

# Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis

Citation for published version (APA):

Cole, S. N., Smith, D. M., Ragan, K., Suurmond, R., & Armitage, C. J. (2021). Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis. *Psychonomic Bulletin & Review*, 28(5), 1514–1537. <https://doi.org/10.3758/s13423-021-01880-6>

**Document status and date:**

Published: 01/10/2021

**DOI:**

[10.3758/s13423-021-01880-6](https://doi.org/10.3758/s13423-021-01880-6)

**Document Version:**

Accepted author manuscript (Peer reviewed / editorial board version)

**Document license:**

Unspecified

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## **Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis**

**Cole, S. N.<sup>1</sup>, Smith, D. M.<sup>2,3</sup>, Ragan, K.<sup>4</sup>, Suurmond, R.<sup>5</sup> & Armitage, C. J.<sup>3</sup>**

<sup>1</sup>York St John University, <sup>2</sup>Leeds Trinity University, <sup>3</sup>University of Manchester, <sup>4</sup>Newcastle University, <sup>5</sup>Maastricht University.

Accepted author version. To cite, please use the following :

Cole, S. N., Smith, D. M., Ragan, K., Suurmond, R. & Armitage, C. J. (2020). Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis. *Psychonomic Bulletin & Review*. *Accepted for publication*

This article is published by *Psychonomic Bulletin & Review* by Psychonomic Society (<https://www.psychonomic.org/>), and is the official record of this article. This author version of the article is to be used for academic or personal purposes, and must not be edited or sold.

Open Science statement:

The data and analysis scripts are available at Open Science Framework: DIO:  
10.17605/OSF.IO/VQZB7

Link: [https://osf.io/vqzb7/?view\\_only=eea2cb1be9ba4bbea0212a3241918ae9](https://osf.io/vqzb7/?view_only=eea2cb1be9ba4bbea0212a3241918ae9)

### **Acknowledgments**

CJA was supported by NIHR Manchester Biomedical Research Centre and the NIHR Greater Manchester Patient Safety Translational Research Centre. SNC was supported by a Postdoctoral Fellowship at the Centre on Autobiographical Memory Research, Aarhus University. We also appreciate help in locating articles provided by Andrew Smith.

## **Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis**

### Abstract

Mental simulation of future scenarios is hypothesized to affect future behavior, but a large and inconsistent literature means it is unclear whether, and under what conditions, mental simulation can change people's behavior. A meta-analysis was conducted to synthesize the effects of mental simulation on behavior and examine under what conditions mental simulation works best. An inclusive systematic database search identified 123 ( $N = 5685$ ) experiments comparing mental simulation to a control group. After applying a multilevel random effects model, a statistically-reliable positive effect of Hedges'  $g=0.49$  [95% CI 0.37; 0.62], which was significantly different than zero. Using a taxonomy to identify different subtypes of mental simulation (along two dimensions, class [process, performance, outcome] and purpose [whether an inferior, standard, superior version of that behavior is simulated]), it was found that superior simulations garnered more reliable beneficial effects than inferior simulations. These findings have implications for integrating theories of how mental simulations change behavior, how mental simulations are classified, and may help guide professionals seeking evidence-based and cost-effective methods of changing behavior.

**Keywords:** mental simulation; mental practice; behaviour change; process simulations; outcome simulations

## **Synthesizing the Effects of Mental Simulation on Behavior Change: Systematic Review and Multilevel Meta-Analysis**

Imagining situations in one's personal future is a common occurrence in humans (D'Argembeau, Renaud, & Van der Linden, 2011) and recent experiments show that mental simulation can change a range of behaviors, including increasing fruit and vegetable consumption (Knäuper, McCollam, Rosen-Brown, Lacaille et al., 2011), improving accuracy on a pointing task (LaCourse, Turner, Randolph-Orr, Schandler, & Cohen, 2004), and increasing speed in a car-racing task (Callow, Roberts, Hardy, Jiang & Edwards, 2013, Exp. 1). Although there have been several attempts to review the literature on the effects of mental simulation on behavior change (e.g., Corbin, 1972; Driskell et al., 1994, Feltz & Landers, 1983; Richardson, 1967) these have typically been limited to unsystematic narrative reviews of the relatively narrow domains of 'mental practice' or 'mental rehearsal' that rely on prior experience of the target behavior, or have confounded mental simulation with other behavior change techniques. The aims of the present systematic review and meta-analysis are to address these limitations and discover: (a) the unique effects of mental simulation and whether these effects are robust enough to hold across a range of domains and behavior types, and (b) under what circumstances mental simulation works best. In this way, this research harnesses the meta-analytic approach to examine the cross-disciplinary nature of a specific psychological phenomenon.

We believe there are four main reasons why this meta-analysis may have particular significance for behavior change research, and related fields (e.g., sports science, cognitive psychology) within the current literature.

First, within the behavior change literature, mental simulation has yet to be fully appreciated as a key behavior change technique in its own right, despite continued empirical research (e.g., Oh & Larose, 2015) and a recent systematic review and meta-analysis focusing on health interventions (see Conroy & Hagger, 2018). It may have received little attention

because, despite some positive findings (e.g., Hagger, Lonsdale & Chatzisarantis, 2011), well-designed experiments in the field have resulted in non-significant findings (e.g., Conroy, Sparks & de Visser, 2015). This could lead to uncertainty among researchers as to the reliability and generality of the effects, perhaps due to the way in which mental simulation is typically conceptualized. For instance, in a taxonomy of 93 behavior change techniques developed by expert agreement, mental simulation was reduced to a sub-component (15.2), labelled ‘mental performance of successful performance’ (see Supplemental Materials of Michie et al., 2013). Although simulation of positive/optimal performance is an important aspect of mental simulation, the broader literature around mental simulation points to a plethora of subtypes that are used across diverse fields (i.e., inferior and superior, see Table 1 below). One aim of the current meta-analysis was to determine whether these subtypes substantially moderate the effects of mental simulation on behavior across varied domains – a question not addressed previously, and certainly not across a diverse range of studies.

The second reason for conducting this systematic review and meta-analysis now, is that several approaches and theories of episodic future thinking – the cognitive ability to mentally place oneself in a future context (see Atance & O’Neill, 2001) - have emerged in recent years. These theories draw upon findings from cognitive and neuroscientific research that are highly relevant to understanding the cognitive processes underlying behavior change resulting from manipulating mental simulation (see Baumeister, Vohs & Oettingen, 2016; Schacter & Addis, 2007; Seligman, Railton, Baumeister & Sripada, 2013). The current review aims to provide initial steps toward integrating more recent theoretical literature with meta-analyzed data, distinguishing fine-grained sub-classifications of mental simulation (see below).

Thirdly, the present meta-analysis improves on previous reviews by adjusting terminology to account for more recent literature adopting mental simulations. For example,

when Driskell and colleagues (1994) published their meta-analysis, most mental simulation research emanated from the motor skill and sports domains. Driskell and colleagues' definition of mental simulation was necessarily specific to these studies (restricted to 'mental practice' / 'mental rehearsal'), which often had highly-scripted motor tasks (e.g., pointing tasks), physically practiced extensively before being mentally simulated. Recently, the term 'mental simulation' has been used to simulate more varied behaviors relating to social behaviors (e.g., cooperation, Meleady et al., 2013) and education (e.g., studying, Pham & Taylor, 1999). Here, we adopted a broader conceptualization, defining mental simulation as mental representations of a behavior that can be hypothetical *or* familiar (practiced), and which can represent the future behavior in different ways, but in the context of the person not physically moving whilst simulating (Taylor, Pham, Rivkin & Armor, 1998). More specifically, we examined simulations of possible future scenarios (cf. 'episodic future thinking', Atance & O'Neill, 2001) – which could be either be based directly upon a previous personal physical experience (i.e., familiar) or mentally simulated based on instructions from an experimenter but with no direct experience of the scenario. This definition enables a meta-analysis such as this greater scope to capture the various ways researchers have studied mental simulation.

Finally, our meta-analysis utilized a sophisticated hierarchical or multilevel approach, which is useful in meta-analysis because, like with single studies (e.g., D'Argembeau, Renaud & Van der Linden, 2011), there may be multiple experimental and comparison groups. In prior studies, this dependency has gone untreated, simply assuming that all effect sizes were independent in these meta-analyses (Driskell et al., 1994). Alternatively, some have dealt with the increase in power when assuming independence by dividing the control group  $N$  by the number of conditions (Toli, Webb & Hardy, 2015). However, this can result in a misleading increase in power (as univariate models count each study as independent).

Thus, by including multilevel data, we model the dependency (Field & Gillett, 2010), ensuring our analyses are appropriate to the structure of the data.

### **Types of Mental Simulation**

In order to overcome limitations of previous mental simulation reviews, we explored the various forms of mental simulation in the literature. We summarize below the possible types and provide a breakdown of these along two dimensions (class and purpose, Table 1). This is a conceptual structure aimed to help understand mental simulation and reduce the risk of confounding conceptually distinct forms of mental simulation.

To further distinguish mental simulations, it is known that they focus either on process, performance or outcome. *Process simulations* involve imagining the procedural steps necessary to achieving one's goal (e.g., "in order to run five miles every day, I will make sure I have the appropriate kit, book out time in my schedule, and jog to work in the mornings ", see Table 1 for a taxonomy of mental simulations). It is hypothesized that process simulations enable people to make concrete plans, developing a volitional action plan toward their goal (Taylor & Pham, 1996). *Performance simulations* are conducted when participants are asked to rehearse mentally or mentally practice a behavior after observing or performing that specific action (see Table 1). For example, Pascual-Leone and colleagues (1995) asked participants to repeatedly simulate a well-learned 5-finger piano sequence to improve precision and accuracy. Thus, performance simulations involve the participant mentally 'running through' a specific task in chronological order, which is then tested behaviorally in a criterion task in identical way to the simulation. *Outcome simulations* involve envisioning a desired outcome (e.g., "I vividly imagine feeling relieved and satisfied having completed a marathon next year"), and are hypothesized to motivate people towards achieving their goal (Pham & Taylor, 1999; Hagger, Lonsdale & Chatzisarantis, 2011; Vasque & Buehler, 2007). Process, performance and outcome simulations can be carried out individually or in any combination

and one of the aims of the present research is to tease apart in which combinations (if any) process, performance and outcome simulations differentially influence behavior. We shall call these different *classes* of mental simulation (see Table 1).

In addition, mental simulations differ on another dimension; how the achievement-level of the simulated behavior is framed, namely; [a] below-average (inferior) behavior, [b] average/unspecified behavior, or [c] above-average (superior) behavior. We called this the *purpose* of the mental simulation. The above-average *purpose* of outcome simulations (*superior outcomes*) can be manipulated through simulations of successfully completed behaviors, often associated with positive emotions and imbued with personal meaning (for similar concepts and interventions, see *approach goals*, Elliot & Harackiewicz, 1996; *possible selves*, Markus & Nurius, 1986; *fantasizing*, Oettingen, 2012). In contrast, *inferior outcome simulations* represent a poor outcome or complete failure to achieve one's goal (for a similar concept, see *feared selves*, Markus & Nurius, 1986), but the latter may garner a motivational 'incentive' as one may want to avoid this possible reality (see *avoidance goals*, Elliot & Harackiewicz, 1996). Each variant may trigger behavior change by highlighting discrepancies between current and possible future selves (Markus & Nurius, 1986). In our taxonomy, we define mental simulations based on what participants are explicitly instructed to imagine.

Additionally, studies considering the purpose of performance simulations have examined positive and negative simulations *of the criterion task itself*, with researchers expecting better behavioral outcomes following superior versus inferior performance simulations (Budney & Woolfolk, 1990; Woolfolk, Murphy, Gottesfeld & Aitken, 1985).

To date, process simulations have been studied in an undifferentiated way. That is, researchers define process simulations as functional or beneficial steps, which make goal



attainment more likely (Pham & Taylor, 1999; Taylor & Pham, 1996). In Pham & Taylor's words (1999), "simulating the steps to reach a goal provides information about the actions needed to attain the goal." (p. 251). Although performance and outcome simulations can be positive or negative, few studies have examined steps which might be ineffective or counterproductive in goal behaviors. Nevertheless, we include superior, standard and inferior process simulations in our classification, creating nine distinct potential subtypes of mental simulation (see Table 1). The aim in developing this taxonomy was to avoid redundancy and create orthogonal dimensions (i.e, class and purpose) (see Hall, Mack, Paivio & Hausenblas, 1995 for a similar approach). In other words, it is reasoned here that any particular class of simulation (e.g., performance), can represent the future behavior in either a positive, neutral or negative way (see Table 1). Although this taxonomy is based on previous categorizations (e.g., positive versus negative outcomes, Marszał-Wiśniewska & Jarczewska-Gerc, 2016), these fine-grained distinctions are novel and have yet not been assessed as moderators in a meta-analysis. Further, we did not label inferior simulations as negative, and superior simulations as positive to avoid possible terminological confusion with affective simulations (see Gamble, Moreau, Tippett, & Addis, 2019a).

Previous reviews have tended to confound these different subtypes of mental simulation (e.g., Driskell et al., 1994; Feltz & Landers, 1983) thereby inappropriately aggregating various distinct categories into one abstract phenomenon. A recent meta-analysis has helped to understand the effects of mental simulation in the domain of health interventions (and specifically, how intervention characteristics, such as follow-up simulations, moderate behavioral effects (Conroy & Hagger, 2018)), but conceptually, simulations were only distinguished as process or outcome simulations (thus not addressing the effects of performance simulations). Finally, a meta-analysis on specificity of future thinking and depression, subdivided mental simulations based on their affective qualities

(Gamble, Moreau, Tippett, & Addis, 2019a). We devised a taxonomy that was similar, yet distinctive, and that incorporated two dimensions (*class* and *purpose*, see Table 1), enabling us to incorporate subtypes from different fields and examining methodological characteristics involved in the effects of mental simulation.

**Table 1:** Possible subtypes of mental simulation

		<u>Classes of mental simulation</u>		
		<i>Process simulations</i>	<i>Performance simulations</i>	<i>Outcome simulations</i>
<u>Purpose</u>	<i>Inferior</i>	Simulating a plan which is ineffective	Simulating a performance which is below average	Simulating an undesired or feared outcome
	<i>Standard</i>	Simulating a standard /average plan	Simulating a standard /average performance	Simulating a standard/average outcome
	<i>Superior</i>	Simulating a plan which is effective	Simulating a performance which is above average	Simulating a desired or wished-for outcome

**Note:** Although there could plausibly be studies involving each subtype, some have yet to be studied. Nevertheless, here we list some examples of the subtypes above from existing studies: Superior Process Simulations: Pham & Taylor (1999); Inferior Process Simulations: No known example in existing literature; Inferior Performance Simulations: Alden et al. (2001); Standard Performance Simulations: Andre & Means (1986); Superior Performance Simulations: Callow et al (2013); Inferior Outcome Simulation: Marszał-Wiśniewska & Jarczewska-Gerc (2016 Exp2); Standard Outcome Simulation: No known example in existing literature; Superior Outcome Simulation: Johannessen et al (2012).

## Methodological Characteristics

### Measurement of Task Performance

Performance on any criterion task can be operationalized in a variety of ways, whether it is the speed with which a task is completed or how effective one is within a given timeframe. More specifically, following simulation, a person may perform a task in a faster time, enact more behaviors over a set time or hit a target with precision. It is plausible that the method of how one measures task performance itself may alter the relation between mental simulation and performance. Herein, these three ways of measurement are labelled *speed*, *frequency* and *accuracy*, respectively. It is well-known across psychology's sub-disciplines that physical practice enhances later performance – consequently one will become *faster* on a specific task. One may predict similar facilitatory effects after mental practice. Mental simulation may increase the *frequency* with which a behavior is enacted by instantiating a 'mental set'; one might physically enact a simulated behavior every time an appropriate opportunity arises, especially when the physical situation shares commonalities with the simulation (for a similar mechanism involving *implementation intentions*, see Gollwitzer, 1999). Epitomized by the phrase 'practice makes perfect', one may expect that simulation increases the ability to enact precise or accurate behaviors. Critically, although previous meta-analytic reviews (Corbin, 1972; Driskell et al., 1994; Feltz & Landers, 1983) examined type of task (cognitive *versus* physical, reactive *versus* self-paced), they did not go further in coding how behaviors are measured. Thus, our understanding of how mental simulation may differentially affect behavior as a function of outcome type is still weak. The current meta-analysis seeks to remedy this by examining its possible moderating effect/s.

