). Thirteen percent is an intermediate rate of opting out relative to prior problem-solving research (Ackerman, 2014; Law et al., 2022; Law et al., 2025); I expected it to allow variability across participants and sensitivity to motivational manipulations. The pilot study revealed no significant differences in the other process indicators examined.[1] Thus, in all current study experiments, participants working with opting out options were required to solve the task before conveying their decision to submit or opt out.

In Experiment 1a, one group of undergraduates solved each problem, rated their confidence, and then decided whether to submit their answer or opt out by clicking "I don't know." Participants were instructed to aim at collecting as many imagined points as possible, under a balanced incentive structure, with +2/-2 points for correct/wrong responses, respectively, and zero points for opting out. In line with the loss aversion (see Mrkva et al., 2020, for a review) and loss attention literature (Yechiam & Hochman, 2013), this incentive structure encourages opting out over submitting potentially incorrect answers, supporting assessing participants' strategic regulation (Higham, 2007; Koriat & Goldsmith, 1996).

Experiment 1b compared three conditions. Its DK condition replicated Experiment 1a with paid participants drawn from the general public, while allowing for a comparison between DK and other conditions: a control group and a Skip group. The control group had to submit solutions to all items. This condition was included to allow probing into how the availability of opting out in the other groups affects submitted responses. The procedure for the Skip group replicated the DK group, the only difference being that opting out was described as skipping the item. Based on research dealing with individual traits (e.g., mindsets regarding intelligence, Dweck & Yeager, 2019) and threats to self-image (Holtgraves et al., 1997), I expected that some people might perceive skipping as easier than clicking an "I don't know" response, which explicitly admits ignorance. If true, this prediction should be supported by higher opt-out rates and higher opt-out criterion levels in the Skip group than in

[1] One measure that did show a significant difference between the pilot groups was control sensitivity, which was lower in the group that could opt out instead of answering (M = .49, SD = .68) than in the group that had to answer and rate confidence for all problems (M = .87, SD = .16), p < .001. However, this is a trivial finding given that confidence variability was trimmed at the lower end in the group that could opt out quickly without confidence ratings.

the DK group. Participants in all groups received a base payment, and points above zero were translated into a monetary bonus.

2.1. Method

2.1.1. Participants

A power analysis was conducted using G*Power (Faul et al., 2009) for a linear multiple regression (fixed model, single regression coefficient). Assuming a medium effect size (0.15) and an alpha of 0.05, the results indicated that a total sample size of 53 would be required to achieve a power of 0.80.

For Experiment 1a, which used a relatively homogenous undergraduate population, I aimed for a more modest power (0.70; $N = 44$). Forty-five native Hebrew-speaking undergraduates were invited. After exclusion of three participants (see exclusion criteria below), 42 remained ($M_{\text{age}} = 25.4$ years, 71.4% reported being female). Demographics (gender, age) were collected on a dedicated form. Participants received course credit points for their participation, with no monetary compensation.

For Experiment 1b, I targeted the standard power per group (0.80; $N = 53$) to account for the increased variance expected from the more diverse Prolific population. One hundred and sixty native English speakers living in predominantly English-speaking countries were invited. After exclusion of nine participants, 151 participants remained, with 50 each in the DK and Skip groups, and 51 in the control group ($M_{\text{age}} = 34.9$ years, 52% reported being female). Their demographics were taken from the Prolific platform. They received 1.5 GBP as their base payment and could earn up to 60 pence as a bonus.

Exclusion criteria for all participants were as follows: (a) failure in most easy analogies used for attention verification; (b) no variability in confidence ratings (e.g., all 75%); (c) RTs and success rates below 2SD of the sample; (d) switching to a different computer window for more than 25% of the task time; and two item-level exclusion criteria if met for more than 20% of the items: (e) focus away from the experiment window for more than 50% of measured RT; and (f) particularly lengthy RTs (3SD or more from RT mean), suggesting that the participant was distracted or multi-tasking when solving these specific problems. Participants were excluded if they failed at least two of the six criteria.

2.1.2. Materials

The Hebrew analogies used for Experiment 1a included 26 problems and one example, developed by piloting with fluent Hebrew speakers. The analogies used for Experiment 1b included 35 English analogies and two examples used by Ackerman, Yom-Tov, et al. (2020). Five analogies were easy (success > 90% in a pilot study) and designated for attention verification. Success rates in the rest of the items were between 35% and 85%, ensuring room for confidence variability at both ends of the confidence scale.

2.1.3. Procedure

In the invitation for both sub-experiments, participants were informed that the task would involve solving challenging problems based on vocabulary. The experimental session opened with a description of the task. Participants were instructed to complete the task based on their own knowledge, without consulting the Internet, books, or other people. The task was first demonstrated by one example with a brief explanation of the associative similarity between the two-word pairs. In the English version, the initial example was "Chick–Hen is like Calf–_______" (Cow). Participants were then given an easy example to practice on, "Paw–Cat, Pincer–______," with the answer options Dog, Scorpion, Elephant, and Horse, ordered randomly. In each item, clicking one of the answer options activated a confidence scale and deactivated the answer options, so that the answer could not be changed. The confidence scale ranged between "I was definitely wrong" (0%) and "I was definitely correct" (100%). Participants had to move a marker on the confidence scale away from its initial middle point. After reporting their confidence, participants clicked either a "Submit" button or "I don't know" / "Better skip it…" when available. This triggered the appearance of the next analogy.

In Experiment 1a, the incentive structure was explained in terms of imagined points to be gained or lost. In Experiment 1b, the incentive structure was explained in terms of points entitling the participant to a bonus of up to 60 pence, with each balance point above zero worth a penny. For all groups, the incentive structure was balanced, with 2 points added to the participant's balance for each submitted correct response and 2 points deducted for each submitted incorrect response. For both the DK and Skip groups, opting out caused no change in the participant's points balance. A legend graphically presenting the incentive structure appeared below the confidence scale throughout the task. The analogies were presented in a random order generated for each participant.

2.2. Results and Discussion

Table 1 presents group averages for all measures. Descriptively comparing the results of Experiment 1a and Experiment 1b's DK group reveals very close similarities in success rates, opt-out frequency, and output-bound accuracy (success in submitted responses). As expected, consistent with meta-memory research (e.g., Koriat & Goldsmith, 1996), control sensitivity was very high in both experiments. These findings support generalization across languages and populations (students vs. the general population). Nevertheless, confidence was higher and RTs were much shorter in the Prolific sample than among the undergraduates, hinting at differences between the populations' thinking styles and/or effects of the incentive structure, which had monetary implications in Experiment 1b only.

Within Experiment 1b, comparing global success rates, mean confidence, and mean RT across the groups revealed no significant differences, all *p*s > .25. This similarity between the groups is important theoretically, as it suggests that the mere possibility of opting out did not substantially change participants' retrieval and thinking processes. A one-way analysis of variance (ANOVA) on output-bound accuracy (success rate of submitted responses) did reveal significantly lower accuracy in the control group relative to those who could opt out, $F(2, 148) = 6.95$, $MSE = 1558.63$, $p = .001$, $\eta_p^2 = .086$. This finding suggests that participants benefited from opting out. Notably, there were no differences whatsoever between the DK and Skip groups, all *p*s > .25. In particular, in contrast to predictions based on traits' literatures (Dweck & Yeager, 2019; Holtgraves et al., 1997), there were also no differences between the two opt-out wordings in opt-out rates or opt-out confidence levels, supporting rational decision-making in terms of the monetary incentive structure.

Table 1

Experiment 1 - Analogies means (SD).

Group Decision	N	Success rate (%)	Points[1]	Control sensitivity	Confidence mean (%)	RT mean (sec)	RT–Conf. corr.[2]
1a – Hebrew analogies, undergraduates							
With DK	42	68.2 (14.0)	21.6 (12.3)	.90*** (.17)	67.7 (11.2)	30.0 (11.2)	
- Submitted – output-bound		73.6 (12.7)			73.1 (10.2)	30.1 (11.2)	-.37*** (.26)
- Opted out (DK) (M = 13.3%, SD = 10.6)[3]		29.7 (28.0)			30.1 (13.0)	32.1 (19.1)	.10ns (.63)
1b – English analogies, Prolific							
With DK	50	69.3 (14.6)	28.1 (19.8)	.94*** (.09)	75.9 (14.8)	12.4 (4.0)	
- Submitted – output bound		74.4 (16.8)			81.2 (14.8)	11.9 (4.0)	-.29*** (.22)
- Opted out (DK) (M = 12.8%, SD = 15.3)		33.9 (28.2)			39.4 (15.1)	17.3 (6.6)	.03ns (.62)
With skip	50	70.6 (13.0)	31.1 (16.6)	.94*** (.08)	73.6 (10.7)	13.9 (5.3)	
- Submitted – output bound		76.9 (14.4)			80.4 (10.4)	13.1 (5.0)	-.40*** (.22)
- Opted out (Skip) (M = 14.8%, SD = 10.2)		36.6 (25.8)			36.5 (14.8)	19.6 (9.7)	-.06ns (.50)
Control group	51	66.2 (13.5)	22.6 (18.9)	---	75.1 (12.2)	12.8 (5.2)	-.35*** (.25)

Note: RT = Response time, DK = "I don't know." Control sensitivity and RT–Confidence correlations are the means of within-participant correlations.

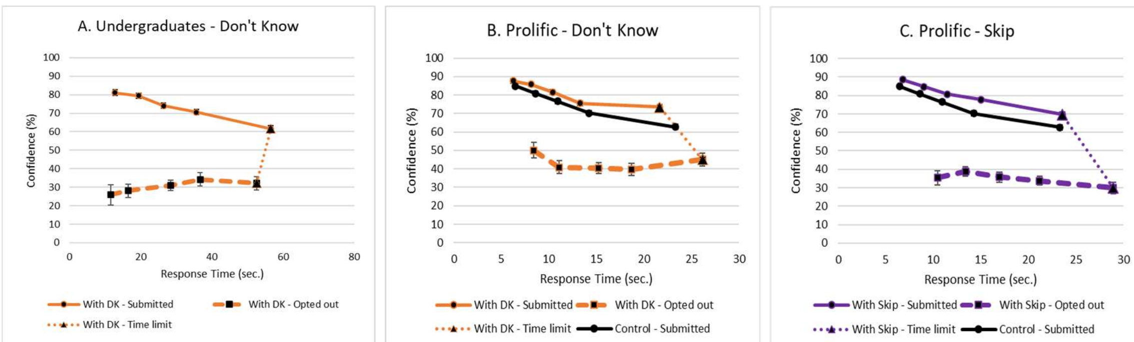
** $p < .01$, *** $p < .001$ for the difference of within-participant correlations from zero. ns – non-significant.

[1] Although the actual paid bonus could not be negative or exceed 60 pence, points could be negative or above 60.

[2] In terms of stopping rules, for the submitted responses, this correlation reflects the confidence criterion slope. For the opt-out responses, this correlation reflects the opt-out criterion slope.