### Dosage

One key moderator in medical trials is the frequency and duration of an intervention or its dosage (Higgins & Green, 2011). Establishing the minimal requirements for an effective mental simulation intervention is especially important here, where the cost-effectiveness of an intervention is often paramount (Hagger, Lonsdale & Chatzisarantis, 2011). Two points concerning dosage have emerged: Neither duration or frequency have a clear positive relation with effectiveness of mental simulation. First, long periods of mental simulation seem to reduce the benefits gained, despite an overall positive effect compared with controls (Driskell et al., 1994). The diminishing return of administering mental simulation over a long duration has been explained by a loss of interest (Driskell et al., 1994) and reactive inhibition (Corbin, 1972; see Feltz & Landers, 1983 for similar results and conclusions). Based on these findings, Driskell and colleagues recommended that mental simulation-based interventions be restricted to 20 minutes in duration. In a more recent meta-analysis, Conroy & Hagger (2018) found that, in health interventions, longer durations predicted more positive behavioral outcomes. Plausibly, the motor skill studies included by Driskell et al (1994) contained studies with tens, hundreds, or thousands of simulation trials, leading to diminishing return effects. In contrast, health studies, the focus of Conroy & Hagger (2018), have far fewer simulation trials, on average (e.g., Hagger, Lonsdale & Chatzisarantis, 2011). Due to differences between domains sampled in previous meta-analyses, it was necessary to assess dosage across multiple domains here.

Concerning frequency, Feltz and Landers (1983) and Driskell and colleagues (1994) found no moderating effect of mental simulation upon the direction or size of the behavioral effect. However, Corbin (1972) theorized that there may be an ‘optimal’ amount of mental simulation trials that lead to benefits on later behavior, after which, the beneficial effects decrease. This is reasoned to occur on the same grounds as *duration*, such as loss of interest in the simulation (after excessive numbers of simulation trials). However, theories of habit

formation (for a review, see Carden & Wood et al., 2018) would suggest that repeated mental simulation might strengthen associations and lead to greater behavior change. We conducted comparable analyses of mental simulation duration and frequency in the present, broader meta-analysis. Such analyses will be necessary to infer general principles of mental simulation and provide guidance for future intervention design.

### **Incentive**

Payment, either by cash, course credit or voucher is a common occurrence in psychological research, yet material reward is a behavior change technique in its own right (e.g., Michie et al., 2013) that influences performance on a variety of tasks/behaviors (e.g., Brase, 2009; Giuffrida & Torgenson, 1997). It would be valuable to see whether, for example, persistence with mental simulation tasks are influenced by the levels of material incentive offered for participation. This is important because payment to engage with mental simulation tasks would limit the widespread adoption of mental simulation in the wider world (e.g., in public health interventions).

### **Delay**

Delay (also termed retention interval or follow-up duration) is defined here as the time between the last mental simulation and the assessment of behavior. In their review, Driskell and colleagues (1994) examined 62 experiments and found that the mental simulation effect decreased linearly as the delay between the last mental simulation and outcome measure increased. Specifically, at 14 days the mental simulation effect had reduced by half, and by 21 days it had fallen below the  $d < .10$  typically denoting a small effect size (Cohen, 1988). Conroy & Hagger's (2018) analysis of 24 control-simulation comparisons revealed no moderating effect of delay. It should be acknowledged that Driskell and colleagues did not include studies in domains such as health and social psychology and Conroy and Hagger (2018) only assessed delay within health/social psychology – leaving a

gap in meta-analytic data assessing across domains. We will attempt to reconcile these different findings to the broader grouping of mental simulation studies included here, whilst also including domain in the analysis (see below).

### **Domain and Behavioral Task**

One of the aims of the present review was to assess whether the mental simulation effect is transferable across, or is limited to, particular domains of study. Some authors of meta-analyses suggest that certain subtypes of mental simulation can enhance behavior (*motor planning*, Driskell et al., 1994; *health psychology*, Conroy & Hagger, 2018; *sports*, Feltz & Landers, 1983), whereas other studies in other domains have yet to be assessed within a meta-analysis (e.g., the effect of mental simulation within occupational or social domains). In the present review, we categorized studies in terms of their domain, and examined whether each task had a largely cognitive, physical or mixed (cognitive and physical) loading (i.e., Driskell et al., 1994; Feltz & Landers, 1983). Previous reviews indicated an advantage of mental simulation in more cognitive-based tasks as compared with physical-based tasks (see Driskell et al., 1994; Feltz & Landers, 1983).

### **Experimental Control and Design**

Finally, it is now recognized that it is difficult to obviate positively biasing those randomly assigned to experimental groups in psychological interventions because, in contrast to clinical trials involving medicinal interventions, participants are required to engage in psychological techniques (Boot, Simons, Stothart & Stutts, 2013). This issue may be prevalent in mental simulation research, as participants can reasonably ‘guess’ they are in the intervention group. To reduce such effects, studies may include active control groups which involve a task, such as watching a video of an upcoming task. The present review will assess the extent of this issue by measuring studies with passive or active control groups. This will also address the issue of spontaneous engagement in mental simulation in control groups.

Without active controls, or checks on mental activities of control participants, participants could be engaged in mental simulation, possibly to an equivalent degree to those assigned to intervention groups, thus reducing the potential benefit garnered by intentional mental simulation.

### **Overall Aims**

Due to a confounded and fragmented literature, it was clear that a re-appraisal of the mental simulation literature was warranted. This allowed a range of key moderators to be evaluated, such as outcome type, incentive, delay, domain, and task type (some of which have never been meta-analytically examined). Furthermore, as taxonomies are important in organizing knowledge, we developed a taxonomy of mental simulation subtypes. For the first time, this allows subtypes of mental simulation to be compared across studies in a new conceptual structure, which: (1) better reflects the mental simulation interventions used, and (2) aids a clearer understanding of the possible mechanisms involved in the simulation process and its impact on behavior. It was envisioned that the present review could inform future research in applied (e.g., life coaching, health interventions) and scientific endeavors (e.g., understanding processes through which behaviors are modified): Less effective simulations can be de-emphasized and optimal simulations more thoroughly investigated. This would have a two-pronged effect: mental simulations could become regarded as a collection of behavior change techniques in applied psychology (Michie et al., 2013), and theoretical work into how and why mental simulation affects behavior would be energized (see Baumeister, Vohs & Oettingen, 2016).

Based on previous work, we predicted that mental simulations (regardless of subtype) would exert larger behavioral effects where there is a longer intervention duration, participants were incentivized, and with shorter delay (Driskell et al., 1994; Feltz & Landers, 1983; Michie et al., 2013). We further explored whether subtype of mental simulation from



our taxonomy would moderate effect sizes. Publication bias, and study quality (including manipulations checks) were also examined to assess any systematic biases in the existing peer-reviewed literature. The aims of the present systematic review and meta-analysis are, for the first time, to: (a) synthesize the effects of mental simulation on behavior change across a range of domains, and (b) discover under what conditions mental simulation works best.

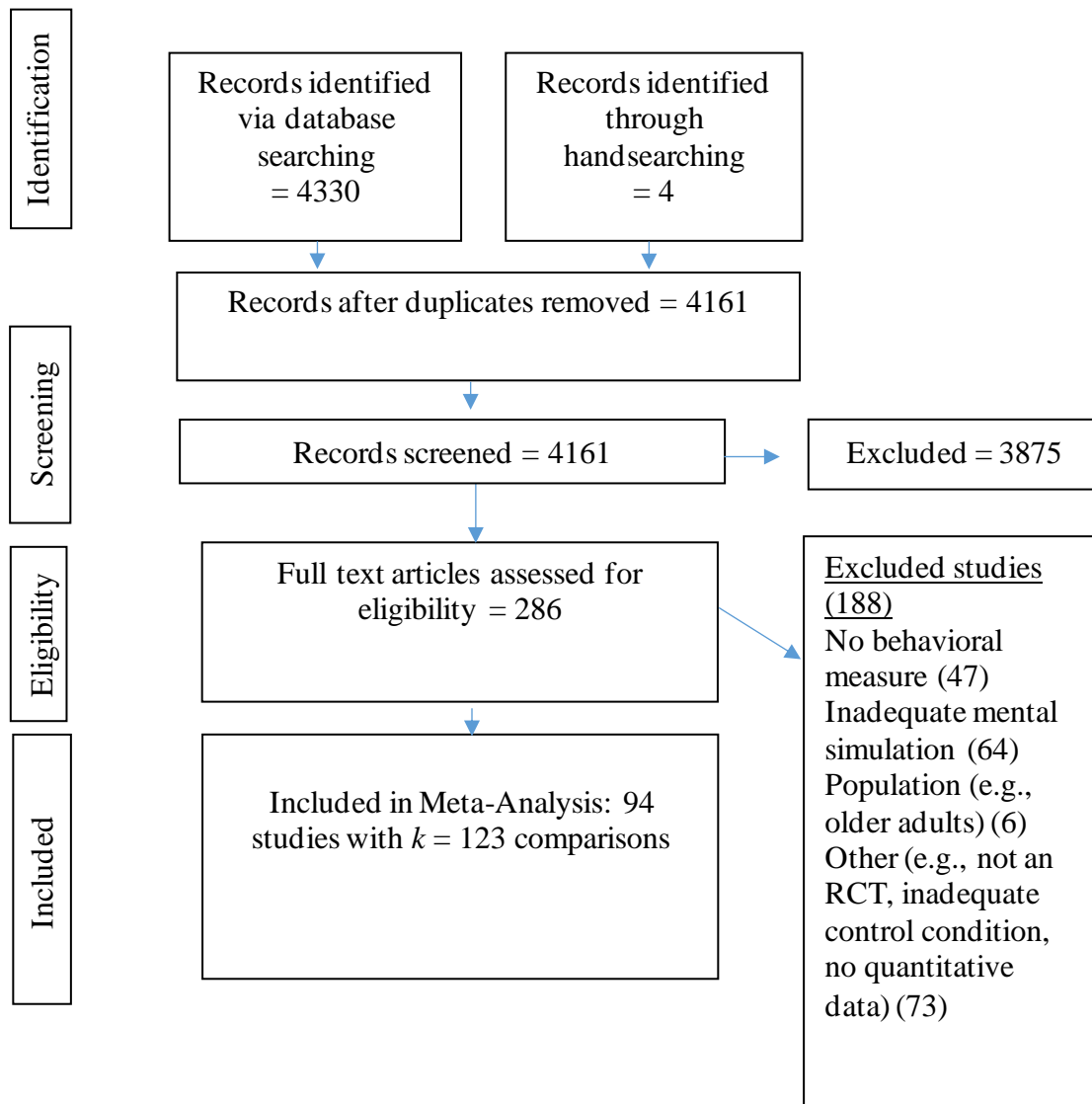
## Method

### Search Strategy and Process

In terms of selection, a systematic search of PubMed (including MEDLINE), PsychINFO and Web of Science (including Arts & Humanities index; Science Citation Index Expanded; Social Science Citation Index) academic databases was conducted of all relevant English-language peer-reviewed articles published before May, 2020. Our definition of mental simulation incorporated its various forms, while maintaining a balance between sensitivity and specificity (see Cochrane Collaboration, 2011). The search terms consisted of descriptors of the intervention and outcome: (mental\* AND rehears\* OR simul\* OR imagin\* OR practi\* OR prepar\* OR possible sel\* OR mental contrasting<sup>1</sup>) AND (behav\* OR perform\*) filtering to include peer-reviewed articles of human participants. To find randomized controlled trials, we included the search terms random\* or experim\*. In total, 4161 articles were identified at the selection phase.

---

<sup>1</sup> One type of intervention relevant here is that of mental contrasting (see Oettingen 2012 for a comprehensive review). This future-thinking intervention involves individuals generating a positive fantasy and then contrasting them with the present barriers to its realization. In doing so, effective plans are devised (so long as certain preconditions are met, such as positive expectations of success, see Oettingen, 2012). The present meta-analysis had an inter-disciplinary focus, and thus we were not able to examine possible mediators and moderators (e.g. expectations, attitudes) from specific domains (health and social psychology). Nevertheless, it was expected that several studies utilizing mental contrasting were necessarily included in this review, and could be categorized within the aforementioned taxonomy, if they fulfilled the inclusion criteria (e.g., behavior as the outcome variable). It will be valuable to assess how this taxonomy maps onto other mental simulation interventions used within the current literature.



**Figure 1.** Flow Diagram of Systematic Search Results

In the eligibility phase, abstracts were screened based on the following four criteria: Included experiments had to: (1) have a between-group experimental design to which participants had been randomly allocated. Hence, only the most rigorous experiments (randomized controlled trials) were included, which provided a more stringent criteria than previous reviews (Feltz & Lander, 1983; Conroy & Hagger, 2018); (2) a control condition (passive or active) that only differed from the experimental condition through absence of mental simulation (e.g., mental simulation only versus passive control group; mental simulation plus feedback versus feedback only control); (3) an outcome measure that

incorporated measurable behavior or performance (self-reported or objective); and (4) include a mental simulation intervention that was related to the main outcome measure. Studies were excluded from the review if they involved; counterfactuals (past simulations) or memories, computer simulations, children (<18 yrs) because mental simulation ability changes through childhood (Skoura, Vinter & Papaxanthis, 2009), older adults (>64 yrs) because age can affect mental simulation capacity (Zapparoli, Invernizzi, Gandola, Verardi, Berlinger et al., 2013) or clinical groups (as the present review is concerned with ‘normal’ cognitive-behavioral processes, and clinical disorders affect mental imagery; see Jeannerod (1994) for a review). As we were interested in randomized controlled trials and because conference proceedings can introduce repeated data from various sources (and more dependency in the data), we decided to only include studies from peer-reviewed articles, and test for publication bias thereafter. Specifying behavior as the main outcome measure as an inclusion criterion meant that experiments assessing outcomes only (e.g., body weight) were excluded unless they also measured the component behaviors (e.g., physical activity, avoiding snacks). In summary, our search strategy aimed to capture experimental studies from a variety of domains. In the abstract screening phase, the first author included or excluded experiments for the next stage using title and abstracts only. A subset of 10% of these were also categorized by the second author to check for reliability. This resulted in moderate to good agreement, as determined by Cohen’s kappa (85% agreement, Cohen’s kappa  $\alpha = .58$ ). All inconsistencies were resolved through discussion between the first and second author and then if a conflict remained, the final decision was made by a third co-author.

Thereafter, the full text of each experiment was examined by one co-author to establish all relevant experiments that fit the review’s criteria (some included experiments which were derived from the same article, see Figure 2). At this stage, the most frequent reason for exclusion was either the mental simulation manipulation not being tested uniquely

(e.g., the mental simulation condition involved other additional components such as implementation intentions, e.g., Stadler et al., 2010) or an absence of a behavioral measure (e.g., body weight was measured, Lukaszewski & Jarczewska-Gerc, 2012, experiment 5 & 6). Although fulfilling the criteria of this review, one experiment did not have sufficient quantitative data for inclusion in the meta-analysis (Bachman, 1990; an attempt was made to contact the author for further information). Finally, all co-authors agreed on the 94 studies (123 effect sizes) that were included in the following quantitative analyses, and the methods of this meta-analysis conformed to the PRISMA guidelines (see *Supplementary Information*). We also conducted manual searches of the reference lists of all included articles retrieved from these searches to identify additional articles. This led to four additional articles being identified (a summary of search results is given in Figure 1).

### **Data Extraction**

In this phase, one co-author (all trained on the coding extraction and mental simulation taxonomy) extracted data from included experiments included in the full text phase. To ensure high standards of data reporting and thoroughness, data extraction was based on guidelines and a proforma by the Cochrane Collaboration (see Higgins & Green, 2011), covering *N* size, sample characteristics, experimental conditions, outcome variables, quality (i.e., manipulation check of mental simulation and control), delivery method, and control condition. Some supplementary categories were added based on the moderator variables described above (e.g., incentive, delay between last mental simulation and behavior).

### **Coding the Subtype of Mental Simulation**

As per the taxonomy specified in the Introduction, in which we identified nine distinct classes of mental simulation, varying along two dimensions; *class* (process/performance/outcome) and *purpose* (i.e., inferior/standard/superior), each

experiment was categorized into one of these subtypes. The first author coded all studies, and a subset (at least 10%) were agreed with a second coder by consensus. There were no disagreements in the reviewed subset.

### **Coding of the Outcome Variable**

We also coded the outcome variable based on the assumption that mental simulation could affect behavior three main ways, *speed* (e.g., reaction time on a pointing task), *accuracy* (an all-or-nothing variable involving someone meeting a target or not, e.g., passing an exam) or *frequency* (the amount of times a behavior is recorded, e.g., amount of times a target is hit, amount of fruit consumed). For a full list of all included studies and their study characteristics, see *Supplementary Materials*. To ensure consistency across co-authors, a standardized form was used for all data extraction. Any disagreements between co-authors were resolved through discussion with the senior co-author (CA) where necessary.

## **Study and Methodological Characteristics**

### **Coding of Domains**

We categorized the studies into the following domains: *health*, *social*, *occupational*, *sports*, *pain* and *motor learning*. For the purposes of the present review, studies examining the mental simulation effect in *educational* contexts (e.g., Pham & Taylor, 1999) were classified as occupational due to the occupational requirement of students to learn.

### **Coding Type of Task**

As in Driskell and colleagues (1994), we assessed the quality of the task employed in each study to assess whether mental simulation was more effective for tasks with either mainly cognitive or mainly physical attributes, or both. Cognitive tasks were defined as those that required mental operations, perceptual input or decisions. Physical tasks were those that contained strength, endurance or coordination. Mixed tasks were those that required both cognitive and physical components (definitions in line with Driskell et al., 1994).

### **Coding of Dosage**

Several metrics were used to determine dosage. As duration of mental simulation was recorded in minutes, studies in which the duration of mental simulation was estimated as ‘short’ but not further specified were assigned a value of *1 minute* except studies in which the duration of mental simulation was calculable. All mental contrasting interventions were given 5 minute duration by default (based on studies in which duration *was* stipulated, e.g., Marszał-Wiśniewska & Jarczewska-Gerc, 2016 Exp2), unless otherwise stated. Duration was coded into three categories: Short=1 (1-5 mins), Medium = 2 (6-20 mins), Long =3 (21+). We also assessed the effect of frequency of mental simulation. Finally, we also multiplied length of (a single) mental simulation (e.g., 2 min) by the total frequency of mental simulations in the intervention (e.g., 5, thus total duration = 2 x 5 = 10).

### **Methodological Quality**

First, we coded whether mental simulation was assessed in the control condition because without this measure, it is difficult to know whether, or the extent to which, participants spontaneously adopted mental simulation, minimizing the effect of any experimental manipulation (for example, after practicing a task, physically, one might spontaneously engage in mental simulation). Secondly, some studies ensured control participants were prevented from using mental strategies (for example, they were asked to carry out mental calculations or read a section of text). Thus, the subtype of control condition was also coded into two categories; either (a) a wait/passive control or (b) an active control that required considerable attention (e.g., verbal tasks, mental arithmetic, writing tasks). Thirdly, studies were then coded for whether mental simulation was assessed in the experimental condition, to assess whether compliance was verifiable (e.g., reporting the contents of the simulation, see, Meslot et al., 2016; or measuring the subjective vividness of the imagery, see Kornspan et al., 2004).

Potential bias was assessed based on criteria from the Cochrane Collaboration (Higgins & Green, 2011) and Chambers (1990). Whether the methods incorporated: (a) participant blinding, (b) experimenter blinding, (c) true randomization, (d) allocation concealment, and (e) outcome concealment, was assessed as yes/no. In line with recommendations by Cochrane Collaboration (Higgins & Green, 2011), if no information was available on these aspects of bias, a separate code ('not described') was assigned. Hence, in this analysis, *ks* differed considerably depending on reporting and the codes assigned.

The first author coded all studies, and a subset (at least 10%) were agreed with a second coder by consensus. There were no disagreements in the reviewed subset.

### **Meta-Analytic Strategy**

Our meta-analysis required that we obtain a standardized mean difference (Hedges' *g*) effect size (ES) for each control and mental simulation condition comparison. A positive ES



indicated the degree of the mental simulation effect in favor of the experimental condition (a  $g$  of zero indicated no benefit; and a negative  $g$  represented a detrimental effect of mental simulation compared with a control group). As is typical in recent meta-analyses (e.g., Harkin et al., 2016), we corrected for small sample sizes by converting all effect sizes to Hedges'  $g$  (Hedges & Olkin, 1985), which was used in all subsequent analyses.

The following descriptions apply to comparisons that had a continuous outcome variable. Where pre-intervention (baseline) and post-intervention measures were reported, these were used to calculate pre-post change scores (by subtracting the pre from the post measure); otherwise post-test scores alone were used. To ensure conservative estimates, where multiple timepoints were recorded, only the farthest follow-up test was used to calculate the effect size.

Where means and standard deviation were not reported, but sample sizes were, we calculated Hedges'  $g$  using the relevant  $p$  or  $t$  value. Where no comparison statistics were reported (such as an absence of means and standard deviations), study authors were contacted via email ( $k = 13$ , 11% of total studies), with 3 responding with data, 4 responding but unavailable data and 6 not responding. With no comparison data, the  $p$  value was assumed to be .05 when reported as "significantly different" and Hedges'  $g = 0.00$  when not (see Harkin et al., 2016 for a similar strategy). Where dependent variables were similar (e.g., binge and overall alcohol consumption, Hagger et al., 2012b), an average was calculated. In addition, some studies measured qualitatively different outcome measures (e.g., speed and frequency), which may react differently to mental simulation. Where this was the case, up to three dependent variables (and effect sizes) were calculated per study to assess whether mental simulation is more effective for specific outcome measures<sup>2</sup>. Where appropriate (e.g.,

---

<sup>2</sup> However, it is notable that the majority of control/mental simulation comparisons had only one type of outcome measure (96%) with five having two outcome measures.

duration not on target, for frequency of errors), dependent variables were recalculated (or reversed) so that *positive effect sizes always represented positive effects of mental simulation upon behavior versus a control condition* (e.g., duration on target, more efficient and effective behavior). Dichotomous data were converted into Hedges' *g*, by entering proportions of the control and treatment groups successful in the task using the online 'Practical Meta-Analysis Effect Size Calculator' (<https://campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php>, see 'binary proportions') accompanying Lipsey and Wilson (2001). Heterogeneity of effect sizes was determined by the Q statistic and a significant heterogeneity indicates a need to uncover potential moderating effects.

We employed a multilevel meta-analysis to account for the dependency between effect sizes from one or more experiments from the same article (of the 94 studies included, 20, or 21%, contributed two or more effect sizes; all other studies contributed one effect size each). Using a univariate model would have forced us to drop observations and aggregate to one effect size per sample (or falsely treat each comparison as independent), whereas the multi-level approach allows us explicitly to incorporate this dependency into the model structure. In other words, multiple experiments are often embedded in a single study and therefore modeled as a multi-level structure using clustering through random coefficients. As appropriate in the social sciences due to the natural heterogeneity of findings, we apply a mixed effects model that incorporates random coefficients (Viechtbauer, 2010, Overton, 1998). Random effects models are suited to meta-analyses that aim to generalize beyond the included studies in contrast to fixed effect models (Field & Gillett, 2010) - an aim held at project start. We also use inverse-variance weighting such that studies with more precision (usually those with larger samples) receive relatively more weight.

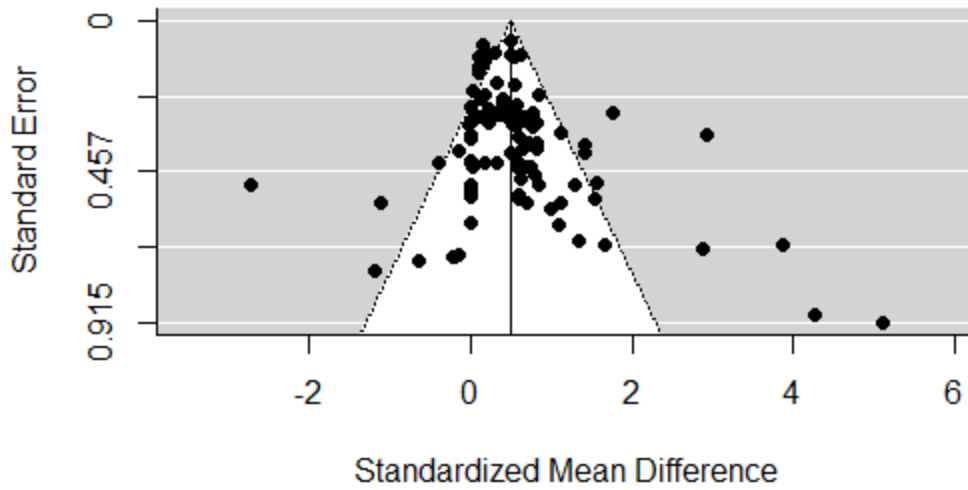
To explore the influence of moderators on the main relationship, we break the meta-analysis down into subgroups, including for different mental simulation subtypes and for other variables such as; outcome type, domain and methodological characteristics of the studies. This allows us to provide specific effect sizes for each of the subgroups and examine whether there are (significant) differences between these subgroups using a multilevel meta-regression approach. We use the *metafor* package for R to compute our models.

Although the protocol of this meta-analysis was not pre-registered, the analysis script and data files are accessible via the open science framework (DOI 10.17605/OSF.IO/H2F7E) in line with the open science agenda (Munafò, Nosek, Bishop, Button, Chambers et al., 2017).

Publication bias was assessed by the statistical and visual inspection of a funnel plot (Figure 2). In this plot, the effect sizes (x-axis) from the studies are plotted against their standard errors (y-axis) and we find no evidence of asymmetry in this plot (asymmetry is assumed to indicate publication bias, where less precise studies (with smaller sample sizes and larger standard error) produce more positive effects). Additionally, we conducted an “Egger-style” analysis (Egger, Davey Smith, Schneider, & Minder, 1997), by including the total sample size of the effects as a predictor in a meta-regression model. This approach allows us to maintain the integrity of the multilevel random effects model and still assess the influence of publication bias on our results directly. The results show that there is no relationship between the observed effects and the precision of the studies as estimated by the total sample size ( $N$  of experimental group and control group combined) ( $\beta = -.001, p = .2284$ ) and therefore, there is no evidence that publication bias poses a major threat to the validity of our results.

Note that other, more traditional, publication bias analyses are not applicable to the multilevel multivariate meta-regression model employed here. In particular, standard Failsafe

Numbers such as Rosenthal's (1979) do not account for heterogeneity and hence, reflect a fixed effect model that is deemed inadequate across most social sciences.



**Figure 2:** Main Funnel Plot for all included control-mental simulation comparisons ( $k = 123$ )

## Results

### Characteristics of Included Studies

The included studies yielded 123 independent effect sizes associated with mental simulation-control comparisons among 5685 participants (control condition *ns* range 5 to 269; experimental condition *ns* range 5 to 181, see *Supplementary Materials*). The average age of participants was 24.97 (SD = 6.34), with the vast majority of studies using student samples (78%). The grand mean female:male participants ratio was 45:55 demonstrating a roughly even distribution across studies. Most behaviors were measured in terms of frequency of occurrence ( $k = 84$ ); with the next common measuring speed ( $k = 35$ ), then accuracy ( $k = 4$ ).

### Overall Effect of Mental Simulation on Behavior

Here, and in all subsequent analyses, we applied a multilevel random effects meta-analysis, in order to account for the dependency originated from several experiments conducted within one study (observable in Table 1). The weighted average effect of mental simulation upon subsequent behavior was Hedges'  $g = 0.49$  [0.37; 0.62],  $k = 123$  (94 studies), demonstrating an overall medium-sized positive effect of the intervention which was statistically significant ( $z(122) = 7.64, p < .0001$ ). Overall, there was significant variability in effect sizes across the 123 comparisons ( $Q = 375.38, p < 0.001$ ). Indeed, although the majority of studies found small-medium positive effects (as defined by Cohen, 1988), neutral ( $g = 0.00$ ) and negative effects were also evident (e.g.,  $g = -2.73$ , Rozand et al., 2016)<sup>3</sup>. A forest plot of all effect sizes (and the overall effect size) is represented in Figure 3.

---

<sup>3</sup> The hypothesis that the degree of mental imagery would mediate the mental simulation effect has been proposed previously (Lang, 1979; Richardson, 1967). As an exploratory measure, we aimed to identify whether mental imagery capacity (as an individual difference measure) mediated the relation between mental simulation intervention and the behavioral outcome. However, after inspecting each study for quantitative data, only six comparisons reported the relation between mental imagery and behavior. It was therefore unfeasible to carry out meta-analyses of this data (similar to Conroy & Hagger, 2018), so we briefly

### Mental Simulation Subtypes

In coding the subtypes of mental simulation, some subtypes were more common than others. At least one study was found for each of the following subtypes; *process (standard)*, *performance (inferior standard, superior)* and *outcome (inferior, superior) simulations*. Additionally, it was found that some studies used composite subtypes, which were *inferior outcome plus process simulations* and *superior outcome plus process simulations*. No subtype overlapped two categories, and no studies were found in the literature utilizing *inferior process*, *superior process*, or *standard outcome simulations*. The full list of subtypes identified (with relevant illustrative examples) is presented in Table 2.

---

summarize their results thus: Of the eight that assessed the relation between mental imagery and behavior in simulation conditions, only one reported a significant correlation – indicating a positive association between level of mental imagery and behavior (spearman’s  $\rho = .62$ ,  $p < .01$ , Arora et al., 2011). All other studies reported no significant relation between mental imagery and behavior (all  $ps > .77$ , Epstein et al., 1980 Exp 1; Callow et al., 2013, Exp 1 & Exp 3; Hayter et al., 2003; Ruffino et al., 2017; Woolfolk et al., 1985).

**Table 2:** Subtypes of mental simulation (and related examples) featured in interventions in the meta-analysis.

<b>MS Class</b>	<b>Process</b>	<b>Performance</b>			<b>Outcome</b>		<b>Composites</b>	
<i>Label</i>	Standard/ unspecified	Inferior/ below average	Standard/ unspecified	Superior/ above average	Inferior	Superior	Inferior outcome + process	Superior outcome + process
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
<i>Number of effect sizes</i>	5	4	52	34	3	15	2	8
<i>Description</i>	Imagining the steps or plan involved in achieving a desired outcome (i.e., planning)	Imagining a performance which is below average	Mimic actual performance with no mention of effectiveness or efficiency of performance	Imagining a performance which is above average	Imagining an undesired or feared outcome	Imagining a desired or wished-for outcome	Simulation of a undesired outcome and simulation of ways to avoid this outcome	Simulation of a desired outcome and simulation of a barrier impeding the outcome
<i>Relevant behaviors within the context of the literature</i>	Imagining ‘sub-goals’ or ‘steps’ toward a goal	Imagining ‘errors’, ‘being slow’, or ‘feeling uncomfortable’	Imagining same ‘responses’, ‘actions’ and ‘behaviors’ as actual performance	Imagining ‘perfect’, ‘fast’ or ‘comfortable’ performance	Imagining ‘receiving a low score’ or ‘being unhealthy’	Imagining ‘breaking a habit’, or ‘getting a good grade’	Simulating a poor result on an attention task, and consideration of ways to avoid this outcome. (Lukaszewski2012, experiment 1)	Mentally simulating a desired outcome with a potential barrier (Adriaanse et al., 2010).

We report the results for the weighted effect sizes of each mental simulation subtype separately in Table 3. When considering individual subtypes, it was found that the following mental simulations led to positive effects significantly different than controls: *Process*, *standard performance*, *superior performance*, *superior outcome* and *superior outcome combined with process simulations*. Mental simulations that included a negative purpose (*inferior performance*, *inferior outcome*, *inferior outcome combined with process simulations*) failed to demonstrate a reliable positive effect. While some subtypes appear to have different effects than others, an omnibus test of moderators shows that the effects are not statistically significantly different between subtypes of mental simulation ( $QM(df = 7) = 10.47, p = 0.1636$ ). In other words, mental simulation appears to have a similarly large and positive effect on task performance across various subtypes. However, note that the number of effects included in several of these subgroups was very small ( $k$ ) and reliable inferences cannot be drawn from those numbers.

### **Exploratory Analysis of Mental Simulation Purpose**

As an exploratory analysis, we conducted an aggregated analysis of purpose (inferior, standard and superior), finding that superior and standard simulations were more effective than inferior simulations.