[3] Opting out rate mean (SD), as a percentage of the problems.

Figure 3. Experiment 1: Response time (RT)–confidence association for submitted and opted-out responses among (A) undergraduates with a "don't know" (DK) option, and the general public from Prolific, with (B) DK or (C) "Skip" as the opt-out wording. Panels (B) and (C) also show results for their common control group, who had to submit all solutions.



Note. Error bars represent the standard errors of the means (the error bars for submitted responses are present, but too small to be seen).

To examine opt-out behavior, Table 1 separates the experimental groups where opting out was permitted into submitted and opt-out responses. Figure 3 visualizes the RT–confidence associations for submit and opt-out responses. It was drawn by splitting the RTs of each participant within each decision, submit or opt out, into five bins. The figure shows the means of RT relative to the X-axis and the means of confidence relative to the Y-axis. Importantly, the division into fifths has no theoretical meaning. It is meant to illustrate the pattern of the RT–confidence association in a way that does not enforce linearity, as correlations do, while still eliminating biases introduced by extreme values. Nevertheless, it is notable that the lines in the figure are almost straight.

In all cases, the confidence criterion diminished as thinking time lengthened, replicating previous findings with and without opting out (Ackerman, 2014, see Figure 1). In Figure 3, Panel B and Panel C, from Experiment 1b, show graphically how minor was the difference in confidence for submitted responses between the control group and those who could opt out. This finding also replicates Ackerman's (2014) findings, as presented in Figure 1 here, despite the differences in design between the two studies—specifically, the collection of intermediate confidence ratings only in Ackerman (2014), and the fact that in Ackerman (2014), but not here, participants used opting out without answering first. As a secondary point, note the dotted lines in the figure, which connect the submit and opt-out responses with the longest

RTs. These dotted lines visually support the theoretical time limit posited by the DCM (Ackerman, 2014). This aspect of the findings is considered in the General Discussion.

Importantly, in Figure 3, where the data are split for purposes of illustration into fifths, it is evident that although there were some slow opt-out times (the lengthiest fifth of the opt-out responses), these were the minority. Thus, participants largely did not reach their time limit, as would be expected by the DCM alone, but opted out before the time point when, in the absence of a DK or Skip option, they would have submitted their answer. In fact, they opted out throughout the time they allowed themselves to think about each problem. This detailed analysis complements prior findings regarding the potential of opting out to reduce labor in vain (Bae et al., 2021). Still, the figure also shows that participants did not always opt out as quickly as one might expect, but sometimes lingered over the problem before opting out.

As for the opt-out criterion level, confidence levels for opted-out responses were consistently lower than those for submitted responses. This is in line with the almost perfect control sensitivity, suggesting that, when possible, people prefer to avoid submitting lower-confidence responses. Probing more closely, clearly, opting out took place with confidence levels higher than zero (the lower end of the confidence scale). More substantial is the finding that participants opted out with confidence levels higher than 25%—that is, chance level: the likelihood of getting the answer right based on a wild guess among the four options in both Experiment 1a, $t(33) = 2.29$, $p = .014$, $d = 0.39$, and in Experiment 1b, $t(33) = 5.56$ $p < .001$, $d = 0.95$ with DK and $t(42) = 5.10$, $p < .001$, $d = 0.78$ with Skip. In terms of rational thinking, this finding may suggest that participants could rule out one answer option and still preferred to opt out over submitting low-confidence answers. Within Experiment 1b, no difference in opt-out confidence levels was found between the two opt-out wordings, $t(75) = 0.85$, $p = .20$, $d = 0.20$.

Finally, supporting the parsimonious model assuming a constant opt-out criterion, the opt-out criterion slope was flat—not associated with thinking time (non-significant RT–confidence correlations), despite the vast room participants had on the confidence scale to allow it to tilt upwards or downwards. Notably, however, calculating the significance of the mean correlation across participants does not take into account the significance of the correlation for each participant. A closer examination of participants who opted out more than twice (and whose slope significance was therefore meaningful) revealed that in Experiment

1a, two out of 25 slopes were significant, and in Experiment 1b, two out of 34 in the DK group and three out of 43 in the Skip group were significant. These findings further support the robustness of the non-significant slopes, while demonstrating that significance could be achieved, but was rare. Experiment 2 further examined these opt-out patterns.

3. Experiment 2 – Motivation and Individual Differences

The main goals of Experiment 2 were to (a) examine whether the opt-out criterion would tilt upwards, tilt downwards, or remain flat (the null hypothesis) with larger samples, (b) examine stronger motivations to opt out, and (c) examine whether individual characteristics associated with monetary incentives as a motivational force would predict this tendency.

Typically, in both metacognitive and decision-making research, incentives are framed such that participants are rewarded for success in the task, with or without losses in case of errors (see Goldsmith, 2016; Yechiam & Zeif, 2023, for reviews). In Experiment 2, the *Control group* was a replication of the DK group from Experiment 1b, using a classic balanced incentive structure and no explicit reference to thinking time. In addition, this study examined two incentive structures expected to affect opt-out behavior, one focused on efficiency and the other on loss. In light of the high similarity between the DK and Skip opt-out wordings in Experiment 1, all conditions in Experiment 2 used DK as the opt-out wording.

Examining the sunk-costs fallacy, Soman (2001) compared the effects of investment in money vs. time. He found that monetary investment had a greater effect than investment in time, unless time was associated with money as well. For the *Efficiency group* in the present study, I aimed to explicitly associate money and time by incentivizing participants to work quickly, but without explicitly forcing a time frame, to avoid artificially affecting time regulation. For this purpose, this group worked under the balanced incentive used for the control group, but with additional monetary incentives based on correct solutions per minute (see also Ackerman, 2023). This incentive structure was expected to elicit more and quicker opting out.

The second incentive structure examined here focused on losses. Carsten et al. (2019), using a Stroop task, found that loss-based incentives led participants to invest more time than reward-based incentives (see also Massar et al., 2020). It is as yet unknown how emphasizing losses affects thinking time, confidence, and success when opting out is available. In the *Loss group* of the present study, participants began the session with the maximum possible bonus and lost four pence for every incorrect response. Here, as in all other studies, opting out was cost-free (i.e., it led to neither gains nor losses). Notably, with this incentive structure, the most rational approach is to opt out of all problems as soon as possible. In fact, a second pilot study (N = 30) examined the effects of the loss only incentive structure. It revealed that participants rarely exploit this opportunity to obtain the maximum bonus while investing the minimum amount of time. Specifically, although the total opt-out rate rose from 13% to 22%, only five participants opted out in more than half the items, and no participant opted out in all problems.

The theoretical question was how the efficiency and loss manipulations would affect the opt-out rate, opt-out confidence criterion level, and opt-out RT–confidence slope. The opt-out rate and confidence level were expected to rise in both groups. In the efficiency group, they were expected to rise due to the attempt to work quickly and the social legitimacy of opting out under the efficiency condition (Tsui, 1991).

As for the opt-out criterion slope, by the sunk cost effect (Arkes & Blumer, 1985) we would expect the opt-out slope to tilt downwards, so to reduce late opting outs; this would be especially so under the loss condition. By contrast, under the Region of Proximal Learning model (Metcalfe & Kornell, 2005) we would expect it to tilt upwards for quitting items with little confidence gain, in particular when the aim is to work efficiently. Again, as detailed earlier, the present study takes as a starting point the parsimonious model, assuming that the criterion would remain flat despite these motivational differences between the groups. These predictions were preregistered (https://osf.io/n26gw/?view_only=9cb9e6fa490742c5823a093ae8a52711). It is worth mentioning that beyond the examination of strategic opt-out behavior, higher opt-out rates expected under the efficiency and loss conditions were expected to increase the power of analyses examining the opt-out criterion slope.

In addition to the expected effects of the motivational manipulations, in this experiment, I also considered individual differences. Prior metacognitive research regarding individual differences has been focused on factors associated with success or with metacognitive judgments (e.g., Grabman & Dodson, 2024; Kruger & Dunning, 1999; Pennycook et al., 2017; Stankov et al., 2014; Strudwicke et al., 2023). Highly relevant is the study by Law et al.

(2022), who suggested that opt-out rates reflect a consistent personality trait independent of confidence and intelligence. Notably, they measured opting out in different tasks from those in which confidence was measured. I am not aware of prior metacognitive research considering individual differences in the association between metacognitive monitoring and opt-out decisions in the same task, specifically not while taking the regulation of time into account.

Here, continuing the line of thought focused on efficiency and losses, I examined how opt-out behavior relates to participants' age and perceived socioeconomic status (SES). Typically, studies examining individual differences control for age and SES (e.g., Kleitman & Moscrop, 2010). However, previous research has revealed potential associations between sensitivity to rewards and losses, on the one hand, and both age and SES, on the other (Albert & Duffy, 2012; Guttman et al., 2021; Knutson et al., 2011; Mikels & Reed, 2009; Weissberger et al., 2022; White et al., 2022). For instance, Gächter et al. (2022) found that loss aversion increases with age and socioeconomic variables. Therefore, in this experiment, I used participants' age obtained from their Prolific profiles, and their perceived SES and self-reported income using a self-report questionnaire. I then examined how these variables are associated with the regulation of time and the three opt-out measures—opt-out rate, opt-out criterion level, and RT–confidence slope.

From a methodological perspective, the confidence scale captions used in Experiment 1 started at 0% confidence ("I was definitely wrong"). Rouy et al. (2022) highlighted that a full 0–100 confidence range might confound with the target effects examined, in particular in cases of low knowledge, and recommended using scales that start at chance level (i.e., the actual point reflecting maximum uncertainty). In order to examine whether Experiment 1's scale framing was responsible for the opt-out behavior seen there (e.g., by blurring the differences among uncertainty levels), in Experiment 2 the confidence rating scale started at 25%—the chance of guessing correctly when answering a four-item multiple-choice question. The caption at the low end of the scale read "A wild guess," while the caption at the high end was, as in Experiment 1, "Definitely correct" (100%).