Some subtypes in the original analysis had only very few observations (inferior performance,  $k = 4$ ; inferior outcome,  $k = 3$ ; combined process and inferior outcome,  $k = 2$ ). Here, we explain an exploratory analysis in which, due to lack of precision when examining the effect of mental simulation subtype, subtypes were aggregated into 3 instead of 8 mental simulation subtypes. Mental simulation subtypes were collapsed into three main types; inferior, standard and superior (referring to mental simulations with an inferior purpose such as imagining poor performance on a task; standard purpose, where the performance within



the mental simulation was not specified; and superior purpose in which participants are asked to imagine a good, optimal or best performance within their simulation).

Similar to the main analysis, a multilevel meta-regression using a random effects model was applied to investigate whether the effect of mental simulation on behavior is moderated by the ‘purpose’ of the mental simulation itself (inferior, standard, superior).

The results suggest that the purpose of the simulation influences the effect it has on behavior. In particular, the meta-regression shows that the purpose aspect of the simulation explains part of the heterogeneity in effects ( $QM = 6.13, p = 0.047$ ). The effect of inferior mental simulation is indistinguishable from zero ( $g = 0.01, 95\% \text{ CI: } -0.41; 0.44, p = 0.94$ ). However, the effects of both standard and superior mental simulations are significantly larger (and positive): standard ( $\beta = 0.46, 95\% \text{ CI: } 0.02; 0.91, p = 0.04$ ) and superior ( $\beta = 0.53, 95\% \text{ CI: } 0.11; 0.94, p = 0.014$ ).

### **Outcome Measures: Speed, Frequency and Accuracy**

As outlined in the Method, we aimed to assess whether the type of outcome measure moderated the size of the effects of mental simulation on behavior. It was found that the type of outcome measurement did not significantly affect the mental simulation effect size. Regardless of whether the outcome measure was based on speed, frequency or accuracy of behavior, a positive effect significantly above zero was found. Thus, mental simulation appears to have positive behavioral effects, regardless of the outcome type.

**Table 3.** Effect size and heterogeneity as a function of mental simulation subtype

Moderator	<i>k</i> ( <i>s</i> )	<i>g</i>	95% CI	QE	QM	<i>p</i>
Mental Simulation Subtype	123 (94)	-	-	342.25*	11.53	.12
Process (standard)	5 (5)	.17	[-.09; .43]	1.64		.20
Performance (inferior)	4 (4)	-.40	[-1.03; 0.22]	4.84		.20
Performance (standard)	52 (45)	.48	[.31; .65]	147.41*		<.0001
Performance (superior)	34 (33)	.67	[.34; 1.00]	161.94*		<.0001
Outcome (inferior)	3 (3)	.16	[-.51 ;.83]	2.41		.63
Outcome (superior)	15 (15)	.23	[.13; .34]	16.18*		<.0001
Combined (inferior outcome + process)	2 (2)	.00	[-.79; .79]	0.00		1.00
Combined (superior outcome + process)	8 (8)	.61	[.40; .81]	7.83		<.0001

**Note:** *k* = number of effect size estimates; *s* = number of studies; *g* = Hedges’ standardized mean difference; 95% CI refers to the lower and upper bounds of the 95% confidence interval around *d*; QE refers to the residual heterogeneity (Cochran’s Q) with \* indicating that the probability of homogeneity of the data is less likely than *p* = 0.05; QM refers to the omnibus test of parameters for all moderators; *p* refers to the approximate *p* value of either the *z*-value of the individual effect (*d*) or to the omnibus test of moderators (QM).

**Table 4.** Moderator analyses

Moderator	<i>k</i> ( <i>s</i> )	<i>g</i>	95% CI	QE	QM	<i>p</i>
Outcome	123 (94)	-	-	366.65*	.47	.7890
Speed	35 (31)	.373	[.137, .606]	118.77*	-	.0017
Frequency	84 (67)	.537	[.391, .683]	243.96*	-	<.0001
Accuracy	4 (4)	.329	[.030, .628]	3.93	-	.0310
Domain	123 (94)	-	-	374.00*	.31	.5774
Health	26 (16)	.371	[.241, .502]	41.67*	-	<.0001
Motor Learning	52 (44)	.588	[.319, .856]	194.08*	-	<.0001
Occupational	12 (9)	.701	[.113, 1.289]	62.56*	-	.0195
Pain	2 (1)	.537	[-.551, 1.624]	2.15	-	.3334
Social	7 (7)	.434	[.197, .672]	10.36	-	.0003
Sports	24 (17)	.338	[.113, .562]	53.87*	-	.0032
Task	123 (94)	-	-	366.59*	1.04	.5934
Motor	52 (42)	.425	[.279, .571]	108.25*	-	<.0001
Cognitive	20 (16)	.313	[.161, .465]	21.99	-	<.0001
Mixed	51 (45)	.569	[.320, .819]	236.35*	-	<.0001
Duration (coded)	115 (88)	-	-	362.84*	0.05	.8204
Short	88 (71)	.480	[.354, .607]	252.03*	-	<.0001
Medium	16 (11)	.794	[.049, 1.539]	105.20*	-	.0368
Long	11 (6)	.481	[.315, .647]	5.25	-	<.0001
Frequency	123 (94)	-	-	374.75*	1.17	.2800
Total Time (duration X frequency)	123 (94)	-	-	374.93*	1.01	.3147
Delay	112 (87)	-	-	308.37*	0.08	.7679
Incentive	49 (33)	-	-	63.23	2.48	.1151
with incentive	41 (30)	.368	[.273, .464]	62.13*	-	<.0001
no incentive	8 (3)	.061	[-.276, .397]	1.10	-	.7247
Post-only or Pre-post Change Scores	122 (93)	-	-	372.35*	0.34	.5582
Change scores	43 (30)	.455	[.163, .747]	173.34*	-	.0023
Post-only scores	79 (64)	.499	[.368, .630]	199.01*	-	<.0001
Participant Blinded	29 (26)	-	-	130.18*	0.53	.4673
Blinded	26 (23)	.599	[.331, .868]	124.93*	-	<.0001
Non-blinded	3 (3)	.833	[.317, 1.350]	5.25	-	.0016
Experimenter Blinded	23 (19)	-	-	111.71*	0.72	.3948
Blinded	14 (10)	.382	[.118, .647]	36.54*	-	.0047
Non-blinded	9 (9)	.718	[.125, 1.312]	75.18*	-	.0177
Randomization	41 (31)	-	-	175.32*	3.05	.0807
Randomized	23 (16)	.612	[.276, .948]	128.01*	-	.0004
Non-randomized	18 (15)	.176	[-.171, .523]	47.30*	-	.3214
Allocation Concealment	17 (13)	-	-	93.64*	0.06	.8044
With alloc concealment	13 (9)	.727	[.124, 1.329]	84.58*	-	.0181
No alloc concealment	4 (4)	.571	[-.059, 1.200]	9.05*	-	.0755
Mental imagery in control	123 (94)	-	-	366.28*	3.63	.1626
With imagery check	20 (16)	.311	[.141, .481]	30.12	-	.0003
Without imagery check	98 (77)	.545	[.392, .697]	334.76*	-	<.0001
Mental imagery in experimental	123 (94)	-	-	374.06*	0.68	.4107
With imagery check	70 (51)	.457	[.262, .651]	259.24*	-	<.0001
Without imagery check	53 (43)	.482	[.351, .613]	114.82*	-	<.0001

**Note:** alloc = allocation; *k* = number of effect size estimates; *s* = number of studies; *g* = Hedges' standardized mean difference; 95% CI refers to the lower and upper bounds of the 95% confidence interval around *d*; QE refers to the residual heterogeneity (Cochran's Q) with \* indicating that the probability of homogeneity of the data is less likely than *p* = 0.05; QM refers to the omnibus test of parameters for all moderators; *p* refers to the approximate *p* value of either the z-value of the individual effect (*d*) or to the omnibus test of moderators (QM).

### **Domain**

The experiments were mostly derived from the motor skills domain ( $k = 49$ ), and fewer were from health ( $k = 26$ ), sports ( $k = 24$ ) and occupational ( $k = 12$ ) psychology. Far fewer studies investigated the effect of mental simulation on withstanding pain ( $k = 2$ ) or social behaviors ( $k = 7$ ). Domain was found not to be a significant moderator of the mental simulation effect in the omnibus test. Effects sizes significantly above zero were demonstrated in almost every domain except pain which did not significantly differ from zero (see Table 4). However, the lack of a statistical significance in this type of studies should be taken with caution due to the limited amount of studies included here ( $k = 2$ ) and the effect estimate is similarly positive.

### **Type of Task**

Experiments were divided into those that contained mainly cognitive or physical attributes, or a combination of both, to assess the type of the task that participants engaged with. However, type of task was not a significant moderator of the mental simulation effect (see Table 4).

### **Dosage: Duration and Frequency of Mental Simulation**

To examine the link between dosage of mental simulation and the effect on behavior, we conducted a meta-regression on two components of dosage independently; *duration* and *frequency* of mental simulation training (see Driskell et al, 1994 and Feltz & Lander, 1983 for a similar approach). We assessed whether each factor significantly predicted effect sizes.

The mean duration of mental simulation was 23.59 mins (SD = 73.02; range = 1-350 mins; median = 3), with the most common being 1 minute ( $k = 36$ ) and the next most common being 3 minutes ( $k = 28$ ) then 5 minutes ( $k = 17$ ). Mental simulation durations were converted into a categorical variable, coded as either *short* (1-5 mins), *medium* (6-20 mins) or

*long* (21 mins+)<sup>4</sup>. It was found that effect size was not moderated by the duration of a mental simulation.

To determine frequency of mental simulation, we recorded the amount of mental simulations. The average amount of simulations was 246.08 (SD= 1680.28, range = 1-18000, median = 3) indicating substantial variability. Meta-regression with total amount of mental simulations established that frequency of mental simulation did not moderate effect size. A final analysis with dosage using total time (duration X frequency of mental simulation) similarly indicated no moderating effect.

This illustrates that, typically, mental simulation interventions are of a short duration and that simply increasing frequency of simulation does not systematically alter observed effects.

### **Delay**

It is possible that delay (or time/retention interval) between the last mental simulation and the final measure of behavior accounts for variability in the effect sizes. Although the behavioral test was most frequently administered on the same day as the intervention ( $k = 66$ , 59% of all 112 comparisons which reported delay), some studies were designed with long delays, with a maximum of 210 days delay (mean delay across studies= 12.61 days, SD = 33.96). Nevertheless, it was found that delay did not predict the size of the mental simulation effect (see Table 4).

### **Incentive**

Of all comparisons, only  $k = 49$  (40%) provided enough information to categorize the study as either offering an extrinsic incentive (e.g., course credit, gifts, cash) or offering none. When no incentive was provided, the effect size was not significantly different than zero ( $g =$

---

<sup>4</sup> Duration was coded as opposed to using the continuous data of minutes of each simulation, because there were some extremely long durations (e.g., Pascual-Leone, 1995), that meant the data would be dramatically skewed. For the sake of completeness, when minutes of each mental simulation was used, the same result was found: That duration of mental simulation did not significantly moderate effect size.

0.06,  $k = 8$ ) whereas when participants were offered a form of incentive the effect size was 0.37 ( $k = 41$ ) and significantly different than zero. Nonetheless, the difference between this non-significant (no incentive) and significant effect is not itself statistically significant (see Table 4).

### **Moderation Analyses: Overall Summary**

To briefly summarize the main results of the moderation analyses (displayed in Table 3 and 4), none of the omnibus tests for the variables tested showed statistically significant effects at the .05 level. Thus, our conclusions are primarily based on identifying and comparing different levels of each variable, based on their Hedges'  $g$  and whether each ES was significantly different than zero, or not. The exception was our exploratory moderation analyses which showed that purpose (inferior, standard, superior) had a significant moderating effect on effects sizes, and specifically, that standard and superior simulations have more beneficial effects on behavior than inferior simulations.

### **Methodological Characteristics**

Risk of bias was assessed based on five indicators (see Method). Across each of these indicators, the majority of studies (67-86%) did not provide enough information to classify, and meta-analyze. The resulting data were tested as moderators in Table 4.

Studies using a pre-post design in which effect sizes represent pre-post change scores, typically reported more conservative estimates ( $g = 0.46$ , 95% CI = [.163, .747]) than effect sizes incorporating only post-intervention scores ( $g = 0.50$ , 95% CI [.368, .630]), although the difference was not statistically significant (see Table 4). In addition, studies using pre- and post-scores showed effects that are also more uncertain and dispersed, as indicated by the wider-ranging confidence interval for this effect that approaches zero on the lower limit.

Participant and experimenter blinding did not significantly affect the results. While experiments in which participants were not blinded typically reported much larger effects ( $g$

= 0.83) than when participants were blinded ( $g = .60$ ), the difference is not significant (see Table 4). Similarly, while studies where experimenters were not blinded ( $g = .72$ ) typically reported (much) larger effects than when experimenters were blinded ( $g = .38$ ), the difference was not statistically significant (see Table 4). These results, however, lack power to detect significant differences (only 29 and 23 experiments respectively) and for those that do give information, non-blinding is rare (only 3 and 9 experiments are non-blinded to participants and experimenters, respectively).

Finally, *allocation* concealment did not affect the results. Studies in which the participants were not aware of which condition they were allocated to typically showed slightly larger effects ( $g = .72$ ) than non-concealment ( $g = .57$ ) but the difference is not significant (see Table 4). As there are no studies for which we have information that the *outcome* was not concealed, we cannot examine any differences for this methodological characteristic.

### **Manipulation Checks**

As stated in the Introduction, it was important to assess the percentage of studies that measured mental simulation in the control and mental simulation conditions, to verify fidelity to the manipulation. An assessment of all included control-experiment comparisons established that only 16% measured mental imagery in the control condition to assess spontaneously used mental imagery. However, this did not modify the effect size. As seen in Table 4, the effect size was significantly above zero regardless of this manipulation check. Furthermore, approximately half of the experiments (56%) employed active control conditions, whereas the remainder utilized a passive or ‘wait’ control condition (active controls may be more effective at eliminating imagery use in control conditions). Fifty-seven percent control-experiment comparisons measured mental imagery to confirm basic

compliance to the intervention, but, as Table 4 shows, this did not significantly moderate effect size.



### Discussion

In this meta-analysis, our aims were to elucidate for the first time the unique effects of mental simulation on behavior change across a range of domains and uncover which mental simulations, and which intervention characteristics, moderated this effect. In brief, are there identifiable characteristics that enhance the behavioral effect of mental simulation? To answer this question, we improve on previous reviews by: (1) conducting a thorough search of all randomized controlled studies on the mental simulation effect and (2) using multi-level modelling to assess the overall effect and potential moderators. The following discussion considers the theoretical and practice implications of the present findings.

We confirmed findings from previous reviews, indicating an overall moderate positive effect of mental simulation compared with non-mental simulation control conditions (Conroy & Hagger, 2018; Feltz & Landers, 1983; Driskell et al., 1994). In an initial meta-analysis, Feltz and Landers (1984) established that mental simulation had a reliable moderate effect ( $d = .48$ , 60 experimental-control comparisons) on behavior when compared with no practice. Driskell and colleagues' (1994) analysis replicated these results with more robust inclusion criteria ( $d = .53$ , 62 comparisons). In a more recent meta-analysis focusing on health studies by Conroy & Hagger, 2018, with 33 comparisons, a similar effect size was found, showing that, on average, mental simulation substantially improves behavior by approximately half of one standard deviation. In the present review, we found a highly similar effect to these reviews ( $g = .49$ , 123 comparisons) when only including randomized controlled experiments.

The greater power (and multilevel nature) of our analyses, based on over double the number of comparisons of previous reviews, highlights the robustness of this effect (although we should note that actual physical practice has been found to be more effective than simulation, Driskell et al., 1994). That this was consistently demonstrated across such a diverse range of behaviors - from laboratory-based reaction time tasks to slalom skiing and to

social behaviors - underlines the reliability and robustness of this effect<sup>5</sup>. This finding corresponds well with proposals that mental simulations can be adopted as an effective behavior change technique that links thought or intentions with action (Johnson & Sherman, 1990; Taylor & Pham, 1996; Pham & Taylor, 1999) and validates continued professional applications of mental simulation in sports (e.g., Nicholls, 2017) and health (e.g., Hagger, 2016). The size of the effects associated with mental simulation are much larger than those associated with other behavior change techniques, such as self-incentives and self-rewards ( $d = 0.17$ , Brown, Smith, Epton, & Armitage, 2017). The implication is that mental simulation may be a particularly potent behavior change technique that could be a starting point for the development of more complex behavior change interventions.

The present review has multi-disciplinary implications for mental simulation research. Namely, unlike previous reviews we provide an integrative approach, and identify key classes of mental simulations. To this end, three classes of mental simulations were identified and compared (performance, outcome and process simulations). This integrative approach will enable researchers to assess which types of mental simulations are most effective, and which are relatively inconsequential. It also allows generalities across domains to be interpreted, highlighting under-researched categories of simulation. For example, this meta-analysis showed that process and outcome simulations, but not performance simulations, have been successfully applied in health psychology. This review also highlights the opportunity to develop behavioral interventions across domains where mental simulation is in its infancy (such as social psychology, see Meleady et al., 2013).

The multilevel meta-analytic approach employed here allowed us to examine whether effect sizes were moderated by the subtype of mental simulation. First, we categorized all

---

<sup>5</sup> Note that the overall effect size is a conservative estimate of the mental simulation effect, as only RCTs were included and analytical measures were taken to reduce over-estimation of the average effect size (see *Methods: Meta-Analytic Strategy*).

experimental conditions into one of the categories identified in a taxonomy based on possible subtypes of simulation (Table 1). Examination of existing studies led to the addition of composite categories (e.g., *process with superior outcomes*), and reduction of others (no *positive process* or *inferior process* simulations were found; although it remains to be seen whether future studies utilize these). All experimental conditions identified were labelled with one of the remaining eight subtypes. Although the planned omnibus test used in the analysis did not reveal a significant effect, we examined whether each level of the categorical moderator leads to either a null *or* statistically significant positive effect, in a one-by-one fashion, replicating a common approach (e.g., Adriaanse, Vinkers, De Ridder, Hox, & De Wit, 2011; Conroy & Hagger, 2018). Adopting this technique, some subtypes were more effective than others. These differences warrant interpretation and discussion.

### **Superior mental simulations improve targeted behavior**

One finding was clearly distinguishable and specific: subtypes of mental simulations with positive components (*superior performance*, *superior outcome* and *superior outcome combined with process simulations*) consistently led to reliable positive behavioral effects (see Table 3). This was confirmed by an exploratory multilevel meta-regression analysis of purpose (inferior, standard, superior) showing that superior simulations led to better behavioral outcomes than inferior ones. What mechanisms explain these effects? Here, it is important to apply explanations to each subtype in turn as they represent different ways humans conceptualize the future (see Introduction and Table 1).

Firstly, the effects of *superior performance* simulations may be understood in terms of the idea that mental and physical practice of a specific behavior are assumed to rely on overlapping neural correlates, dynamic qualities and temporal durations (Allami et al., 2008; Decety & Grèzes, 2006; Jeannerod, 2001; referred to herein as the *overlapping systems hypothesis*). If performance simulations share common neuropsychological processes with

actual performance, reasonably, simulating a superior version of that behavior should garner more beneficial changes in behavior than other versions (standard, inferior). Thus, when one imagines optimal performance on a finger pointing task (Debarnot, Clerget & Oliver, 2011) subsequent behavior on the task improves, and across studies included here, such simulations led to a large behavioral effect compared to controls (Hedges'  $g = .67$ ).

The same logic cannot be applied to *superior outcome* simulations, however, because such simulations are – by definition – focused upon the consequences of the enacted behavior rather than the behavior itself (i.e., *if you perform well, you will attain outcome X*, such as feeling positive and physically active, Andersson & Moss, 2011). Benefits of superior outcome simulations are often explained by the motivational ‘pull’ of simulating a desired scenario (Hagger, Lonsdale & Chatzisarantis, 2011; Vasque & Buehler, 2007). Superior outcome simulations typically focus upon simulating emotions, rather than action plans; they are motivational rather than volitional (Taylor & Pham, 1996). This may be why simply thinking about a desired future is seen as inadequate as a behavior change technique on its own (Taylor & Pham, 1996; Pham & Taylor, 1999). Indeed, if the simulation involves fantasizing, motivation may *decrease* due to the individual feeling a premature sense of achievement in-the-now, reducing the drive to strive for an end-goal (Oettingen, 2012). The results of the present meta-analysis are in line with research on possible selves (Markus & Nurius, 1986) – that is, simulations of positive scenarios which increase motivation towards that goal, breeding energisation and action. Thus, it is apparent here that simulation of superior outcomes (and positive simulations in general) are among several beneficial mental simulation subtypes which can be used as behavior change techniques. However, it is important to state that the effect of superior *outcome* simulations (Hedges'  $g = .23$ ) was smaller than effects found for other subtypes - possibly because the effect of positive

simulations are moderated by factors such as personality type and imagery content (Vasque & Buehler, 2007) or because these simulations are *by nature* less effective.

**No evidence found for an effect of inferior mental simulations on behavior**

In contrast, when including composite categories, all subtype incorporating a negative purpose (*inferior outcome, inferior performance, process + inferior outcome*) failed to show a significant positive behavioral effect. The difference between inferior and standard/superior simulations was also evident in the moderation analysis of purpose. Thus, in the context of this meta-analysis comparing only experiments on behavioral measures in healthy population, mental simulations incorporating an inferior component (e.g., imagining not withstanding pain, Alden et al, 2001; or gaining weight, Marszał-Wiśniewska & Jarczewska-Gerc, 2016) do not appear to benefit future behavior. With inferior outcome simulations (used separately or combined with process simulations), one might expect a *reduction* in undesirable behaviors (e.g., unhealthy eating) – sometimes called ‘avoidance behaviors’ – by simulating the negative consequences of those behaviors (e.g., weight gain, see Lang, 1995; Elliot, 2006): an ironic effect. Indeed, Taylor’s (1991) examination of the *mobilization-minimization hypothesis* supports the idea that negatively-valenced information, in general, recruits stronger cognitive, social and behavioral responses than does neutral or positive information. However, it is important to clarify that our conceptualization of inferior simulations does not have a one-to-one mapping on to what has been termed negative or negatively-valenced simulations (Gamble et al., 2019).

In the context of this meta-analysis, we found that simply imagining *inferior outcomes* was largely inconsequential for behavior. No subtype of simulation with an inferior component garnered reliably positive behavioral effects, whether mapped to the actual performance or the outcomes of behavior. Thus, by including for the first time inferior subtypes of mental simulation in a meta-analysis (when conceptualized differently than

emotionally negative simulations, see Gamble et al., 2019a), we present novel evidence that disentangles effects that were previously studied homogenously.

It is clear, as anticipated by Pham and Taylor (1999) almost 20 years ago, that more systematic examinations of positive *and* negative outcome simulations (and their potential mediators) are required. Although researchers may want to demonstrate positive behavioral effects, a nuanced understanding of how and why simulations affect behavior (both negatively and positively) requires more complex experimental studies which randomise participants to superior, standard and inferior simulations within the same study.

The (contrasting) effects of superior and inferior simulations found in this review should also be observed in the context of detailed reviews of mental imagery emphasizing the powerful physiological and cognitive effects of emotional imagery (Decety & Grèzes, 2006). This is because it is likely that superior and inferior simulations are associated with positive and negative emotions, respectively. One area that has examined the effects of negatively emotional simulations is clinical psychology. For example, recent studies have showed an inverse relationship between negative simulations of goals at baseline and goal attainment after a delay of two months (Gamble Tippett, & Moreau, & Addis., 2019b) and that highly negative emotional simulations can predict or maintain clinical symptoms (in depression, Holmes, Blackwell, Burnett Heyes, Renner, & Raes, 2016, and bipolar disorder, Hales, Deepröse, Goodwin, & Holmes, 2011). Nevertheless, we are cautious about over-generalizing the findings of this meta-analysis to other domains, and stress that the absence of evidence (of an effect of inferior simulations) is not evidence of absence (of an effect). Future research would benefit from determining the independent effects of emotion on effects of these types of simulations.

**Limited and equivocal evidence that mentally simulating the process improves targeted behaviour**

In the present meta-analysis, when participants focused upon instrumental plans toward one's goal, rather than goal outcomes; the means rather than the end<sup>6</sup>, there was a small positive effect that did not reach statistical significance within our multilevel moderation model. This does not therefore clearly validate Taylor and colleagues' (Pham & Taylor, 1999; Taylor, Pham, Rivkin & Armor, 1998) theoretical work on the effectiveness of process simulations for subsequent behavior and self-regulation. The benefits of process simulations are often mentioned (across cognitive and health psychology, Schacter, 2012; Murru & Martin Ginis, 2010) with several authors assuming that such simulations are able to improve one's decision making about future acts – helping one decide when, where and how to achieve an end goal. Thus, a positive behavioral effect is implied. However, results from this meta-analysis show that caution is needed in making any strong assumptions about process simulations, at least until further empirical research conclusively replicates positive effects following process simulations. It is noteworthy in this context that the empirical study often cited (e.g., by Epstude, Scholl & Roese, 2016; Freund & Hennecke, 2015) as key evidence supporting this principle - Pham and Taylor (1999) - failed to find a statistically significant effect of process simulation on exam performance (behavior) over the control group ( $p < .09$ , coded herein as  $d = 0$  based on our meta-analytic strategy). Nevertheless, the small amount of studies identified herein ( $k = 5$ , with only 4 studies conducted since Pham & Taylor, 1999) which fit our criteria and focus on process simulations, shows that this research is in its infancy, and we recommend that more empirical research is needed before claims are made about the behavioral benefits of process simulations.

### **Mentally simulating performance improves behavior**

---

<sup>6</sup>However, note that Taylor and colleagues' comparisons between process and outcome were based on positive outcomes, whereas we additionally compare positive and negative outcomes.

The present meta-analysis found a significant and reliable effect of standard performance simulation on behavior. Also, an exploratory analysis found that standard simulations worked similarly to superior simulations, and outperformed inferior ones. There are clear, well-documented streams of research that allow explanation of the positive effects of simply imagining a task one will later perform – i.e., *performance simulations*. Traditional explanations of mental simulation’s effects on behavior focused on the idea that humans can mentally replicate a physically practiced/modeled behavior using similar cognitive and neurobiological processes as those necessary for its physical enactment (Corbin, 1972; Feltz & Landers, 1984). In terms of cognition, internalized representations of a specific action purportedly follow similar biomechanical and cognitive rules as actual behavior (Driskell, Copper & Moran, 1994). Furthermore, neuroscientific findings have lent much support to the *overlapping systems hypothesis* (see Decety & Grèzes, 2006; Jeannerod, 2001 for reviews). Applying this principle to the studies included here, through repeated mental practice, one can ‘engrain’ certain behavior-chains (or ‘behavioral schemata’), from initiation to completion, such that behavior-chains are retrieved and enacted *automatically* (for example, once a golf putt is repeatedly *simulated*, the same process can be activated and enacted, relatively automatically, when a golfer holds a starting position *in an actual game*, Budney et al., 1990).

This mechanism accords well with dual-process theories of behavior which emphasize that behavior is coordinated via two parallel routes; an automatic process, which is rapid and stimulus-driven, and a volitional process which relies on conscious intention and reflection (Strack & Deutsch, 2004, see also Hagger, Trost, Keech, Chan, & Hamilton, 2017)<sup>7</sup>. Under

---

<sup>7</sup> Although we adopt a neuro-cognitive explanation here, it is important to note the dearth of social/health research utilizing *performance simulations* (see *Supplementary Materials*). Hence, it is possible that explanatory variables typically adopted in health/social psychology (e.g., *self-efficacy*, Bandura, 1997; *framing* and *self-appraisal*, see Baumeister, Masicompo and Vohs, 2011) may play a - hitherto undiscovered - role in explaining performance simulations.



this framework, performance simulations change behavior via a mostly automatic process. In contrast, *outcome simulations* do not involve rehearsal of the to-be-performed-behavior, and thus cannot be explained by the instantiation of an automatic process.

Due to their ability to explain a diverse range of psychological phenomena, dual-process models such as the one proposed by Strack and Deutsch (2004) could be useful to distinguish the processes through which *process*, *performance* and *outcome simulations* lead to changes in behavior (which may utilize reflective and/or impulsive processes, respectively). Applications of this dual-process model to behavior change (such as Hagger et al., 2017), specifically concerning mental simulations, represent a possible fruitful avenue of research revealed by the current meta-analysis and a recent meta-analysis by Conroy & Hagger (2018).

### **Combining Subtypes of Simulations**

A sizeable minority of studies combined subtypes of simulation; specifically, here we found several studies which integrated process and outcome simulations. Within this category, the most numerous were mental contrasting interventions— a technique devised by Oettingen and colleagues (see 2012) to combine volitional and motivational aspects of mental simulation (typically aligned with *process* and *outcome simulations*, respectively, yet distinct from this conceptualization). There exists supportive evidence spanning over 15 years showing positive behavioral effects of mental contrasting (e.g., Oettingen & Wadden, 1991; Oettingen & Mayer, 2002; see Oettingen, 2012 for a review). Our data largely agreed. Although not the largest, a homogenous positive effect was found when examining combined process *and* superior outcome simulations. On the other hand, only two studies (both indicating null effects) were identified incorporating *process and inferior outcome* simulations. Thus, here we did not have enough data to draw firm conclusions.

### **Summary of Effects as a Function of Subtypes of Mental Simulation**

In summary, although the present meta-analysis replicates similar positive effects as found previously (e.g., Driskell, Copper & Moran, 1994), we show that this effect depends on different classes of mental simulation. Furthermore, these different classes could rely on different mechanisms. This is important for theoretical developments and practical applications of this technique, because specific classes of mental simulation could be more appropriate for specific situations and populations (e.g., especially clinical populations with cognitive or emotion dysfunctions).

However, there are limits to what these data say about ‘standard’ mental simulations. Fundamentally, because our types of mental simulation were based on differences in instruction, this opens the question of what kinds of simulation participants engaged in when not directed to imagine either a superior or inferior behavior (‘standard’ simulations, see Table 1). Although we assumed these simulations would fall ‘in-between’ inferior and superior simulations, it is possible that when undirected, participants imagine idealized scenarios (i.e., ‘future positivity bias’, see Berntsen & Bohn, 2020). It would be important for future research to confirm whether such a systematic positivity bias exists in the types of mental simulation employed in studies included in this meta-analysis.

### **Do incentives moderate the mental simulation effect?**

We found a significant moderating effect for incentive, a variable selected for analysis for the first time based on a related theoretical framework (Michie et al., 2013). Extrinsic incentives (i.e., vouchers, money, course credit) influenced the effectiveness of mental simulation interventions such that mental simulations without incentives failed to produce a significant effect ( $g = .06$  on average), whereas those including rewards garnered significant positive effects. Although based on a small sample of studies (some with small  $N$  sizes) this represents a novel avenue to examine in the mental simulation literature, and supports our hypothesis that offering rewards represents a behavior change technique based on increased

motivation (Brase, 2009), and would ‘boost’ any effects of mental simulation. Importantly, this ‘boost’ was not the result of contingencies within studies whereby the reward was tied to the mental simulation. Indeed, in all relevant studies, incentives were given independently of group assignment. This finding raises the possibility that public health programs utilizing mental simulation interventions should strongly consider combining them with extrinsic reward and indicates that primary research in this area would be fruitful.

### **Other methodological characteristics**

For the remaining moderating variables (*dependent variable, domain, dosage, delay, task type, incentive* and *study quality*), no moderating effects were found. It was somewhat surprising that increased frequency and duration of mental simulation show no impact on behavior (although see Driskell et al., 1994 for a similar finding). It is probable that a minimal dosage is required for interventions to be cost- and time- effective (the ‘optimal duration’ was found to be 20 minutes in Driskell et al., 1994), after which successive simulations do not garner additional benefits. We recommend that future empirical work directly tests the effect of dosage. Specifically, well-controlled experiments manipulating the temporal spacing (e.g., massed versus distributed) of mental simulation sessions whilst holding amount of dosage constant will be required. One testable hypothesis would be that massed practice leads to less effective outcomes than distributed practice (Corbin, 1972) – a robust finding recently shown in studies of learning (see Dunlosky et al. 2013, for a review).