3.1. Method

3.1.1. Participants

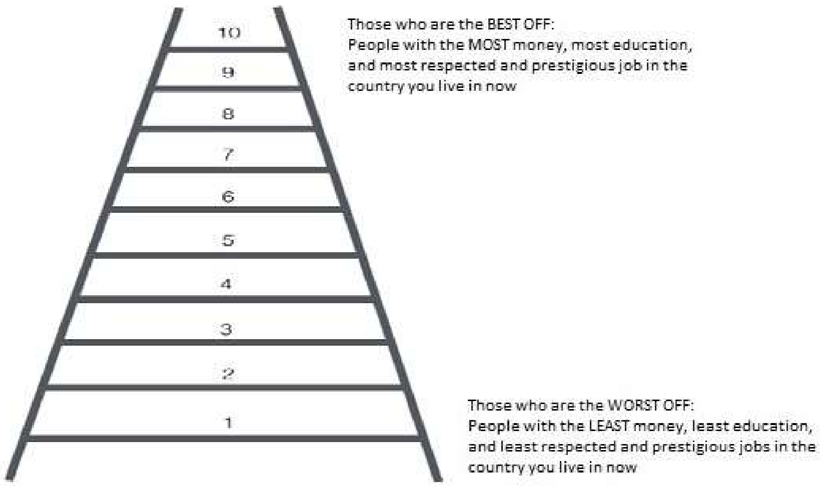
To calculate the opt-out criterion's slope while allowing enough variability to expose individual differences, the sample size was designed to be even larger than required based on G*Power for Experiment 1. The sample comprised 191 Prolific participants (M_{age} = 36.6 years, 56% reported being female); nine additional participants were screened out based on the same screening criteria as in Experiment 1. Of the final sample, 66 were randomly assigned to the balanced (control) group, 62 to the efficiency condition, and 63 to the loss condition. Participants received 2 GBP as a base payment and could earn up to 60 pence as a bonus.[2]

3.1.2. Materials

The analogies used in Experiment 1b were used here as well. Regarding the individual characteristics, as mentioned above, participants' age was taken from Prolific reports. For perceived SES, the main measure used was a ladder (Figure 4), with instructions adapted from O'Shea and Ueda (2021): "Please type the number (1–10) where you think you stand at this time in your life, relative to other people in the country you live in now." To complement this with a more concrete measure, participants were also asked to self-report their annual income (from 1, *Less than $20,000 per year*, up to 8, *More than $140,000 per year*).

[2] The base payment was set by Prolific, which updates its payment standard periodically. Hence, the payment was slightly higher than that offered in Experiment 1.

Figure 4. Perceived Socioeconomic Status (PSES) ladder.



3.1.3. Procedure

The procedure was similar to that of Experiment 1b, except for the 25%–100% confidence scale. A legend presenting the relevant incentive structure appeared on the screen throughout the task. The balanced incentive group replicated Experiment 1b, with +/-2 pence for correct and wrong submitted responses, respectively, and a maximum reward of 60 pence. For the efficiency group, the response options were the same. However, the incentive structure was different: participants could gain only up to 40 pence based on the balanced structure described above, rather than 60; and they could earn an additional 20 pence if they submitted more than four correct solutions per minute. This target efficiency level was chosen based on the average efficiency shown in Experiment 1b (*M* = 3.6, *SD* = 1.2), rounded upwards. Thus, I expected that most participants would indeed be able to meet this criterion if they sincerely aimed to work efficiently.

For the loss group, the instructions explained that they would start a priori with a bonus of 60 pence but would lose four pence for each wrong solution submitted. Submitting a correct solution or opting out did not affect the bonus.

The self-report and demographic questions were presented after participants completed the problem set. These items were presented in the same order for all participants.

3.2. Results and Discussion

Descriptive statistics appear in Table 2. As planned, the balanced incentive control group replicated most measures of the DK group of Experiment 1b. Although overall success and output-bound accuracy were highly similar, success in opted-out responses (i.e., correct solutions chosen but not submitted) was higher in this experiment than in the previous one. Confidence was not substantially higher than in Experiment 1b, despite the confidence scale starting at 25% rather than at 0%. Nevertheless, RTs were longer and RT–confidence slopes were more strongly negative. Most importantly, the opting-out patterns were fully replicated.

The two manipulations—i.e., the efficiency and loss incentive framing—resulted in several differences from the balanced control group, indicating that the participants took them into account and attempted to act accordingly. Notably, these differences were not in the classical measures of success rate and confidence, both *p*s > .10, but in the regulation of time and opting out, which are at the focus of the present study. As predicted, RT showed a significant effect of the manipulations, $F(2, 188) = 9.04$, $MSE = 248.58$, $p < .001$, $\eta_p^2 = .088$, with the efficiency group investing significantly less time than the other two groups, both *p*s < .01, and no difference between the loss and control groups, $p = .54$. This is, in fact, a manipulation check.

Turning to the main purpose of this experiment, let us consider the opt-out rate first. Although the most rational behavior for the loss condition was to opt out of all problems as quickly as possible, only one participant indeed exploited this option. On the other hand, only 19% (12 participants) in the loss group did not opt out at all, fewer than in both the efficiency group (35%) and the balanced control group (36%). As predicted, an ANOVA revealed a significant effect of the incentive condition on the opt-out rate, $F(2, 188) = 8.05$, $MSE = 2687.28$, $p < .001$, $\eta_p^2 = .079$, with the loss group opting out almost twice as much as the other two groups, both *p*s < .005, and no difference between the efficiency and control groups, $p = .99$. The finding that the efficiency group worked quickly but did not opt out more than the control group, as expected, is interesting, because it suggests that they succeeded in solving the problems more efficiently without compromising on the overall quality of their solutions.

Table 2

Analogies - Experiment 2 Means (SD).

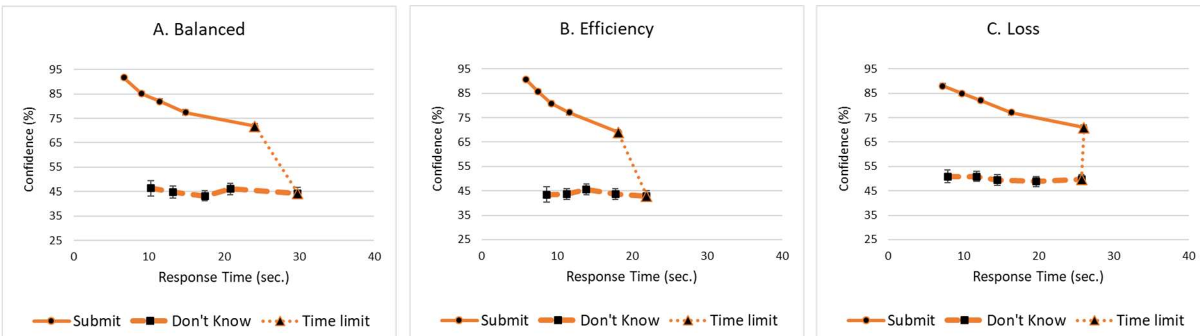
Group Decision	N	Success rate (%)	Points	Control sensitivity	Confidence mean (%)	RT mean (sec.)	RT–Conf. corr.
Balanced control group	66	70.0 (15.2)	27.8 (19.5)	.90*** (.18)	77.2 (11.0)	13.9 (4.3)	
- Submitted – output-bound		73.6 (16.2)			81.7 (10.5)	13.3 (4.1)	-.41*** (.19)
- Opted-out (DK) (*M* = 12.0%, *SD* = 14.1)		48.4 (31.8)			42.8 (14.4)	19.7 (9.5)	-.07ns (.49)
Balanced + efficiency incentive	62	66.1 (14.0)	32.3 (25.3)	.86*** (.33)	76.3 (9.1)	11.0 (3.1)	
- Submitted – output-bound		70.2 (16.4)			80.9 (9.5)	10.5 (2.8)	-.46*** (.22)
- Opted-out (DK) (*M* = 11.6%, *SD* = 13.1)		37.6 (30.1)			43.9 (13.8)	16.4 (7.2)	-.11ns (.50)
Loss incentive framing	63	66.4 (16.6)	29.7 (21.0)	.85*** (.32)	73.5 (10.5)	14.9 (7.4)	
- Submitted – output-bound		73.6 (18.2)			81.3 (10.1)	13.9 (7.3)	-.38*** (.26)
- Opted-out (DK) (*M* = 23.1%, *SD* = 25.2)		41.4 (26.6)			46.8 (14.6)	20.5 (11.4)	0.0ns (.59)

Note: RT = Response time, DK = "I don't know." Control sensitivity and RT–Confidence correlations are the means of within-participant correlations.

** $p < .01$, *** $p < .001$ for the difference of within-participant correlations from zero. ns – non-significant.

When splitting the analysis by the decision to submit vs. opt out, no significant differences among the groups were found, beyond those in RT. Looking at Figure 5 reveals the differences in response times among the groups, with the Efficiency group both submitting and opting out more quickly than the other groups. Most importantly, confidence in opted-out responses was always lower than that for submitted responses. Confidence in opted-out responses was higher than in the prior experiment, which might be due to the scale starting at 25% rather than at 0%. Nonetheless, again, there was substantial room for the opt-out criterion slope to tilt either downwards or upwards. Still, in all groups the RT–confidence slopes showed a highly similar pattern, with no significant differences between them, $F(2, 185) = 1.77$, $MSE = .090$, $p = .173$, $\eta_p^2 = .019$ for the Submit decision, and $F(2, 111) = 0.411$, $MSE = .118$, $p = .664$, $\eta_p^2 = .007$ for opting out.

Figure 5. Experiment 2: Response time (RT)–confidence association for submitted and "don't know" responses under (A) balanced incentives control group, (B) balanced with the addition of efficiency incentives, and (C) incentives focused on losses.



Note. Error bars represent the standard errors of the means (the error bars for submitted responses are too small to be seen).

Turning to the individuals' opt-out slopes, significant slopes were again rare. Of the 42 participants who opted out more than twice in the Balanced control condition, only one showed a significant slope, and these ratios were similar in the other conditions (0/39 in the Efficiency group and 3/51 in the Loss group). Thus, the opt-out criterion remained tilted only rarely, despite the strategic adjustments of time management and opt-out rates to task requirements.

Table 3 presents the set of between-participant correlations, which enables examining individual differences potentially related to the perception of loss, which was present in all conditions. Given the consistent opt-out confidence levels and RT–confidence slopes across the three groups, the correlations were examined based on the entire sample. N is the number of participants who opted out more than twice (to allow assessing a meaningful opt-out criterion slope). In addition, one participant failed to answer the self-report questions, and one participant did not have age in their Prolific profile.

Table 3

Between-participant correlations between individual characteristics (1–3, age, SES-related measures) and opt-out behavior (4–6; pink and blue shading) in Experiment 2, for those who opted out more than twice.

No.	Measure	Mean	SD	N	1	2	3	4	5
1	Age	33.7	10.7	105					
2	SES Ladder	5.26	1.52	105	.14				
3	Self income	2.76	1.65	105	.03	**.46****			
4	Opt-out rate	26.9	18.7	106	**-.22***	0.05	-0.07		
5	Opt-out confidence	44.2	11.6	105	-.06	0.12	-.002	**.30****	
6	Opt-out RT–conf. slope	.03	.47	105	-.08	-0.03	0.05	-.02	-.11

Bold - Correlation is significant ** p < .01, * p < .05 (2-tailed).

I highlight here some correlations, while readers are invited to consider others. The strong association between SES (2), which reflects subjective financial standing, and the more concrete self-reported income (3) is a kind of sanity check. Both were uncorrelated with age, stressing the contribution of examining both types of measures. The opt-out rate (4) was negatively correlated with age (1; pink shading). This finding may suggest less willingness to opt out as people age, but may also stem from higher objective or perceived verbal knowledge relevant to the particular task at hand (though see the accumulated data analysis following the reports for all experiments, below). In contrast, neither of the individual characteristics associated with economic status (2, 3) correlated with any of the opt-out measures (4–6; blue shading). This finding adds to the existing literature regarding the effects of SES on loss aversion, which thus far has yielded ambiguous results (e.g., Gächter et al., 2022; Guttman et al., 2021; Hansla & Johansson, 2020).