There was also no evidence for a reduction of the mental simulation effect over time, consistent with Conroy and Hagger (2018) (this conflicts with, Driskell et al., 1994). The lack of an effect here may be due to the fact that our delay effect was based on the time-lag between last mental simulation and *farthest* timepoint measured to avoid overestimating the mental simulation effect. In Driskell et al., (1994), it was unclear how delay was measured and whether intermediate or ‘booster’ procedures were incorporated. Thus, here we show

when the final impact of the intervention is considered, the positive effects of mental simulations are not limited by the delay.

### **Limitations**

Two main limitations of this meta-analysis require elaboration. First, although we assessed subtype differences in a one-by-one fashion, the reason why these apparently substantial differences in effect sizes (see Table 3) did not lead to a significant moderation effect on the omnibus test should be considered. We note that this null effect could be due to low power to detect large effects due to small  $k$  sizes per subtype and small sample sizes within the included studies. This meta-analysis also included studies using different tasks, outcome measures, and domains. Although this allowed us to analyze which effects stand across experimental variation, it also increases the chance that some real effects, specific to certain tasks, outcomes and domains, were masked by general methodological variability (a common issue in meta-analyses, Lipsey & Wilson, 2001).

Secondly, we restricted our search to peer-reviewed publications. The exclusion of unpublished work and the ‘grey’ literature (e.g., conference proceedings), that are sometimes included in meta-analyses (e.g., Conroy & Hagger, 2018; Harkin et al., 2016), probably did not affect the validity of our findings: we did not find evidence of publication bias (i.e., there was no relation between low precision and effect size in the studies included, see Figure 2).

### **Future Directions**

In future research, it will be important to expand the present analysis of mental simulation subtypes, which was restricted here due to the commonality of some interventions (e.g., performance simulations) and the scarcity (e.g., negative process simulations) of others. Indeed, three of the nine cells identified in possible subtypes of mental simulation (see Table 1) could not be examined due to lack of empirical research. This sets the scene for new

research to examine untested subtypes of simulation, such as inferior or superior process/plans versus an appropriate control.

Another fruitful avenue for future work would be exploring the longitudinal effects of mental simulation. The majority of studies included in this review assessed behavioral effects within one day of the intervention, and none examined effects surpassing one year. Similarly, the finding that almost all studies are conducted with young student participants drives the need for more mental simulation studies among diverse and more inaccessible populations.

Furthermore, our analysis of domain indicated the need for studies to examine the effect of ‘open’ and ‘closed’ behaviors. For example, one study included here examined this factor within badminton (serving = closed; returning = open, Jeon, Kim, Ali & Choi, 2014), defining open skills as “generally more variable, unpredictable and externally paced” whereas closed skills are “more predictable, stable and self-paced” (p. 157). It is possible that some mental simulations (e.g., performance simulations) might be less effective for open skills, being less able to map directly onto the future behavior, and draw upon the overlapping systems underlying behavior and mental imagery of behavior (Decety & Grèzes, 2006). Similarly, in health psychology, researchers may require interventions that advocate flexible, open behaviors (e.g., in selecting locations to eat healthily), driving a need to consider whether the type of simulation utilized for an intervention lends itself to ‘open’ or ‘closed’ behaviors.

Moreover, the difference between open and closed behaviors links with a more conceptual issue: whether some of the simulations can be defined as future-oriented or if some are merely reactivated memories. We outline two types of simulation that may or may not be reactivations of memory. Firstly, where a participant is asked to simulate a circumscribed fixed pattern of behavior, they may indeed rely on a memory activation when simulating the task (e.g., a ‘closed’ behavior, see above). This would involve no ‘temporal

tag' as a future event. Secondly, where a participant must put him or herself into a future context or where behaviors are more flexible (e.g., 'open' tasks), an explicit simulated event is required with a future 'temporal tag'. Research has shown that such 'episodic future thoughts' (Atance & O'Neill, 2001), although relying on memory processes and brain structures, involve a constructive recombination of episodic details into a novel simulated scenario (see Schacter et al., 2012 for a review). Current models of mental simulations do not fully account for these differences, highlighting the need for further theoretical work.

This meta-analysis highlighted a need for studies to systematically include a manipulation check to assess whether participants in experimental conditions actually created a mental image in their 'mind's eye' - whether that be a verbal description of the image (Hagger et al., 2011), a rating of vivid imagery (Andre & Means, 1986) or a judgment of imagery duration (Debarnot, Clerget & Olivier, 2011) related to the simulation. Furthermore, the moderation analyses herein rests somewhat on the assumption that participants were indeed simulating in the way they were instructed. Although there was no strong evidence to the contrary (i.e., that participants were regularly excluded for simulating different scenarios/behaviors), and over half of the included studies did carry out a manipulation check on simulation content, it would be important for reliability that all future studies verify the content of the simulation with such a manipulation check. Importantly, very few studies checked whether those in the control condition engaged in mental simulation (in which opportunities exist to simulate the subsequent behavior), indicating a need for studies to include a retrospective check. Although a moderation effect of verification was not found, if a check does not happen, the mental simulation effect may be reduced due to imagery use across experimental and control conditions.

Finally, according to a review of placebo effects (Boot, Simons, Stothart & Stutts, 2013), an adequate control condition in psychological interventions should ensure that control

participants' expectations about potential benefits of the 'intervention' should match those in experimental conditions. Although approximately half of all studies included here adopted active controls (e.g., reading poetry), it is likely that those in the control and experimental conditions differed on whether and how much they expected positive behavioral effects from 'interventions' they received. For instance, those in true experimental conditions may realize similarities between the simulation and task, and guess they are in the experimental arm; those in control conditions may, in contrast, realize the 'irrelevant' nature of the control 'intervention'. This would affect expectations of task success. This limitation of previous studies can be overcome by measuring outcome expectations after intervention/control (Boot et al., 2013), or by informing those in the experimental group they are in the control group, and assessing feedforward effects on behavior<sup>8</sup>.

### **General Summary**

Several authors have argued and built models around the proposal that mental simulation is a desirable method to increase optimal behaviors (e.g., Oettingen, 2012; Taylor & Pham, 1996). Here, we analyzed if those benefits extended across several conditions, as defined by a taxonomy informed by extant empirical research. We found that across a range of behaviors, classifications identified provided a useful framework to differentiate subtypes of simulations and their effects. These findings not only have clear implications for theoretical understanding of mental simulation effects but may aid professionals seeking evidence-based and cost-effective methods of changing behavior. It is hoped that a common language will motivate more cross-pollination across sub-disciplines. Furthermore, new research programs will benefit from delineating not only *in which circumstances* but also *how* mental simulation changes behavior (see Michie et al., 2013). These new research programs may benefit from a variety of methodological approaches; incorporating individual

---

<sup>8</sup> We thank Mike Page for making the latter point.

differences, experimental and neuroscientific approaches. Finally, important insights may be gained by building upon effective applications of mental simulation in the domains of psychotherapy and neuropsychology (Holmes, Arntz & Smucker, 2007; Liu et al., 2004).



**References**

Adams, D. M., Stull, A. T., & Hegarty, M. (2014). Effects of mental and manual rotation training on mental and manual rotation performance. *Spatial Cognition & Computation*, *14*(3), 169-198. doi: 10.1080/13875868.2014.913050

Adriaanse, M. A., Oettingen, G., Gollwitzer, P. M., Hennes, E. P., De Ridder, D. T., & De Wit, J. B. (2010). When planning is not enough: Fighting unhealthy snacking habits by mental contrasting with implementation intentions (MCII). *European Journal of Social Psychology*, *40*(7), 1277-1293. doi: 10.1002/ejsp.730

Adriaanse, M. A., Vinkers, C. D. W., De Ridder, D. T. D., Hox, J. J., & De Wit, J. B. F. (2011). Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. *Appetite*, *56*(1), 183–193. doi: 10.1016/j.appet.2010.10.012

Alden, A. L., Dale, J. A., & DeGood, D. E. (2001). Interactive effects of the affect quality and directional focus of mental imagery on pain analgesia. *Applied Psychophysiology and Biofeedback*, *26*(2), 117-126. doi: 10.1023/A:1011387122804

Allami, N., Paulignan, Y., Brovelli, A., & Boussaoud, D. (2008). Visuo-motor learning with combination of different rates of motor imagery and physical practice. *Experimental Brain Research*, *184* (1), 105-113. doi: 10.1007/s00221-007-1086-x

Andersson, E. K., & Moss, T. P. (2011). Imagery and implementation intention: A randomised controlled trial of interventions to increase exercise behaviour in the general population. *Psychology of Sport and Exercise*, 12(2), 63-70. doi: 10.1016/j.psychsport.2010.07.004

Andre, J. C., & Means, J. R. (1986). Rate of imagery in mental practice: An experimental investigation. *Journal of Sport Psychology*, 8(2), 124-128.

Armitage, C. J., & Reidy, J. G. (2008). Use of mental simulations to change theory of planned behaviour variables. *British Journal of Health Psychology*, 13(3), 513-524. doi: 10.1348/135910707X227088

Arora, S., Aggarwal, R., Sirimanna, P., Moran, A., Grantcharov, T., Kneebone, R., ... & Darzi, A. (2011). Mental practice enhances surgical technical skills: a randomized controlled study. *Annals of surgery*, 253(2), 265-270. doi: 10.1097/SLA.0b013e318207a789

Atance, C. M. & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Sciences*, 5 (12):533-539. doi: 10.1016/S1364-6613(00)01804-0

Austenfeld, J. L., Paolo, A. M., & Stanton, A. L. (2006). Effects of writing about emotions versus goals on psychological and physical health among third-year medical students. *Journal of Personality*, 74(1), 267-286. doi: 10.1111/j.1467-6494.2005.00375.x

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211. doi: 10.1016/0749-5978(91)90020-T

Bachman, K. (1990). Using mental imagery to practice a specific psychomotor skill. *The Journal of Continuing Education*, *21*, 125-128. doi: 10.3928/0022-0124-19900501-10

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173. doi: 10.1037/0022-3514.51.6.1173

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baumeister, R.F., Vohs, K.D., & Oettingen, G. (2016). Pragmatic prospection: How and why people think about the future. *Review of General Psychology*, *20*, 3-16. doi: 10.1037/gpr0000060

Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thoughts cause behavior?. *Annual review of psychology*, *62*, 331-361. doi : <https://doi.org/10.1146/annurev.psych.093008.131126>

Berntsen, D., & Bohn, A. (2010). Remembering and forecasting: The relation. *Memory & Cognition* **38**, 265–278. doi: <https://doi.org/10.3758/MC.38.3.265>

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The Pervasive Problem With Placebos in Psychology: Why Active Control Groups Are Not Sufficient to Rule Out Placebo

Effects. *Perspectives on Psychological Science*, 8(4), 445–454.

<https://doi.org/10.1177/1745691613491271>

Brase, G. L. (2009). How different types of participant payments alter task performance. *Judgment and Decision Making*, 4(5), 419-428.

Brown, E. M., Smith, D. M., Epton, T., & Armitage, C. J. (2018). Do Self-Incentives and Self-Rewards Change Behavior? A Systematic Review and Meta-Analysis. *Behavior Therapy*, 49(1), 113–123. DOI: <https://doi.org/10.1016/j.beth.2017.09.004>

Budney, A. J., & Woolfolk, R. L. (1990). Using the wrong image: An exploration of the adverse effects of imagery on motor performance. *Journal of Mental Imagery*. 14(3-4), 75-86.

Callow, N., Roberts, R., Hardy, L., Jiang, D. & Edwards, M. G. (2013). Performance improvements from imagery: evidence that internal visual imagery is superior to external visual imagery for slalom performance. *Frontiers in Human Neuroscience*, 7, 1-10. doi: 10.3389/fnhum.2013.00697

Callow, N., Jiang, D., Roberts, R., & Edwards, M. G. (2017). Kinesthetic Imagery Provides Additive Benefits to Internal Visual Imagery on Slalom Task Performance, *Journal of Sport and Exercise Psychology*, 39(1), 81-86. DOI: <https://doi.org/10.1123/jsep.2016-0168>

Carden, L. & Wood, W. (2018). Habit formation and change. *Current Opinion in Behavioral Sciences*, 20, 117-122 doi: 10.1016/j.cobeha.2017.12.009

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354. doi: 10.1037/0033-2909.132.3.354

Chambers, M. J. (1990). Forecasting with demand systems: A comparative study. *Journal of Econometrics*, *44*(3), 363-376. doi: 10.1016/0304-4076(90)90064-Z

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2<sup>nd</sup> Edition. London: Academic Press. DOI: 10.1126/science.169.3941.167

XXXDATAXXX (2020, January 21). Mental Simulation Meta-Analysis. Retrieved from XXXXXX anonymised for blind review [data OSF link]

Conroy, D., Sparks, P., & de Visser, R. (2015). Efficacy of a non-drinking mental simulation intervention for reducing student alcohol consumption. *British journal of health psychology*, *20*(4), 688–707. <https://doi.org/10.1111/bjhp.12133>

Conroy, D., & Hagger, M. S. (2018). Imagery interventions in health behavior: A meta-analysis. *Health Psychology*, *37*(7), 668. doi: 10.1037/hea0000625

Corbin, C.B. (1972). Mental practice. In W.P. Morgan (Ed.), *Ergogenic aids and muscular performance* pp. 93-118). New York: Academic Press.

Creelman, J. (2003). Influence of mental practice on development of voluntary control of a novel motor acquisition task. *Perceptual and Motor Skills*, 97, 319-337. doi: 10.2466/pms.2003.97.1.319

D'Argembeau, A., Renaud, O., & Van der Linden, M. (2011). Frequency, characteristics and functions of Future-oriented thoughts in daily life. *Applied Cognitive Psychology*, 25, 96-103. doi: 10.1002/acp.1647

Debartot U, Clerget E, Olivier E (2011) Role of the Primary Motor Cortex in the Early Boost in Performance Following Mental Imagery Training. *PLoS ONE* 6(10): e26717. doi: 10.1371/journal.pone.0026717

Decety, J. & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, 1079, 4-14. doi: 10.1016/j.brainres.2005.12.115

Driskell, J. E., Copper, C., Moran, A. (1994). Does mental practice enhance performance? *Journal of Applied Psychology*, 79, 481–492. doi: 10.1037/0021-9010.79.4.481

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4-58. doi: 10.1177/1529100612453266

Egger, M., G. Davey Smith, M. Schneider, and C. Minder. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315: 629–34. doi: 10.1136/bmj.315.7109.629

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*, 461–475. doi: 10.1037/0022-3514.70.3.461

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation & Emotion*, *30*, 111–116. doi: <https://doi.org/10.1007/s11031-006-9028-7>

Epstein, M. L. (1980). The relationship of mental imagery and mental rehearsal to performance of a motor task. *Journal of Sport Psychology*, *2*(3), 211-220. doi: 10.1123/jsp.2.3.211

Epstude, K., Scholl, A., & Roese, N. J. (2016). Prefactual thoughts: Mental simulations about what might happen. *Review of General Psychology*, *20*(1), 48. doi: 10.1037/gpr0000064

Feltz, D.L., & Landers, D.M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of Sport Psychology*, *5*, 25-57.

Feltz, D.L., Landers, D.M., & Becker, B.J. (1988). A revised meta-analysis of the mental practice literature on motor skill learning. In D. Druckman & J. Swets (Eds.), *Enhancing human performance: Issues, theories and techniques* pp. 1-65). Washington, DC: National Academy Press.

Fery, Y. (2003). Differentiating visual and kinaesthetic imagery in mental practice. *Canadian Journal of Experimental Psychology*, *57* (1), 1-10. doi: 10.1037/h0087408

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63(3), 665-694. doi: 10.1348/000711010X502733

Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary?. *Psychological methods*, 10(4), 444. doi: 10.1037/1082-989X.10.4.444

Fontani, R., Facchini, A., Casini, M. & Corradeschi, F. (2007). Effect of mental imagery on the development of skilled motor actions. *Perceptual and Motor Skills*. 105, 803-826. doi: 10.2466/pms.105.3.803-826

Fradkin, I., Strauss, A. Y., Pereg, M. & Huppert, J. D. (2018). Rigidly applied rules? Revisiting Inflexibility in Obsessive Compulsive Disorder using Multilevel Meta-Analysis. *Clinical Psychological Science*, 6 (4), 481-505. doi: 10.1177/2167702618756069

Freund, A. M., & Hennecke, M. (2015). On means and ends: The role of goal focus in successful goal pursuit. *Current Directions in Psychological Science*, 24(2), 149–153. doi: 10.1177/0963721414559774

Gamble, B., Moreau, D., Tippett, L. J., & Addis, D. R. (2019a). Specificity of Future Thinking in Depression: A Meta-Analysis. *Perspectives on Psychological Science*, 14(5), 816–834. <https://doi.org/10.1177/1745691619851784>

Gamble, B., & Tippett, L., & Moreau, D., & Addis, D. (2019b). The Futures We Want: How Goal-Directed Imagination Relates to Mental Health. doi: 10.31234/osf.io/j3n96.



Geoffrion, R., Gebhart, J., Dooley, Y., Bent, A., Dandolu, V., Meeks, R., ... & Robert, M. (2012). The mind's scalpel in surgical education: a randomised controlled trial of mental imagery. *BJOG: An International Journal of Obstetrics & Gynaecology*, *119*(9), 1040-1048. doi: 10.1111/j.1471-0528.2012.03398.x

Giuffrida, A. & Torgerson, D. J. (1997). Should we pay the patient? Review of financial incentives to enhance patient compliance. *British Medical Journal*, *315*, 703-707. doi: 10.1136/bmj.315.7110.703

Grouios, G. (1992). On the reduction of reaction time with mental practice. *International Journal of Sport Behavior*, *15*(2), 141-157.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Gomes, T. V., Ugrinowitsch, H., Marinho, N., Shea, J. B., Raisbeck, L. D., & Benda, R. N. (2014). Effects of mental practice in novice learners in a serial positioning skill acquisition. *Perceptual and motor skills*, *119*(2), 397-414. doi: 10.2466/23.PMS.119c20z4

Hafenbrädl, S., & Woike, J. K. (2018). Competitive escalation and interventions. *Journal of Behavioral Decision Making*, *31*(5), 695–714. <https://doi.org/10.1002/bdm.2084>

Hagger, M.S., Lonsdale, A., & Chatzisarantis, N.L.D. (2011). Effectiveness of a brief intervention using mental simulations in reducing alcohol consumption in corporate

employees. *Psychology, Health & Medicine*, 16(4), 375-392. doi:

10.1080/13548506.2011.554568

Hagger, M. S., Lonsdale, A., Koka, A., Hein, V., Pasi, H., Lintunen, T., & Chatzisarantis, N.L.D. (2012a). An intervention to reduce alcohol consumption in undergraduate students using implementation intentions and mental simulations: A cross-national study.

*International Journal of Behavioral Medicine*, 19, 82-96. doi: 10.1007/s12529-011-9163-8

Hagger, M. S., Lonsdale, A., & Chatzisarantis, N.L.D. (2012b). A theory-based intervention to reduce alcohol drinking in excess of guideline limits among undergraduate students.

*British Journal of Health Psychology*, 17, 18-43. doi: 10.1111/j.2044-8287.2010.02011.x

Hagger, M. S. (September-November, 2016). *Guidelines on changing behaviour using mental imagery and mental simulation for health educators*. Development of guidelines on behaviour change for Health Promotion Division, Ministry of Health, Malaysia funded by the World Health Organization.

Hagger, M. S., Trost, N., Keech, J. J., Chan, D. K. C., & Hamilton, K. (2017). Predicting sugar consumption: Application of an integrated dual-process, dual-phase model.

*Appetite*, 116, 147–156. doi: 10.1016/j.appet.2017.04.032

Hales, S. A., Deeprose, C., Goodwin, G. M., & Holmes, E. A. (2011). Cognitions in bipolar affective disorder and unipolar depression: imagining suicide. *Bipolar disorders*, 13(7-8),

651–661. <https://doi.org/10.1111/j.1399-5618.2011.00954.x>

Hall, C. R., Mack, D. E., Paivio, A., Hausenblas, H. A. (1995). Imagery use by athletes: development of the Sport Imagery Questionnaire. *International Journal of Sport Psychology*, 29 (1), 73-89.

Harkin, B, Webb, TL, Chang, BPI, Prestwich, A., Conner, M., Kellar, I., Yael, B. & Sheeran, P. (2016) Does Monitoring Goal Progress Promote Goal Attainment? A Meta-Analysis of the Experimental Evidence. *Psychological Bulletin*, 142 (2). pp. 198-229.  
doi:10.1037/bul0000025

Hayter, M. A., Bould, M. D., Afsari, M., Riem, N., Chiu, M., & Boet, S. (2012). Does warm-up using mental practice improve crisis resource management performance? A simulation study. *British Journal of Anaesthesia*, 110(2), 299-304. doi: 10.1093/bja/aes351

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.

Higgins, J. P., Thompson, S. G., Deeks, J.J. & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327 (7414), 557-560. doi: 10.1136/bmj.327.7414.557

Higgins J.P.T. & Green, S. (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.  
Available from [www.handbook.cochrane.org](http://www.handbook.cochrane.org)

Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behavior therapy: Images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry*, 38(4), 297–305. doi: 10.1016/j.jbtep.2007.10.00

Holmes, E. A., Blackwell, S. E., Burnett Heyes, S., Renner, F., & Raes, F. (2016). Mental imagery in depression: Phenomenology, potential mechanisms, and treatment implications. *Annual Review of Clinical Psychology*, 12, 249–280. <https://doi.org/10.1146/annurev-clinpsy-021815-092925>

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, 14, 103-109. doi:10.1006/nimg.2001.0832

Jeon, H. J. Kim, J., Ali, A. & Choi, S. (2014). Noise distraction and mental practice in closed and open motor skills. *Perceptual and Motor Skills: Motor Skills & Ergonomics*, 119 (1), 156-168. doi: 10.2466/25.23.PMS.119c14z7

Johannessen, K. B., Oettingen, G., & Mayer, D. (2012). Mental contrasting of a dieting wish improves self-reported health behaviour. *Psychology & Health*, 27(sup2), 43-58. doi: 10.1080/08870446.2011.626038

Johnson, M. K., & Sherman S. J. (1990). Constructing and reconstructing the past and the future in the present. (In: Higgins ET, Sorrentino RM, Eds). *Handbook of motivation and cognition: Foundations of social behavior*. (Vol. 2). (pp. 482–526) Guilford Press: New York.

Jungmann, F., Gockel, I., Hecht, H., Kuhr, K., Räsänen, J., Sihvo, E., & Lang, H. (2011). Impact of perceptual ability and mental imagery training on simulated laparoscopic knot-tying in surgical novices using a Nissen fundoplication model. *Scandinavian Journal of Surgery, 100*(2), 78-85. doi: 10.1177/145749691110000203

Kirk, D., Oettingen, G., & Gollwitzer, P. M. (2013). Promoting integrative bargaining: Mental contrasting with implementation intentions. *International Journal of Conflict Management, 24*(2), 148-165. doi: 10.1108/10444061311316771

Knäuper, B., McCollam, A., Rosen-Brown, A., Lacaille, J., Kelso, E., & Roseman, M. (2011). Fruitful plans: adding targeted mental imagery to implementation intentions increases fruit consumption. *Psychology & health, 26*(5), 601–617.  
<https://doi.org/10.1080/08870441003703218>

Komesu, Y., Urwitz-Lane, R., Ozel, B., Lukban, J., Kahn, M., Muir, T., Fenner, Dee., & Rogers, R. (2009). Does mental imagery prior to cystoscopy make a difference? A randomized controlled trial. *American journal of obstetrics and gynecology, 201*(2), 218-e1. doi: 10.1016/j.ajog.2009.04.008

Kornspan, A. S., Overby, L. Y. & Lerner, B. S. (2004). Analysis and performance of Pre-performance imagery and other strategies on a golf putting task. *Journal of Mental Imagery, 28* (3 & 4), 59-74.

LaCourse, M. G. Turner, J. A. Randolph-Orr, E., Schandler S. & Cohen, M. J. (2004). Cerebral and cerebellar sensorimotor plasticity following motor imagery-based mental

practice of a sequential movement. *Journal of Rehabilitation Research & Development*, 41(4), 505-524. doi: 1938-1352

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5), 372–385. <https://doi.org/10.1037/0003-066X.50.5.372>

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Liu, K. P., Chan, C. C., Lee, T. M. & Hui-Chan, C. W. (2004). Mental imagery for promoting relearning for people after stroke: A randomized controlled trial. *Archives of Physical Medicine Rehabilitation*, 85, 1403-1408. doi: 10.1016/j.apmr.2003.12.035

Littell, J., Corcoran, J., & Pillai, V. (2008). *Introduction. In Systematic Reviews and Meta-Analysis*: Oxford University Press. doi: 10.1177/1049731508318552

Lukaszewski, W. & Jarczewska-Gerc, E. (2012). Mental simulation and persistence in action. *Journal of Russian and East European Psychology*, 50 (3), 26-46. doi: 10.2753/RPO1061-0405500302

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41(9), 954–969. doi: 10.1037/0003-066X.41.9.954

Marszał-Wiśniewska, M. & Jarczewska-Gerc, E. (2016). Role of mental simulation in the weight loss process. *The Journal of Psychology*, 150(1), 1-14. doi: 10.1080/00223980.2014.987102

McNeil, D. G., Spittle, M. & Mesagno, C. (2019). Imagery training for reactive agility: performance improvements for decision time but not overall reactive agility. *International*

*Journal of Sport and Exercise Psychology*. Advance online publication. DOI:

<https://doi.org/10.1080/1612197X.2019.1696866>

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954.

Meleady, R., Hopthrow, T., & Crisp, R. J. (2013). Simulating social dilemmas: Promoting cooperative behaviour through imagined group discussion. *Journal of Personality and Social Psychology*, *104*(5), 839-853. doi: 10.1037/a0031233

Mendoza, D., & Wichman, H. (1978). “Inner” darts: Effects of mental practice on performance of dart throwing. *Perceptual and Motor skills*, *47*(3\_suppl), 1195-1199. doi: 10.2466/pms.1978.47.3f.1195

Meslot, C., Gauchet, A., Allenet, B., François, O., & Hagger, M.S. (2016). Theory-Based Interventions Combining Mental Simulation and Planning Techniques to Improve Physical Activity: Null Results from Two Randomized Controlled Trials. *Frontiers in Psychology*, *7*, 1789, 1-16. doi: 10.3389/fpsyg.2016.01789

Michie, S., Richardson, M, Johnston, M. & Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., Wood, C. E. (2013). The Behavior Change technique taxonomy (v1) of 93 hierarchically structured techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*, 81-95. doi: 10.1007/s12160-013-9486-6

Morin, L., & Latham, G. (2000). The effect of mental practice and goal setting as a transfer of training intervention on supervisors' self-efficacy and communication skills: an exploratory study. *Applied Psychology, 49*(3), 566-578. doi: 10.1111/1464-0597.00032

Munafò, M.R., Nosek, B.A., Bishop, D.V., Button, K.S., Chambers, C.D., Sert, N.P., Simonsohn, U., Wagenmakers, E.M., Ware, J.J., & Ioannidis, J.P. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*. doi: 10.1038/s41562-016-0021

Murru, E. C., & Martin Ginis, K. A. (2010). Imagining the possibilities: the effects of a possible selves intervention on self-regulatory efficacy and exercise behavior. *Journal of sport & exercise psychology, 32*(4), 537–554. <https://doi.org/10.1123/jsep.32.4.537>

Nicholls, A. R. (2017). *Psychology in sports coaching: theory and practice*. New York; Routledge. doi: 10.1260/1747-9541.8.1.255

O, J. & Munroe-Chandler, K. J. (2008). The effects of image speed on the performance of a soccer task. *Applied Research, 22*, 1-17. doi: 10.1123/tsp.22.1.1

Oettingen, G., & Wadden, T. A. (1991). Expectation, fantasy, and weight loss: Is the impact of positive thinking always positive? *Cognitive Therapy and Research, 15*, 167–175. doi: 10.1007/BF01173206

Oettingen, G., & Mayer, D. (2002). The motivating function of thinking about the future: Expectations versus fantasies. *Journal of Personality and Social Psychology, 83*, 1198–1212. doi: 10.1037/0022-3514.83.5.1198



Oettingen, G. (2012). Future thought and behaviour change. *European Review of Social Psychology*, 23 (1), 1-63. doi: 10.1080/10463283.2011.643698

Oh, H. J. & Larose, R. (2015). Tell Me a Story About Healthy Snacking and I Will Follow: Comparing the Effectiveness of Self-Generated Versus Message-Aided Implementation Intentions on Promoting Healthy Snacking Habits Among College Students, *Health Communication*, 30(10), 962-974. doi: 10.1080/10410236.2014.910289

Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology*, 91, 188-204. doi: 10.1037/0022-3514.91.1.188

Overdorf, V., Page, S. J., Schweighardt R, McGrath RE. (2004). Mental and physical practice schedules in acquisition and retention of novel timing skills. *Perceptual and Motor Skills*, 99, 51-62. doi: 10.2466/pms.99.1.51-62

Overton, R. C. 1998. A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods* 3 (3): 354–79. doi: 10.1037/1082-989X.3.3.354

Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., & Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *Journal of Neurophysiology*, 74(3), 1037–1045. <https://doi.org/10.1152/jn.1995.74.3.1037>

Peluso, E. A., Ross, M. J., Gfeller, J. D., & LaVoie, D. J. (2005). A comparison of mental strategies during athletic skills performance. *Journal of sports science & medicine*, 4(4), 543.

Pham, L.B., & Taylor, S.E. (1999). From thought to action: Effects of process- versus outcome- based mental simulations on performance. *Personality and Social Psychology Bulletin*, 26, 250-260. doi: 10.1177/0146167299025002010

Post, P.G., Williams, C., Simpson, D., & Berning, J.M. (2015). The Effects of a PETTLEP Imagery Intervention on the Learning of a Complex Motor Skill. *Journal of Imagery Research in Sport and Physical Activity*, 10, 19 - 30.

Raison, N., Ahmed, K., Abe, T., Brunckhorst, O., Novara, G., Buffi, N., McIlhenny, C., van der Poel, H., van Hemelrijck, M., Gavazzi, A., & Dasgupta, P. (2018). Cognitive training for technical and non-technical skills in robotic surgery: a randomised controlled trial. *BJU international*, 122(6), 1075–1081. DOI: 10.1111/bju.14376

Rawlings, E. I., Rawlings, I. L., Chen, S. S., & Yilk, M. D. (1972). The facilitating effects of mental rehearsal in the acquisition of rotary pursuit tracking. *Psychonomic Science*, 26(2), 71-73. doi: 10.3758/BF03335435

Renner, F., Murphy, F. C., Ji, J. L., Manly, T., & Holmes, E. A. (2019). Mental imagery as a “motivational amplifier” to promote activities. *Behaviour Research and Therapy*, 114, 51-59. DOI: <https://doi.org/10.1016/j.brat.2019.02.002>

Richardson, A. (1967). Mental Practice: A Review and Discussion Part I,

*Research Quarterly. American Association for Health, Physical Education and Recreation*, 38 (1), 95-107. doi: 10.1080/10671188.1967.10614808

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin Vol 86*(3), 638-641. doi: 10.1037/0033-2909.86.3.638

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education*, 33(4), 221-230. doi: 10.2307/3345249

Rozand, V., Lebon, F., Stapley, P. J., Papaxanthis, C., & Lepers, R. (2016). A prolonged motor imagery session alter imagined and actual movement durations: Potential implications for neurorehabilitation. *Behavioural brain research*, 297, 67–75.  
<https://doi.org/10.1016/j.bbr.2015.09.036>

Ruffino, C., Papaxanthis, C., & Lebon, F. (2017). The influence of imagery capacity in motor performance improvement. *Experimental brain research*, 235(10), 3049–3057. DOI: <https://doi.org/10.1007/s00221-017-5039-8>

Ryan, E. D. Blakeslee, T. & Furst, D. M. (1986). Mental practice and motor skill learning: An indirect test of the neuromuscular feedback hypothesis. *International Journal of Sports Psychology*, 17 (1), 60-70.

Saimpont, A., Lafleur, M. F., Malouin, F., Richards, C. L., Doyon, J., & Jackson, P. L. (2013). The comparison between motor imagery and verbal rehearsal on the learning of

sequential movements. *Frontiers in Human Neuroscience*, 7, 773. doi:

10.3389/fnhum.2013.00773

Sanders, C. W., Sadoski, M., Walsum, K, Bramson, Wiprud, R. & Fossum, T. W. (2008).

Learning basic surgical skills with mental imagery: using the simulation centre in the mind.