Turning to intercorrelations among the three opt-out measures (4–6, gray shading), the more confident participants were in their opted-out responses (5), the more they opted out (4). This finding may seem counterintuitive. However, in fact, this finding is derived from the current study's hypotheses based on the 3SRM, regarding the strategic setting of the opt-out criterion. Given two people with the same spread of confidence ratings, the person who sets their opt-out criterion higher is expected to opt out of more items than the person who sets their opt-out criterion at a lower confidence level (because there are more instances in which confidence does not exceed the opt-out criterion). Thus, this correlation supports strategic adjustment of the opt-out criterion in line with the SDT principles adapted by Koriat and

Goldsmith (1996).

To sum up, it is particularly commendable that the efficiency group managed to work quickly, without changing their target confidence level and opting out policy, counteracting the speed–accuracy tradeoff. The loss group opted out more, but still kept the three stopping rules unchanged. In particular, the flat opt-out slope was robust across conditions and individuals.

4. Experiment 3 – General Knowledge Questions

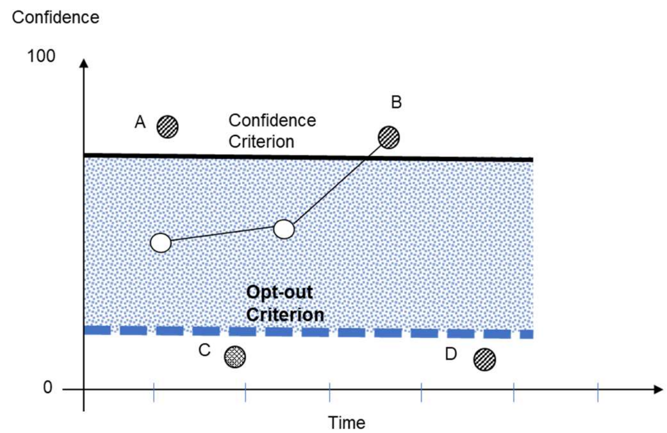
Experiment 3 is a reanalysis of data from Undorf et al. (2021, Experiment 3) in which response times were collected but not analyzed. In this experiment, participants answered general knowledge questions requiring numerical answers (e.g., In what year? How many?). Participants were instructed to respond using fixed ranges (e.g., 10-year intervals) within which they expected the correct answer to be found. Thus, the answering format was open-ended, without specific answer options offered. As in the previous experiments of the present study, participants first answered each question, rated confidence, and then could opt out under a balanced incentive structure (+/-2 pence for correct/wrong submitted answers, and no gains or losses for opting out).

The experiment included four groups: a control group that had to submit all answers; a DK group; a Help group; and a combined group that had both DK and Help options—a condition designed to more closely imitate the real-life range of people's opt-out choices. For the groups with the Help option, it was explained that pressing the Help button would allow them to answer the question again in the following block while being shown an answer from what was described as a credible participant from a previous experiment. Requesting help cost one point, but could allow gaining two points later if the participant successfully solved the item with a hint. The data relevant to the present study were the answers and related measures provided in the initial answering block before those in the Help groups received their second chance. Importantly, Undorf et al. found more opting out using DK responses (41%) than help requests (17%). The overall higher opt-out rate than in the analogies used in the previous experiments of the present study might be due to the greater effort required to come up with an open-ended answer, and/or the greater social legitimacy of admitting ignorance when facing questions about unfamiliar topics. The preference for DK over help could stem from a general reluctance to use help, or from strategic considerations aligned with the cost of using help. For the purpose of examining the characteristics of opting out, higher opt-out rates strengthen the statistical power of the analyses.

Importantly, beyond higher opt-out rates, unlike the case with analogies in Experiment 1 and Experiment 2 in this study, which consistently produced negative time–confidence correlations (Figure 1), the general knowledge task used by Undorf et al. (2021) provides a rare case showing flat or even weakly positive time–confidence slopes. This flat pattern is more in line with the classic pre-set single-criterion Discrepancy Reduction Model (Nelson & Narens, 1990) and Drift-Diffusion Model (Ratcliff et al., 2016). In perceptual decision-making researchers have noted a similar puzzle of different time–confidence slopes for different tasks (Hawkins et al., 2015), and attempted to explain the source for the difference (Smith & Ratcliff, 2022). However, the particular explanation offered is based on the task having a two-alternative forced choice structure, and does not fit the more cognitive demanding thinking course the present study aims to explain. A recent attempt to experimentally transfer a negative correlation situation into a flat one has faced challenges (Olschewski et al., 2025). In the present study, the combination of the two tasks—analogies with negative time–confidence correlations, and general knowledge questions with flat time–confidence correlations and more opting out decisions—allows for characterizing opting out behavior across time-allocation policies.

In terms of delineating a theoretical stopping rule model, the 3SRM was a development of the DCM (Ackerman, 2014), which described stopping rules for tasks that generate negative time–judgment correlations. With the DCM, the flat (or weak) time–confidence association suggests that the confidence criterion does not diminish. See Figure 6. The possibility of a time limit in this situation has not yet been considered in the literature. Thus, I adapt the confidence criterion from the single stopping rule models and add the opt-out criterion.

Figure 6. The opt-out criterion in cases of a flat time–confidence association.



Focusing on the characteristics of opt-out confidence and time, in the data re-analyses done as part of Experiment 3 I aimed to test whether the opt-out patterns are stable across patterns of submitting answers, or change when participants adhere to a constant, rather than diminishing, confidence criterion. Given the stability of opt-out behavior vis-à-vis the various individual characteristics and task-level manipulations examined in this study so far, I hypothesized that opting-out characteristics reflect a robust metacognitive process for dealing with uncertainty in a generalizable manner. If this hypothesis is supported, the substantially higher opt-out rate found here strengthens the findings of the previous experiments.

4.1. Method

4.1.1. Participants

Two hundred and twelve Prolific participants took part. They were randomly split into four conditions, with 52 participants in the control group, 54 in the DK group, 50 in the Help group, and 55 in the combined DK and Help group. A post-hoc power analysis examining the expected power of a two-tail test with one predictor and effect size of 0.15 revealed that with 52 participants the expected power is 0.78, and with 55 participants the expected power is 0.81.

4.1.2. Materials

Materials were 38 English general-knowledge questions that required quantitative or numeric responses in predefined answer intervals (e.g., 10-year ranges) and covered a broad range of topics (e.g., "How many keys does the standard piano have?", "In what year did the French Revolution start?"). One question served as a practice item ("In what year did Barack

Obama's US presidency begin?"), and four very easy questions served as attention-check items (e.g., "How many arms does an octopus have?"). Following Undorf et al. (2021), attention-check items were not included in the analyses. Thus, there were 33 target questions.

4.1.3. Procedure

Participants answered each question by providing a bounded range of values with a required interval. The required interval was presented along with the question. On the same screen, participants rated their confidence in the correctness of their answer on a 0 to 100% scale ("What is the chance that your answer encompasses the correct value?"). Participants were required to select a confidence judgment other than 50%, which was the initial position of the pointer on the scale.

After each confidence rating, in the control condition, participants had to submit all answers. In the DK condition, participants could withhold their answers by clicking an "I don't know" button instead of submitting. For the Help condition, a "Get help" button replaced "I don't know." In the combination condition, participants could choose between "Submit," "Get help," and "I don't know." Instructions in the help-seeking and combination conditions explained that each question for which participants sought help would be presented again in a subsequent phase, together with a response from a previous participant who was correct more often than 80% of participants in that previous study.

Instructions for all groups explained that each submitted correct answer would earn 2p, and incorrect answers would be penalized by 2p. Participants in the two conditions with a DK option were told that they would neither gain nor lose when opting out by clicking "I don't know." Participants in the two conditions with a Help option were told that 1p would be deducted from their score for each help request, and that answers in the help phase would be scored in the same way as answers submitted without seeking help (+/-2p for correct/incorrect answers). A legend detailing the incentives relevant to the respective condition was presented on all screens.

4.2. Results and Discussion

As can be seen in Table 4, the general knowledge task used in Experiment 3 was harder (46% success in the control group) than the analogies used in the previous experiments (66% success). This is not surprising given that in the previous experiments the task format was multiple choice, with four answer options to choose from. Not only does a multiple-choice format mean that partial knowledge may allow ruling out some answers, but mere guessing already makes the chance of success 25% (Vuorre & Metcalfe, 2022). In an open-ended test format, these aids are not available. More importantly, the global RT–confidence correlations were not negative, as seen in the previous experiments and many other problem-solving, decision-making, and memorization tasks. This makes this task highly valuable for further examining opt-out behavior. Against this background, I examined the differences between the groups while focusing on the regulation of time, which was not considered by Undorf et al. (2021).

Table 4

General knowledge - Experiment 3 Means (SD).

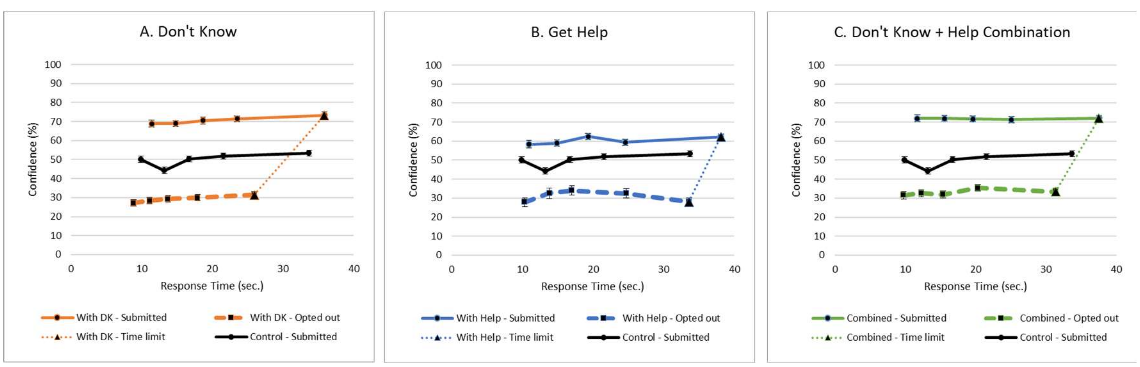
Group Decision	N	Success rate (%)	Points	Control sensitivity	Confidence mean (%)	RT mean (sec)	RT–Conf. corr.
DK option	55	48.9 (17.8)	10.8 (17.8)	.89*** (.13)	54.1 (15.7)	18.6 (9.4)	
- Submitted – output-bound		64.7 (21.5)			71.5 (15.0)	20.1 (9.8)	.06ns (.26)
- Opted-out (DK) (*M* = 40.6%, *SD* = 22.6)		27.2 (18.3)			29.7 (12.2)	16.8 (8.7)	.08ns (.41)
HELP option	50	53.8 (19.2)	8.1 (22.8)	.70*** (.27)	55.4 (13.8)	21.3 (9.1)	
- Submitted – output-bound		60.0 (21.9)			61.8 (15.6)	21.6 (9.3)	.01ns (.28)
- Opted-out (HELP) (*M* = 17.1%, *SD* = 21.7)		35.3 (28.5)			33.5 (15.4)	22.1 (13.2)	0.0ns (.51)
Combined DK+HELP options	55	49.5 (16.6)	10.9 (16.3)	.89*** (.14)	56.7 (12.3)	20.4 (7.9)	
- Submitted – output-bound		63.7 (18.5)			73.4 (11.4)	21.3 (8.1)	.01ns (.29)
- Opted-out (DK or HELP) (*M* = 38.9%, *SD* = 23.5)		25.4 (15.6)			32.5 (13.9)	18.7 (8.6)	0.0ns (.37)
Control group	52	45.8 (17.2)	-5.8 (22.3)	---	49.9 (13.0)	19.0 (8.2)	.11** (.25)

Note: RT = Response time, DK = "I don't know." Control sensitivity and RT–Confidence correlations are the means of within-participant correlations. * $p < .05$, ** $p < .01$, *** $p < .001$ for the difference of within-participant correlations from zero. ns – non-significant.

No RT comparisons reached significance, although the difference between opting out with DK and when requesting help was on the verge of significance, $p = .05$. Control sensitivity showed variability across the groups, with weaker associations in the help-only group than in the groups that had the option to respond with DK. This was an important finding by Undorf et al. (2021) for the control sensitivity literature, given the persistently high values that have typically been found among healthy adults.

Looking at Figure 7 reveals that, unlike in the previous experiments, here the control condition was in the middle of the confidence range, with the possibility of opting out allowing participants to submit answers where their confidence was higher and to withhold those where their confidence was lower. Before we delve further into opting-out behavior, an important finding was that the four groups did not differ in submit RT–confidence correlations, $F(3, 211) = 1.88$, $MSE = 0.14$, $p = .133$, $\eta_p^2 = .027$. This was so despite this correlation being tilted upwards in the control group, while in all groups that could opt out these slopes were flat, not significantly tilted upwards or downwards. See Figure 7.

Figure 7. Experiment 3: Response time (RT)–confidence association for submitted and opted-out responses when the options were (A) "Don't know" (DK), (B) "Get help," and (C) combined DK and Help.



Note. Error bars represent the standard errors of the means.

As for the opt-out patterns, clearly, the general knowledge task used here encouraged higher opt-out rates than the analogies used in Experiment 1 and Experiment 2. Interestingly, DK was used much more than help requests, and the opt-out rate in the combined condition was equivalent to when DK was offered alone, $F(3, 208) = 51.33$, $MSE = 19864.2$, $p < .001$, $\eta_p^2 = .425$, with the help-seeking rate lower than opting out in the other groups, both ps < .001 in a Tukey post-hoc test. Comparing opt-out rates while including only those who opted out at least once did not change this pattern, $F(2, 127) = 8.09$, $MSE = 3324.9$, $p < .001$, $\eta_p^2 = .113$, with the Help condition showing lower opt-out rates ($M = 27.7\%$, $SD = 21.8$) than both the DK ($M = 43.0\%$, $SD = 20.9$) and the combined condition ($M = 45.5\%$, $SD = 18.5$), both *p*s < .005. Thus, regardless of global willingness to opt out, help-seeking was used the least. As explained above, this reluctance to use help might stem from the monetary penalty associated with using it.

Looking into the spread of RTs among opt-out items in Figure 7, unlike the previous experiments, here participants of all groups opted out more quickly than when they submitted answers after the longest thinking time. Comparing the groups in this respect reveals no difference among them, $F(2, 127) = 2.81$, $MSE = 276.24$, $p = .064$, $\eta_p^2 = .042$. Confidence in opt-out responses did not differ among the conditions either, $F(2, 127) = 0.90$, $MSE = 167.4$, $p = .410$, $\eta_p^2 = .014$. Moreover, the opt-out RT–confidence correlation, reflecting the opt-out criterion slope, did not differ among the groups, $F(2, 127) = 0.58$, $MSE = 0.102$, $p = .564$, $\eta_p^2 = .009$, and none of the slopes differed from zero (Table 4). Slopes were significant for 3/52 who opted out more than twice in the DK group, and for 0/31 in the Help group. In the combined group, slopes were significant for 3/44 who opted out more than twice using DK, and for 3/33 who did so by asking for help. Thus, once more, significant slopes were quite rare.

In sum, the problem-solving task used in Experiment 1 and Experiment 2 and the general knowledge task used here clearly differ in many respects. The inclusion of Help as yet another opt-out wording revealed some differences from using DK. The combined condition was more similar to DK than to the Help condition. As for the opt-out behavior, this experiment decisively indicates that the opt-out criterion is lower than the submit criterion and is persistently flat, regardless of the confidence criterion slope.

5. Accumulated Data Analysis

5.1. Purpose and Method

In this section I merged the three experiments into a mega-analysis—a method that allows benefitting from a one large data set while controlling for the differences between experiments (Eisenhauer, 2021). This method is particularly beneficial for examining the robustness of null effects, as it tests for a weak effect that might be exposed with a substantially larger sample size. The main aims were (a) to examine the robustness of the flat RT–confidence association using several regression models and (b) consider potential moderators that might explain some of its variance using correlations across the individual-level variables available in the three experiments. Given the consistency of the flat association so far, I hypothesized that few such correlations would be found.

Overall, the combined data set includes 596 participants.

5.2. Results and Discussion

The analyses in this section were intended to discover predictors of the RT–confidence slope while benefiting from the elevated power of the accumulated data. First, to examine whether the opt-out RT–confidence slope was significant across all sets of experimental data, a regression was calculated based on each participant who opted out more than twice (N = 317) and who showed variability in confidence for opted-out items (N = 315). There were no differences among the experiments in opt-out RT–confidence slopes, $F(3, 311) = 1.55$, MSE = 11.02, $p = .20$, $\eta_p^2 = .010$. A Two One-Sided Test (TOST) revealed that the 95% confidence interval for these slopes was between -0.20 and +0.30, indicating that the mean slope can be considered equivalent to zero if one allows for an equivalence delta as low as 0.30. Another regression analysis considered all relevant dependent variables as predictors of the opt-out criterion slope (RT–confidence correlation), while controlling for the experimental predictors considered in the next correlational analysis (see Table 5). This analysis revealed no significant predictors either, all ps > .13 and $R^2 = 0.028$. Thus, so far, no predictors could be detected, indicating the robustness of the flat opt-out criterion slope.

The correlations between all considered dependent variables for all participants who were included in this analysis are presented in Table 5. My report here is focused on the three opt-out measures. First, those three measures (8–10, gray shading) did not correlate either with each other or with age (1, pink shading), canceling out the two correlations found in Experiment 2 (Table 3).

The opt-out rate (8) negatively correlated with success (2) and points (3) within tasks, supporting Law et al.'s (2022) associations between the opt-out rate and performance when obtained across different tasks for each measure. Focusing on stopping rules, the finding that the opt-out rate (8) had a positive association with the submit RT–confidence slope (7), but a negative association with global response time (6), is notable. By the 3SRM, people who worked faster than others opted out more and compromised less on their confidence in submitted responses.

Table 5

Between-participant correlations between age (1), cognitive and metacognitive measures (2–7), and opt-out behavior (8–10; pink and blue shading) across all three experiments, for participants who opted out more than twice.

No.	Measure	Mean	SD	N	1	2	3	4	5	6	7	8	9
1	Age	33.4	11.3	317									
2	Success rate	61.4	17.6	317	-.01								
3	Points	25.3	20.8	317	-.07	**.79****							
4	Control sensitivity	.89	.16	316	-.04	**.26****	**.23****						
5	Confidence mean	65.1	13.8	317	.04	**.61****	**.51****	**.16****					
6	Response time (RT) mean	17.3	8.8	317	.02	.06	**-.12***	.01	-.09				
7	Submit conf.-RT corr.	-.22	.32	313	.01	**-.44****	**-.46****	-.10	**-.42****	.10			
8	Opt out rate (%)	31.2	19.6	317	-.07	**-.54****	**-.26****	-.05	**-.60****	**-.10***	**.31****		
9	Opt out confidence mean	36.9	13.8	317	.07	**.22****	**.25****	**-.28****	**.59****	**-.19****	**-.16***	-.02	
10	Opt out conf.-RT corr.	.03	.46	315	-.04	-.06	-.05	.04	-.08	-.02	**.12***	.05	-.03

Bold - Correlation is significant ** $p < .01$, * $p < .05$ (2-tailed).

The opt-out confidence means (9) correlated with all cognitive and metacognitive measures (2–7). Particularly, they were most strongly associated with the general confidence mean (5), supporting prior findings of confidence as a general trait (Stankov et al., 2014). The negative correlation between the opt-out confidence means and global RT (6) is notable, as it strengthens once more the framing of the opt-out criterion based on the SDT and the link with Koriat and Goldsmith (1996), who associated opting out with confidence level. Adding the time aspect to the process reveals that people who respond more quickly (6) tend to set their opt-out confidence higher (9) and opt out more (8) than slower responders.

The main purpose of the accumulative analysis was to examine the robustness of the null effects found so far on the opt-out RT–confidence slope (10), and to probe further for individual differences in this respect. This analysis exposed only one, weak, correlation with the submit RT–confidence slope (7). A complementary analysis used the same data while entering all variables into a stepwise linear regression model that allows exploration of variable combinations as predictors. This analysis also revealed only the submit confidence slope as a predictor of the opt-out confidence slope. However, its R^2 was .014, meaning it explained only 1.4% of the variance, making even the submit confidence slope a weak predictor of the opt-out confidence slope. A quadratic model did not contribute much either, bringing the variance explained by the submit confidence slope to $R^2 = 0.019$ (1.9%).

In sum, taking advantage of the elevated power allowed by accumulating the data across experiments did not reveal any consistent effects on the opt-out criterion slope. These additional analyses support the stability of the flat opt-out criterion, given a consistent level of confidence that leads to opting out throughout the thinking time continuum.

6. General Discussion

Metacognitive research has dealt with stopping rules since its inception (Nelson & Narens, 1990), but research dealing with strategic opting out has been scarce and scattered (Koriat & Goldsmith, 1996; Law et al., 2025). Importantly, no prior metacognitive model has considered the time it takes to opt out. Beyond being a strategic behavior, opting out allows people to improve the accuracy of their provided responses (see Goldsmith, 2016). The present study used well-controlled experimentation to delineate the opt-out criterion, and to call attention to its role in everyday thinking as a means of saving time.

The starting point of the present study was initial evidence that people often opt out more quickly than suggested by prior metacognitive models (Ackerman, 2014; Metcalfe & Kornell, 2005), in parallel to the entire time range of submitted answers. Some prior research pointed to the possibility of quick opting out (e.g., Glucksberg & McCloskey, 1981; Metcalfe & Kornell, 2005; Singer & Tiede, 2008), but no prior research delved into the time–confidence association. The present study presents the opt-out stopping criterion as a means of addressing this gap using metacognitive concepts and a comprehensive set of measures across a variety of conditions. Their combination allows analyzing the opting out policy participants adopt in each condition.

6.1. Stopping-Rule Models

The present study extends prior effort regulation models that included one and two stopping rules, but that were unable to explain quick opting out. The consistent pattern of time allocation found here fully supports the 3SRM structure, with three stopping rules: a submit criterion based on confidence, which either diminishes (Figure 2) or does not (Figure 6), a time limit, and an opt-out criterion that is flat, unaffected by thinking time. The results of the three experiments yield nine replications (Figure 3, Figure 5, and Figure 7) and a global persistent pattern across the accumulated data sets (Table 5).

Going beyond all prior effort regulation models, the 3SRM covers cases of negative global RT–confidence correlations, as demonstrated in Experiment 1 and Experiment 2, as well as cases of no correlation, as demonstrated in Experiment 3. As mentioned above, decision-making research has already attempted to identify the reasons a task might show a negative or no correlation between time and confidence, with inconsistent results (Hawkins et al., 2015; Olschewski et al., 2025).

Regardless of the task type showing negative or no RT-confidence slopes of the submitted responses, the slopes were parallel to that found in the control groups, who could not opt out. While in Experiment 1 of the present study the confidence criterion under a DK condition was only slightly higher than in the control group (Figure 3), in Experiment 3 the confidence criterion was substantially higher when opting out was allowed than in the control group (Figure 7). The differences between the two may be due to the lower rate of opting out in Experiment 1 than in Experiment 3, which led opting out to have a smaller effect on the submitted responses. Another possible reason is that confidence was higher in Experiment 1 than in Experiment 3. Although there were no ceiling effects, there was nonetheless less room for confidence in quick responses to rise in Experiment 1 (from 80–90%) than in Experiment 3 (from 50%). Future research is called to consider the effects of opting out on tasks with negative time–confidence correlations and generally lower confidence.

The time limit is the least studied component of the DCM (Ackerman, 2014). Nascent research has started to delineate the factors that affect it (Ackerman et al., 2023; Ackerman & Levontin, 2024; Undorf & Ackerman, 2017). In Figure 6, a time limit was not included, as it was not part of any prior model with a pre-set confidence criterion. However, the results suggest that there was a time limit in both task types. In particular, the maximum time was equivalent in all experiments between the control group and the submitted responses by groups who could opt out.

What is the shape of the time limit? By the DCM, the time limit is a constant maximum time beyond which people are not willing to invest any more effort (Ackerman, Yom-Tov, et al., 2020; Undorf & Ackerman, 2017). However, the results of the present study (see Figure 3, Figure 5, and Figure 7) raise the possibility that different factors affect the time limit for opting out and for submitting. For instance, with open-ended knowledge questions in Experiment 3 (Figure 7), when confidence was generally lower and opting out was used more often relative to the multiple-choice analogies, opting out might be perceived as more legitimate and thus used more quickly than the slowest submitted responses. The potential association between general confidence level in a task and the respective time limits for submission and opting out warrants further investigation.

As for the confidence level at which people opt out, this was not sensitive to most manipulations tested in the present study. This was the case although most manipulations did affect some measures, hinting both that they had enough salience and that the study had sufficient power to detect such effects. Nevertheless, confidence ratings for opt-out responses were higher (42–47%, Experiment 2) when the scale started at 25% than when it started at 0% (30–40%, Experiment 1 and Experiment 3). It is unclear whether this difference is merely technical or carries theoretical meaning. In both cases, the findings suggest that people use a consistent level of confidence as the threshold below which they opt out, across motivational factors and individual characteristics of age and SES. More indicative of effort regulation are the correlational analyses revealing that people who performed better, with higher confidence, and more quickly, also opted out with higher confidence (Table 5, correlations of measure 9 with measures 2–7). Thus, opt-out confidence criterion level seems to be associated with the (perceived) ease of performing the task.

The present study uniquely aimed to delineate the slope of the opt-out confidence criterion while considering variability in the time it takes people to opt out, given the effect of time on the confidence criterion for submission (DCM, Ackerman, 2014). The results consistently show that the opt-out criterion is stable throughout the thinking time people allow for each task item, with opting out occurring with similar confidence levels early and late in people's allocated thinking time. These findings support the parsimonious model presented by Koriat and Goldsmith (1996), who did not take the time into account. Thus, the present study's experiments mostly exposed a list of factors that did not affect the opt-out criterion slope. This pattern was robust across populations and languages (Experiment 1), participant age (all experiments) and socioeconomic characteristics (Experiment 2), motivational conditions (Experiment 1, Experiment 2), tasks that differ in their typical effort regulation pattern (Experiment 3 relative to the other two experiments), confidence scale ranges (all experiments), and variations in opt-out wording that included skipping (Experiment 1), "I don't know" (all experiments), help-seeking, and a combination of the latter two, which approaches the flexibility of real-life communication (Experiment 3). The global opt-out pattern was persistent even when looking at the accumulated data across all experiments.

The persistent flat opt-out slope stands in contrast to the alternative predictions derived from the sunk cost effect, which supports a downward tendency, and the Region of Proximal Learning model, which supports an upward tendency. Thus, the conclusion is that the opt-out criterion functions like a signal-detection confidence threshold, leading to opting out when confidence is lower than this criterion. Unlike the submission confidence criterion, which is determined by both confidence and time, the opt-out criterion is set in a more parsimonious manner, to be based on confidence as the sole parameter, without taking the invested thinking time into account.

Of course, the flat opt-out criterion could also stem from balancing several considerations within and across participants. The individual difference analyses—including those focused on associating the opt-out criterion slope with age and SES in Experiment 2, and the more comprehensive analyses associating the opt-out criterion slope with age, cognitive variables, and metacognitive measures across all the collected data—did not expose any factor that reliably tilts the opt-out criterion upwards or downwards.

The purpose of the comparison between the two task types was to examine whether the opt-out criterion differs in some way between tasks that show diminishing criterion patterns (DCM, Ackerman, 2014) and those for which the fixed criterion fits (Nelson & Narens, 1990; Ratcliff et al., 2016). The results suggest that the two confidence criteria, for submitting and for opting out, are disconnected—each has factors that affect its height and slope. Still, there seems to be variance in the opt-out confidence criterion slope that is yet to be explained, suggested by the weak but significant association between the opt-out criterion slope and the submit criterion slope. That association provides evidence that people do vary in their effort regulation and suggests that the idea of some dependence between the two slopes cannot be ruled out.

The 3SRM addresses a gap in understanding time regulation under uncertainty. It advances prior metacognitive frameworks by revealing that opting out is not merely a by-product of failed problem-solving, but a systematically governed control decision with its own stable confidence threshold. By integrating an opt-out criterion alongside the submit criterion and time limit, the 3SRM clarifies the distinct cognitive signals that drive submission, perseverance, and withdrawal. It thereby captures a broader range of real behaviors than models assuming only confidence-based or time-based stopping. Crucially, the 3SRM explains why people often opt out before reaching their time limit with repeated demonstrations that these decisions occur at a remarkably constant confidence level across tasks, incentives, and individual differences. The model therefore uniquely explains how people counteract the speed–accuracy tradeoff when facing challenging thinking tasks that allow legitimate opting out.

6.2. Motivational Effects on Opt-Out Behavior

Goldsmith (2016) focused on the use of opting out for quality control. His message was that people aim to improve output-bound accuracy (success rate over submitted responses) relative to input-bound accuracy (success rate over all items) by deploying their control sensitivity to opt out of low-confidence answers. Importantly, this finding should not be taken for granted, as a recent study with perceptual decision-making found no such improvement (Al Dowaji et al., 2025). The present study reinforces and generalizes the conclusion that people are motivated toward accuracy, as reflected by the overall high control sensitivity regardless of mean confidence being around 70-75% with multiple-choice analogies (Experiment 1 and Experiment 2) or around 55% with open-ended knowledge questions (Experiment 3). This conclusion is further supported by substantially lower success rate in opted-out responses than in submitted ones in all experiments (Table 1, Table 2, and Table 4). However, if participants aim only for accuracy, they should avoid opting out altogether, in particular in multiple-choice formats, as in Experiment 1 and Experiment 2, because opting out waives any opportunity of coming up with the correct answer just by guessing. When answering knowledge questions calling for numerical answers, as in Experiment 3, aiming solely for accuracy would lead under uncertainty to very coarse answers that are undeniably correct but also meaningless (e.g., "Obama was elected during the last century"). However, people do prefer to opt out over providing ridiculously uninformative answers (Ackerman &

Goldsmith, 2008).

The present study adds the consideration of time. On the one hand, if success were the sole motivation, people would invest time until their confidence reached the level they would require for quick responses, in line with the single stopping rule models. People clearly do not do this (Ackerman & Morsanyi, 2023). On the other hand, if participants were guided solely by a desire to save time, in line with the principle of cognitive miserliness (Stanovich, 2018), they would wildly guess or use opting out opportunities to save time and effort. Rather, participants seem to strategically balance these two motivations.

By the DCM (Ackerman, 2014), while the confidence criterion reflects the aim of answering correctly, the time limit reflects the concurrent aim of not wasting time. Experiment 2 of the present study demonstrated strategic adjustment of these aims in line with the incentive structure. Motivating people to work efficiently, even if this would entitle them to only 20p, led participants to counteract the speed–accuracy tradeoff: they worked faster with no compromise on success or confidence, with equivalent opting out rates. Optimistically, this finding reveals that people are able to effectively cut time investment with no compromise on their confidence target. Pessimistically, this finding indicates that by default people waste time, despite the general aim to work quickly in addition to provide answers that are correct. When losses were the only option, participants doubled their opt-out rates, but time management still remained as in the balanced condition (control group), with no change in the confidence with which they opted out and the association of confidence with time (Table 4, Figure 5).

Why do people work hard to strategically balance success with time? The monetary incentives in all current experiments were small (up to 60 pence) but could lead to a 25–30% rise relative to the baseline compensation. As hinted above, people may be guided by multiple other motivations, even when the incentive structure would suggest otherwise, and despite the task potentially taking more time without proportional compensation. First, Ackerman and Levontin (2024) used priming by reading half a page of text to make people believe that facing challenging tasks does or does not develop a person's intelligence. This subtle manipulation led participants to invest more time in items where more thinking offered a chance for improvement, but not in items with a lower chance for improvement. This finding hints at general cognitive development as an underlying motivation that is combined with avoiding time wastage. Second, in online platforms there are global utilitarian motivations, such as reputation within the platform (e.g., an approval rate that affects invitations for future tasks). This reputation motivation might lead people toward assumed demand characteristics, focusing on success as the inferred goal based on situational norms, even when this stands at odds with the immediate incentive structure, as was the case with the loss condition in Experiment 2 here (see Corneille & Lush, 2023, for a review). Moreover, Prolific participants have been found to display attributes that support a motivation toward success: intrinsic honesty, focused attention regardless of involved losses, curiosity, and high need for cognition (Albert & Smilek, 2023; Peer et al., 2022). Future research is called for to consider the balance between explicit and implicit demands.

The motivation to avoid risky choices has been extensively studied, with loss aversion being a hallmark of those findings (Mrkva et al., 2020). In the present study, all groups could potentially lose points, even the control groups, who had to submit every answer. A motivational dominance of loss aversion could, theoretically, lead participants in all groups to opt out of all items and waive the promised bonus. However, although participants clearly took potential losses into account, they were far from exploiting the potential of opting out. Thus, although there are signs of loss aversion in the present findings, participants clearly take into account other motivations.

Interestingly, the "Skip" framing for opting out did not show an effect relative to "I don't know" (Experiment 1), but help requests (17%) were used less than admitting ignorance using "I don't know" (40%; Experiment 3). The straightforward explanation is that in our experimental design, help requests were costly (1 point) and time-consuming. However, beyond that, unlike skipping using offered help as in the design of Experiment 3 might be associated with social implications (e.g., a desire to maintain a positive self-image, Chen & Son, 2024; Harari et al., 2022; Holtgraves et al., 1997). While some prior research did consider social effects on metacognitive processes (see Ackerman, Bernstein, et al., 2020; Allwood et al., 2016; Mata, 2020; Sidi et al., 2018), studies have only rarely examined opting out while taking social motivations into account (e.g., norms of communication, Grice, 1975), as done with the aim to be informative when communicating with others (Ackerman & Goldsmith, 2008; Yaniv & Foster, 1997).

All these motivations are highly relevant for real-life situations, including education and work contexts, in which the aims of accuracy and time constantly combine to demand efficiency, and beliefs about skill improvement, task design, losses, and social considerations are at play (Bureau et al., 2022; Gagné & Hewett, 2025). Future research is called for to consider how time investment is affected by a willingness to admit ignorance given this variety of underlying, and potentially contradictory, motivations.

6.3. Additional Future Directions

This study clearly represents an initial step in understanding effort regulation in the presence of explicit opt-out opportunities. The fine-tuned examination reported here introduces many questions for future research not raised so far. First, how can people effectively decide to opt out quickly? As explained above, memory research has shown that people initially assess their ability to answer the question before embarking on a thorough search for the answer (Glucksberg & McCloskey, 1981; Singer & Tiede, 2008). In problem-solving too, as Figure 1 demonstrates, rating confidence as low within the first 15 seconds (i.e., the very first intermediate confidence rating) was highly predictive that the solving process would be long, and that solvers would nonetheless end up with little trust in the correctness of the best answer they could provide. The two-response paradigm of reasoning research requires participants to provide the initial solution that comes to mind and rate their Feeling of Rightness judgments before engaging in comprehensive solving attempts (Thompson et al., 2011). This entire body of research is focused on providing the initial solution that comes to mind vs. triggering deliberative thinking (Bago & De Neys, 2017; Thompson et al., 2011; Thompson et al., 2013), while overlooking the question of whether the same judgment would also guide opting out. A small body of prior meta-reasoning research (Ackerman & Thompson, 2017) has considered judgments of solvability after only a glance at problems that include unsolvable ones (e.g., Ackerman & Beller, 2017; Burton et al., 2023; Glucksberg & McCloskey, 1981; Topolinski & Strack, 2009). By the meta-reasoning theory, a low judgment of solvability should trigger quick opting out. Indeed, these brief initial judgments of solvability were found to be predictive of later solving attempts (Topolinski & Strack, 2009), including longer time invested in solving attempts for problems that were judged to be solvable compared with those judged to be unsolvable, regardless of their actual solvability (Lauterman & Ackerman, 2019, 2024).

There is a lot of room for further research into factors that affect initial judgments and their association with regulatory processes. A promising direction is to combine Judgments of Solvability provided before any answer comes to mind and Feeling of Rightness regarding the first considered solution with opt-out options. Such research might also delve into heuristic cues that underlie these early judgments, including those already identified for one but not for the other, such as coherence, nameability, background information, and fluency (Lauterman & Ackerman, 2024; Thompson et al., 2013; Topolinski & Strack, 2008).

Second, one may ask whether confidence starts to accumulate from zero. The starting point is not well-defined as people must understand the requirements and get the information provided before they can address a thinking challenge. All studies that have collected judgments of solvability and those that apply the two-response paradigm with Feeling of Rightness judgments show variability in these early metacognitive judgments, rather than all judgments starting at very low confidence. It is possible that ongoing process tracking based on pupil dilation and mouse tracking (Dhengre et al., in press; Purcell et al., 2023; Travers et al., 2016) can provide sight-holes into early judgments without explicitly soliciting those judgments. If this approach proves effective, studying the temporal dynamics that precede opting out and inferring confidence just before the decision is made would represent a good use of those methods.

7. Summary and Conclusion

This study analyzed opt-out behavior in terms of monitoring confidence as a guide for the allocation of thinking time while offering researchers tools that support considering factors that may affect effort regulation. In particular, the addition of the opt-out criterion to prior theoretical models delineates the time course that guides opting-out decisions. Moreover, the present study joins prior calls for attention to opting out as a legitimate response option available in educational, personal, and work scenarios, and carrying the potential to save time that would otherwise be invested in vain (see Sidi & Ackerman, 2024, for a review). Overall, the present findings provide new insights into metacognitive monitoring and control and open doors for future research into factors affecting time management in terms of theory and training toward improving thinking efficiency when performing tasks under tight time frames.

References

Ackerman, R. (2014). The Diminishing Criterion Model for metacognitive regulation of time investment. *Journal of Experimental Psychology: General*, *143*(3), 1349-1368.

Ackerman, R. (2023). Bird's-Eye View of Cue Integration: Exposing instructional and task design factors which bias problem solvers. *Educational Psychology Review*, *35*(2), 55.

Ackerman, R., & Beller, Y. (2017). Shared and distinct cue utilization for metacognitive judgments during reasoning and memorization. *Thinking & Reasoning*, *23*(4), 376-408.

Ackerman, R., Bernstein, D. M., & Kumar, R. (2020). Metacognitive hindsight bias. *Memory & cognition*, *48*(5), 731-744.

Ackerman, R., Binah-Polak, A., & Lauterman, T. (2023). Metacognitive effort regulation across cultures. *Journal of Intelligence*, *11*(9), 171.

Ackerman, R., Gal, A., Sagi, T., & Shraga, R. (2019). A cognitive model of human bias in matching. *Pacific Rim International Conference on Artificial Intelligence (PRICAI)*.

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting--With and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(5), 1224-1245.

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, *17*(1), 18-32.

Ackerman, R., & Levontin, L. (2024). Mindset effects on the regulation of thinking time in problem-solving. *Thinking & Reasoning*, *30*(3), 479-508.

Ackerman, R., & Morsanyi, K. (2023). We know what stops you from thinking forever: A metacognitive perspective. *Behavioral & Brain Sciences*, *46*.

Ackerman, R., & Thompson, V. A. (2017). Meta-Reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, *21*(8), 607-617.

Ackerman, R., Yom-Tov, E., & Torgovitsky, I. (2020). Using confidence and consensuality to predict time invested in problem solving and in real-life web searching. *Cognition*, *199*, 104248.

Al Dowaji, R., Xu, J., Jin, Y., Porte, A., & Lauwereyns, J. (2025). The limits of metacognitive control during perceptual decision-making: opting out without improving accuracy. *Frontiers in Psychology*, *16*, 1551665.

Albert, D. A., & Smilek, D. (2023). Comparing attentional disengagement between Prolific and MTurk samples. *Scientific reports*, *13*(1), 20574.

Albert, S. M., & Duffy, J. (2012). Differences in risk aversion between young and older adults. *Neuroscience and neuroeconomics*, 3-9.

Allwood, C. M., Karlsson, B. S., & Buratti, S. (2016). Does consulting with others affect answerability judgments of difficult questions? *Social Influence*, *11*(1), 40-53.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124-140.

Bae, J., Hong, S.-s., & Son, L. K. (2021). Prior failures, laboring in vain, and knowing when to give up: Incremental versus entity theories. *Metacognition and Learning*, *16*(2), 275-296.

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual review of psychology*, *64*, 417-444.

Bottemanne, L., & Dreher, J.-C. (2019). Vicarious rewards modulate the drift rate of evidence accumulation from the drift diffusion model. *Frontiers in Behavioral Neuroscience*, *13*, 142.

Bureau, J. S., Howard, J. L., Chong, J. X., & Guay, F. (2022). Pathways to student motivation: A meta-analysis of antecedents of autonomous and controlled motivations. *Review of Educational Research*, *92*(1), 46-72.

Burton, O. R., Bodner, G. E., Williamson, P., & Arnold, M. M. (2023). How accurate and predictive are judgments of solvability? Explorations in a two-phase anagram solving paradigm. *Metacognition and Learning*, *18*(1), 1-35.

Calder-Travis, J., Bogacz, R., & Yeung, N. (2023). Expressions for Bayesian confidence of drift diffusion observers in fluctuating stimuli tasks. *Journal of Mathematical Psychology*, *117*, 102815.

Carsten, T., Hoofs, V., Boehler, C. N., & Krebs, R. M. (2019). Are losses more effective than rewards in improving performance in a cognitive task? *Motivation Science*.

Chen, S., & Son, L. K. (2024). High impostors are more hesitant to ask for help. *Behavioral Sciences*, *14*(9), 810.

Cohen, D., Shavit, Y., & Teodorescu, K. (2024). Don't Give-Up: Why some intervention schemes encourage suboptimal behavior. *Psychonomic Bulletin & Review*, 1-10.

Corneille, O., & Lush, P. (2023). Sixty years after Orne's American psychologist article: A conceptual framework for subjective experiences elicited by demand characteristics. *Personality and Social Psychology Review*, *27*(1), 83-101.

Dhengre, S., Ackerman, R., & Rothrock, L. (in press). Eye-tracking as a lens into meta-reasoning. *Thinking & Reasoning*.

Donkin, C., Little, D. R., & Houpt, J. W. (2014). Assessing the speed-accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1183.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271-280.

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological inquiry*, *6*(4), 267-285.

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, *14*(3), 481-496.

Eisenhauer, J. G. (2021). Meta-analysis and mega-analysis: A simple introduction. *Teaching Statistics*, *43*(1), 21-27.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160.

Fiedler, K., Ackerman, R., & Scarampi, C. (2019). Metacognition: monitoring and controlling one's own knowledge, reasoning and decisions. In R. J. Sternberg & J. Funke (Eds.), *The Psychology of Human Thought: An Introduction* (pp. 89-111). Heidelberg University Publishing.

Funke, J. (2010). Complex problem solving: a case for complex cognition? *Cognitive processing*, *11*(2), 133-142.

Gächter, S., Johnson, E. J., & Herrmann, A. (2022). Individual-level loss aversion in riskless and risky choices. *Theory and Decision*, *92*(3), 599-624.

Gagné, M., & Hewett, R. (2025). Assumptions about human motivation have consequences for practice. *Journal of Management Studies*, *62*(5), 2098-2124.

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(5), 311.

Goldsmith, M. (2016). Metacognitive quality-control processes in memory retrieval and reporting. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 357-385). Oxford University Press.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389.

Grabman, J. H., & Dodson, C. S. (2024). Unskilled, underperforming, or unaware? Testing three accounts of individual differences in metacognitive monitoring. *Cognition*, *242*, 105659.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics* (pp. 41–58). Academic Press.

Guttman, Z. R., Ghahremani, D. G., Pochon, J.-B., Dean, A. C., & London, E. D. (2021). Age influences loss aversion through effects on posterior cingulate cortical thickness. *Frontiers in Neuroscience*, *15*, 673106.
Guzel, M. A., & Higham, P. A. (2013). Dissociating early-and late-selection processes in recall: The mixed blessing of categorized study lists. *Memory & cognition*, *41*(5), 683-697.
Hanczakowski, M., Pasek, T., Zawadzka, K., & Mazzoni, G. (2013). Cue familiarity and 'don't know'responding in episodic memory tasks. *Journal of Memory and Language*, *69*(3), 368-383.
Hansla, A., & Johansson, L.-O. (2020). Risky spending after experienced loss: The moderating effect of socioeconomic background. *Journal of the Association for Consumer Research*, *5*(4), 427-438.
Harari, D., Parke, M. R., & Marr, J. C. (2022). When helping hurts helpers: Anticipatory versus reactive helping, helper's relative status, and recipient self-threat. *Academy of Management Journal*, *65*(6), 1954-1983.
Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, *35*(6), 2476-2484.
Hawkins, G. E., & Heathcote, A. (2021). Racing against the clock: Evidence-based versus time-based decisions. *Psychological review*, *128*(2), 222–263.
Higham, P. A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General*, *136*(1), 1.
Holtgraves, T., Eck, J., & Lasky, B. (1997). Face Management, Question Wording, and Social Desirability 1. *Journal of Applied Social Psychology*, *27*(18), 1650-1671.
Kleitman, S., & Moscrop, T. (2010). Self-confidence and academic achievements in primary-school children: Their relationships and links to parental bonds, intelligence, age, and gender. In A. Efklides & P. Misailidi (Eds.), *Trends and Prospects in Metacognition Research. Part 2* (pp. 293-326). Springer.
Knutson, B., Samanez-Larkin, G. R., & Kuhnen, C. M. (2011). Gain and loss learning differentially contribute to life financial outcomes. *PLoS One*, *6*(9), e24390.
Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological review*, *103*(3), 490-517.
Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, *135*(1), 36-68.
Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology-General*, *131*(2), 147-162.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121-1134.
Lauterman, T., & Ackerman, R. (2019). Initial Judgment of Solvability in non-verbal problems–A predictor of solving processes *Metacognition and Learning*, *14*(3), 365–383.
Lauterman, T., & Ackerman, R. (2024). Initial Judgment of Solvability: Integrating prior expectations with experience-based heuristic cues. *Thinking & Reasoning*, *30*(1), 135-168.
Law, M. K., Stankov, L., & Kleitman, S. (2022). I choose to opt-out of answering: Individual differences in giving up behaviour on cognitive tests. *Journal of Intelligence*, *10*(4), 86.
Law, M. K., Thompson, V. A., Stankov, L., & Kleitman, S. (2025). Systematic adaptive and maladaptive giving-up strategies in cognitive problem-solving. *Thinking & Reasoning*, 1-42.
Lee, D. G., Daunizeau, J., & Pezzulo, G. (2023). Evidence or confidence: What is really monitored during a decision? *Psychonomic Bulletin & Review*, *30*(4), 1360-1379.
Maniscalco, B., Charles, L., & Peters, M. A. (2024). Optimal metacognitive decision strategies in signal detection theory. *Psychonomic Bulletin & Review*, 1-29.

Masís, J., Chapman, T., Rhee, J. Y., Cox, D. D., & Saxe, A. M. (2023). Strategically managing learning during perceptual decision making. *Elife*, *12*, e64978.
Massar, S. A., Pu, Z., Chen, C., & Chee, M. W. (2020). Losses motivate cognitive effort more than gains in effort-based decision making and performance. *Frontiers in human neuroscience*, *14*, 544976.
Mata, A. (2020). Conflict detection and social perception: Bringing meta-reasoning and social cognition together. *Thinking & Reasoning*, *26*(1), 140-149.
Mazor, M., Maimon-Mor, R. O., Charles, L., & Fleming, S. M. (2023). Paradoxical evidence weighting in confidence judgments for detection and discrimination. *Attention, Perception, & Psychophysics*, *85*(7), 2356-2385.
Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463-477.
Metcalfe, J., & Wiebe, D. (1987). Metacognition in insight and noninsight problem solving. *Memory & cognition*, *15*, 238-246.
Mikels, J. A., & Reed, A. E. (2009). Monetary losses do not loom large in later life: Age differences in the framing effect. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *64*(4), 457-460.
Mrkva, K., Johnson, E. J., Gächter, S., & Herrmann, A. (2020). Moderating loss aversion: Loss aversion has moderators, but reports of its death are greatly exaggerated. *Journal of Consumer Psychology*, *30*(3), 407-428.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-173). Academic Press.
O'Shea, B. A., & Ueda, M. (2021). Who is more likely to ignore experts' advice related to COVID-19? *Preventive medicine reports*, *23*, 101470.
Olschewski, S., Mullett, T. L., & Stewart, N. (2025). Optimal allocation of time in risky choices under opportunity costs. *Cognitive Psychology*, *157*, 101716.
Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237-241.
Payne, S. J., & Duggan, G. B. (2011). Giving up problem solving. *Memory & cognition*, *39*(5), 902-913.
Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662.
Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review*, 1-11.
Purcell, Z. A., Wastell, C. A., & Sweller, N. (2023). Eye movements reveal that low confidence precedes deliberation. *Quarterly Journal of Experimental Psychology*, *76*(7), 1539-1546.
Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260-281.
Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, *181*(4099), 574-576.
Reynolds, A., Garton, R., Kvam, P., Sauer, J., Osth, A. F., & Heathcote, A. (2021). A dynamic model of deciding not to choose. *Journal of Experimental Psychology: General*, *150*(1), 42.
Rouy, M., de Gardelle, V., Reyes, G., Sackur, J., Vergnaud, J. C., Filevich, E., & Faivre, N. (2022). Metacognitive improvement: Disentangling adaptive training from experimental confounds. *Journal of Experimental Psychology: General*, *151*(9), 2083.
Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying “don't know” responses enhances interview quality. *Journal of Experimental Psychology: Applied*, *19*(1), 72.
Shapira, A. A., & Pansky, A. (2019). Cognitive and metacognitive determinants of eyewitness memory accuracy over time. *Metacognition and Learning*, *14*, 437-461.
Sidi, Y., & Ackerman, R. (2024). Opting out as an untapped resource in instructional design: Review and implications. *Educational Psychology Review*, *36*(2), 1-31.

Sidi, Y., Ackerman, R., & Erez, A. (2018). Feeling happy and (over) confident: the role of positive affect in metacognitive processes. *Cognition and Emotion*, *32*(4), 876-884.
Singer, M., & Tiede, H. L. (2008). Feeling of knowing and duration of unsuccessful memory search. *Memory & cognition*, *36*(3), 588-597.
Smith, P. L., & Ratcliff, R. (2022). Modeling evidence accumulation decision processes using integral equations: Urgency-gating and collapsing boundaries. *Psychological review*, *129*(2), 235–267.
Soman, D. (2001). The mental accounting of sunk time costs: Why time is not like money. *Journal of Behavioral Decision Making*, *14*(3), 169-185.
Son, L. K., & Sethi, R. (2010). Adaptive learning and the allocation of time. *Adaptive Behavior*, *18*(2), 132-140.
Stankov, L., Kleitman, S., & Jackson, S. A. (2014). Measures of the trait of confidence. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (pp. 158-189). Academic Press.
Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, *24*(4), 423-444.
Strudwicke, H. W., Bodner, G. E., Williamson, P., & Arnold, M. M. (2023). Open-minded and reflective thinking predicts reasoning and meta-reasoning: evidence from a ratio-bias conflict task. *Thinking & Reasoning*, 1-27.
Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107-140.
Thompson, V. A., Prowse Turner, J. A., Pennycook, G., Ball, L., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, *128*, 237-251.
Topolinski, S., & Strack, F. (2008). Where there's a will—there's no intuition. The unintentional basis of semantic coherence judgments. *Journal of Memory and Language*, *58*(4), 1032-1048.
Topolinski, S., & Strack, F. (2009). The analysis of intuition: Processing fluency and affect in judgements of semantic coherence. *Cognition and Emotion*, *23*(8), 1465-1503.
Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, *150*, 109-118.
Tsui, A. B. (1991). The pragmatic functions of I don't know. *Text-Interdisciplinary Journal for the Study of Discourse*, *11*(4), 607-622.
Undorf, M., & Ackerman, R. (2017). The puzzle of study time allocation for the most challenging items. *Psychonomic Bulletin & Review*, *24*(6), 2003-2011.
Undorf, M., & Erdfelder, E. (2011). Judgments of learning reflect encoding fluency: Conclusive evidence for the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1264.
Undorf, M., Livneh, I., & Ackerman, R. (2021). Metacognitive control processes in question answering: help seeking and withholding answers. *Metacognition and Learning*, *16*(2), 431-458.
Vuorre, M., & Metcalfe, J. (2022). Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacognition and Learning*, *17*(2), 269-291.
Walker, A., Turpin, M. H., Fugelsang, J., & Koehler, D. (2019). Intuition speed as a predictor of choice and confidence in point spread predictions. *Judgment and Decision Making*, *14*(2), 148-155.
Weissberger, G. H., Han, S. D., Yu, L., Barnes, L. L., Lamar, M., Bennett, D. A., & Boyle, P. A. (2022). Subjective socioeconomic status is associated with risk aversion in a community-based cohort of older adults without dementia. *Frontiers in Psychology*, *13*, 963418.
White, S. F., Nusslock, R., & Miller, G. E. (2022). Low socioeconomic status is associated with a greater neural response to both rewards and losses. *Journal of cognitive neuroscience*, *34*(10), 1939-1951.
Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 201.

Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgmental estimation. *Journal of Behavioral Decision Making*, *10*(1), 21-32.
Yechiam, E., & Hochman, G. (2013). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, *66*(2), 212-231.
Yechiam, E., & Zeif, D. (2023). Revisiting the effect of incentivization on cognitive reflection: A meta-analysis. *Journal of Behavioral Decision Making*, *36*(1), e2286.