*Medical Education*, 42, 607-612. Doi: 10.1111/j.1365-2923.2007.02964.x

Seligman, M. E. P., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, 8, 119-141. doi:

10.1177/174569161247317.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory:

remembering the past and imagining the future. *Philosophical Transactions of the Royal*

*Society B*, 362, 773–786. doi: 10.1098/rstb.2007.2087

Schacter, D.L., Addis, D. R. ,Hassabis,D., Martin,V.C., Spreng,R. N., and Szpunar,K. K.

(2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, 76, 677–

694. doi: 10.1016/j.neuron.2012.11.001

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American*

*Psychologist*, 67(8), 603–613. <https://doi.org/10.1037/a0029869>

Seif-Barghi, T., Kordi, R., Memari, A. H., Mansournia, M. A., & Jalali-Ghomi, M. (2012).

The effect of an ecological imagery program on soccer performance of elite players. *Asian Journal of Sports Medicine*, 3(2), 81.

Shackell, E. M., & Standing, L. G. (2007). Mind Over Matter: Mental Training Increases Physical Strength. *North American Journal of Psychology*, 9(1).

Shanks, D. R., & Cameron, A. (2000). The effect of mental practice on performance in a sequential reaction time task. *Journal of Motor Behavior*, 32(3), 305-313. doi:

10.1080/00222890009601381

Sheeran, P., Harris, P., Vaughan, J., Oettingen, G., & Gollwitzer, P. M. (2013). Gone exercising: Mental contrasting promotes physical activity among overweight, middle-aged, low-SES fishermen. *Health Psychology*, 32(7), 802. doi: 10.1037/a0029293

Skoura, X., Vinter, A., & Papaxanthis, C. (2009). Mentally simulated motor actions in children. *Developmental Neuropsychology*, 34(3), 356-367. doi:

10.1080/87565640902801874

Smith, D., Collins, D., & Holmes, P. (2003). Impact and mechanism of mental practice effects on strength. *International Journal of Sport and Exercise Psychology*, 1(3), 293-306.

doi: 10.1080/1612197X.2003.9671720

Smyth, M. M. (1975). The role of mental practice in skill acquisition. *Journal of*

*Motor Behavior*, 7(3), 199-205. doi: 10.1080/00222895.1975.10735034

Stadler G., Oettingen G., Gollwitzer P. M. (2010). Intervention effects of information and self-regulation on eating fruits and vegetables over two years. *Health Psychology, 29* 274–283. doi: 10.1037/a0018644

Strachan, S. M., Marcotte, M. M. E. Giller, T. M. T., Brunet, J. & Schellenberg, B. J. I. (2017). An online intervention to increase physical activity: Self-regulatory possible selves and the moderating role of task self-efficacy. *Psychology of Sport and Exercise, 31*, 158-165. doi: 10.1016/j.psychsport.2016.05.001

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247. doi: 10.1207/s15327957pspr0803\_1

Suurmond R, van Rhee, H, Hak T. (2017). Introduction, comparison and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*. 1-17. doi.org/10.1002/jrsm.1260. doi: 10.1002/jrsm.1260

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67-85.

Taylor, S. E., & Pham, L. B. (1996). Mental simulation, motivation, and action. In P.M. Gollwitzer & J.A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 219-235). New York: Guilford Press.

Taylor, S.E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination. Mental simulation, self-regulation, and coping. *American Psychologist*, *53*(4):429-39. doi: 10.1037//0003-066X.53.4.429

Theeuwes, M., Liefoghe, B., De Schryver, M., & De Houwer, J. (2018). The role of motor imagery in learning via instructions. *Acta psychologica*, *184*, 110–123.  
<https://doi.org/10.1016/j.actpsy.2017.05.002>

Toli, A., Webb, T. L., & Hardy, G. E. (2016). Does forming implementation intentions help people with mental health problems to achieve goals? A meta-analysis of experimental studies with clinical and analogue samples. *The British journal of clinical psychology*, *55*(1), 69–90. <https://doi.org/10.1111/bjc.12086>

Van Rhee, H.J., Suurmond, R., & Hak, T. (2015). User manual for *Meta-Essentials: Workbooks for meta-analysis (Version 1.2)* Rotterdam, The Netherlands: Erasmus Research Institute of Management. Retrieved from [www.irim.eur.nl/research-support/meta-essentials](http://www.irim.eur.nl/research-support/meta-essentials)

Vasque, N, A., & Buehler, R. (2007). Seeing future success: Does imagery perspective influence achievement motivation? *Personality and Social Psychology Bulletin*, *33*, 1392-1405. doi: 10.1177/0146167207304541

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software* *36* (3).

Vogt, S. (1995). On relations between perceiving, imagining and performing in the learning of cyclical movement sequences. *British journal of Psychology*, 86(2), 191-216. doi: 10.1111/j.2044-8295.1995.tb02556.x

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132, 249–268. doi: 10.1037/0033-2909.132.2.249

Wohldmann, E. L., Healy, A. F., & Bourne Jr, L. E. (2007). Pushing the limits of imagination: mental practice for learning sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 254. doi: 10.1037/0278-7393.33.1.254

Wollman, N., Hill, J., & Lipsitz, T. (1985). Effects of imagery on track and bowling performance in naturalistic settings. *Perceptual and Motor Skills*, 60 (3), 986-986. doi: 10.2466/pms.1985.60.3.986

Woolfolk, R. L., Murphy, S. M., Gottesfeld, D., & Aitken, D. (1985). Effects of mental rehearsal of task motor activity and mental depiction of task outcome on motor skill performance. *Journal of Sport Psychology*, 7(2), 191-197. doi: 10.1123/jsp.7.2.191

Yue, G., & Cole, K. J. (1992). Strength increases from the motor program: comparison of training with maximal voluntary and imagined muscle contractions. *Journal of Neurophysiology*, 67(5), 1114-1123. doi: 10.1152/jn.1992.67.5.1114



Zapparoli, L., Invernizzi, P., Gandola, M., Verardi, M., Berlingeri, M., Sberna, M., ...Paulesu, E. (2013). Mental images across the adult lifespan: a behavioural and fMRI investigation of motor execution and motor imagery. *Experimental Brain Research*, 224(4), 519+. doi: 10.1007/s00221-012-3331-1

Zecker, S. G. (1982). Mental practice and knowledge of results in the learning of a perceptual motor skill. *Journal of Sport Psychology*, 4(1), 52-63. doi: 10.1123/jsp.4.1.52

Zinatelli, M. & Vogel-Sprott, M. (1990). Learned tolerance to alcohol: Mental rehearsal with imagined consequences. *Alcoholism: clinical and experimental*, 14, 518-521. doi: 10.1111/j.1530-0277.1990.tb01191.x

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, 41, doi: 10.1017/S0140525X17001972. Advance online publication

**Figure Captions**

**Figure 1** Flow Diagram of Systematic Search Results

**Figure 2** Main Funnel Plot for all included control-mental simulation comparisons ( $k = 123$ )

**Figure 3** Forest Plot showing all 123 effect sizes included in the meta-analysis.

Note for Figure 3: Each effect size (Hedge's  $g$ ) is represented with a square, and the overall effect size presented as a diamond. The vertical dotted line represents zero effect (mental simulation versus control) and all error bars represent 95% confidence intervals. All means and confidence interval values are also presented on the right-hand side of each effect.

Multiple types of mental simulation and dependent variables are shown independently, even if they are from within the same study. Study authors and years are presented on the left-hand side and the type of mental simulation is represented by letters A-H (see Table 2 for descriptions of each simulation).

MENTAL SIMULATION META-ANALYSIS

**Supplementary Material**

Characteristics of Included Studies, Including Potential Moderators

Study	N	Age	Sex % ♀	Sample + ethnicity	Incentive	Outcome	MS duration (mins)	MS frequency	Delay (days)	Behavior	Domain
Adams_2014_Exp1	C=20, MS <sup>c</sup> =25	NR	NR	University students	Course credit/\$10	Speed	1	Once	0	Manual rotation of shapes	Motor skills
Adriaanse_2010_Exp 2	C =20, MS <sup>h</sup> =19	19.4 YA	100	University students	\$10	Frequency	3	X 7	7	Eating habits (self-report)	Health
Alden_2001	C=10, MS <sup>d</sup> =10, MS <sup>b</sup> =10	YA	50	University students	NR	Speed	5	Once	0	Pain tolerance of cold pressor	Pain
Allami_2008	C=5, MS <sup>c</sup> =5	28.5 YA	NR	NR	NR	Speed	1	X 120	0	Visuo-motor sequence task	Motor skills
Andersson_2011	C = 15, MS <sup>f</sup> = 15	29 YA	68	White European university students and employees	None	Frequency	5	X 14	14	Physical activity	Health
Andre_1986	C=22, MS <sup>c</sup> =44	20.97 YA	0	University students	Course credit	Frequency	15	X 5	5	Frisbee-throwing	Sports
Arora_2011	C=9, MS <sup>c</sup> =9	22 YA	33	Surgeons	NR	Frequency	30	X 5	0	Surgical skill	Occupational
Austenfield_2006	C=21, MS <sup>f</sup> =21	26.4 YA	45	Students: white (84%), with 2% African American, 2% Latino, 11% Asian, and	\$25-\$100	Frequency	25	X 3	90	Medical visits (before and after intervention in 3 month window); fewer the better	Health

## MENTAL SIMULATION META-ANALYSIS

Baker_1983	C=14, MS <sup>c</sup> =13	YA	NR	2% other ethnicities Counselling students	NR	Frequency	150	Once	7	Counsellor Interview Competence Scale; Responding Skills Proficiency Index; Attending Skills Proficiency Index	Occupational
Budney_1990_Exp1	C=10, MS <sup>d</sup> =10;M S <sup>b</sup> =10	NR	NR	NR	NR	Frequency	1	Once	0	Golf putting	Motor skills
Callow_2013_Exp1	C=15, MS <sup>d</sup> =15	21.5 YA	0	Drivers	NR	Speed	1	Once	0	Driving in a simulator	Motor skills
Callow_2013_Exp3	C=10, MS <sup>c</sup> =10	24.79 YA	23	Recreational skiers	NR	Speed & Frequency	1	10	30	Skiing on an artificial slope	Sports
Callow_2017	C=15 MS <sup>c</sup> =30	21.75 YA	0	University students	NR	Speed	2	X 1	0	Driving on a simulated rally driving circuit	Sports
Conroy_2015	C=60, MS <sup>a</sup> =44, MS <sup>f</sup> =62,M S <sup>h</sup> =45	20.16 YA	74	University students	£25 prize draw	Frequency	3	X 3	28	Weekly Alcohol consumption in Units of alcohol	Health
Creelman_2003	C=13, MS <sup>c</sup> =11	28.8	43	NR	NR	Accuracy	3	X 2	0	Big toe abduction	Motor skills
Debarnot_2011_Exp1	C=8, MS <sup>d</sup> =8	28.0	54	NR	NR	Speed & Frequency	1	X 10	0	Finger tapping sequence	Motor skills
Epstein_1980	C=15, MS <sup>c</sup> = 30	YA	44	University students	NR	Frequency	3	Once	0	Dart-throwing	Sports
Fery2003_Exp1	C=8, MS <sup>d</sup> =8,	21.9 YA	0	University students	NR	Frequency	1	X 36	2	Reproduction of an image (by drawing)	Motor skills

## MENTAL SIMULATION META-ANALYSIS

Fery2003_Exp2	C=8, MS <sup>d</sup> =8	20.5 YA	0	University students	NR	Frequency	1	X 12	2	Reproduction of an image (using a stylus)	Motor skills
Fontani_2007	C=10, MS <sup>c</sup> =10	35	0	Karate students	NR	Speed & Frequency	1	X 3600	1	Power and strength in karate	Sports
Gentilli_2010_ Exp1	C=10, MS <sup>d</sup> =10	20-24 YA	50	NR	NR	Speed	1	X 60	0	Pointing task	Motor skills
Gentilli_2010_ Exp2	C=5, MS <sup>d</sup> =5	24.3 YA	NR	NR	NR	Speed	1	X 60	1	Pointing task	Motor skills.
Geoffrion_2012	C=26, MS <sup>c</sup> =24	Not report ed	74	Novice surgeons	NR	Frequency	3	X 3	1	Medical operation	Occupational
Gomes_2014	C=12, MS <sup>d</sup> =24	24.7 YA	Mixe d	University students	NR	Frequency	3	X 6	NR	Physical sequencing task (error)	Motor skills
Gould_1980_Exp 2	C=20; MS <sup>d</sup> =20	Not report ed	50	University students	NR	Frequency	1	X 4	0	Strength	Motor skills
Grouios_1992	C=20; MS <sup>c</sup> =20	18-25 YA	0	Competitiv e sports people	NR	Speed	1	X 560	NR	Reaction time task	Motor skills
Hafenbräld_2018 _Exp2	C=51; MS <sup>c</sup> =50	36.0	56	MTurk community sample	Up to 1.20 USD	Accuracy	3	X 1	0	Simulated auction task	Social
Hagger_2011	C=73; MS <sup>f</sup> =86	35.7	39	Corporate employees with varied SES backgrounds and positions	Prize draw	Frequency	3	Once	28	Units alcohol consumed, binge drinking episodes (self- report)	Health
Hagger_2012a	C=168, MS <sup>f</sup> =181	20.2 YA	74	Students (from Estonia, UK and Finland)	Prize draw (£100)	Frequency	3	Once	28	Units alcohol consumed, binge drinking episodes (self- report)	Health

## MENTAL SIMULATION META-ANALYSIS

Hagger_2012b	C=269, MS <sup>f</sup> =169	22.8 YA	56	University students	Prize draw	Frequency	3	Once	28	Units alcohol consumed, binge drinking episodes (self- report)	Health
Halvari_1996	C=23, MS <sup>c</sup> =22	YA	42	Sports class students	NR	Speed & Frequency	13	X 3	0	Sequence of physical exercises	Sports
Hayter_2013	C=21, MS <sup>c</sup> =19	Not report ed	40	Postgradua te anaesthesia trainees	NR	Frequency	20	Once	0	Medical Crisis Management (Ottawa Global Rating Scale)	Occupational
Jeon2014_Exp1	C=12; MS <sup>c</sup> =12	23.3 YA	0	University Students	NR	Frequency	1	X 3	0	Serving in badminton (closed skill) Error	Sports
Jeon2014_Exp2	C=12; MS <sup>c</sup> =12	21.1	50	University Students	NR	Frequency	1	X 3	0	Returning in badminton (open skill) Error	Sports
Johannessen_2012	C=30; MS <sup>h</sup> =36; MS <sup>f</sup> =35	19.6	92	University Students	Course credit	Frequency	3	Once	14	Calorie intake, high calorie foods (self- report)	Health
Jungmann_2011	C=20; MS <sup>c</sup> =20	Not report ed	50	Medical students	NR	speed	3	X 4	Variable (not used in meta- analysis)	Virtual reality surgical performance	Occupational
Kirk_2013	C=23; MS <sup>h</sup> =22	Not report ed	45	University Students	Course credit	Frequency	3	Once	0	Effective negotiation (joint gain) between pairs of individuals	Social
Knauper_2011	C=44; MS <sup>f</sup> =34 MS <sup>a</sup> =37	18.3	62	Students. 71.8% 'White', 10.7%	NR	Frequency	5	X 7	7	Fruit intake (self-report)	Health

## MENTAL SIMULATION META-ANALYSIS

Study	C	MS <sup>c</sup>	Mean	SD	Age	Gender	Design	Frequency	Duration	Control	Task	Outcome
Kohl_1980_Exp1	C=21;	MS <sup>c</sup> =19	21.0	0			Frequency	1	X 25	0	Pursuit Rotor task performance (opposite hand to mental simulation)	Motor skills
Kohl_1980_Exp2	C=20;	MS <sup>c</sup> =20	20.3	0			Frequency	1	X 18	0	Pursuit Rotor task performance (opposite hand to mental simulation)	Motor skills
Kohl_1980_Exp3	C=36;	MS <sup>c</sup> =36	16.9	0	YA		Frequency	1	X 9	0	Pursuit Rotor task performance (opposite hand to mental simulation)	Motor skills
Komesu_2009	C=35;	MS <sup>c</sup> =33	29.0	81		gynaecology residents	Frequency	20	Once	0	Cystoscopy	Occupational
Kornspan_2004	C=10;	MS <sup>d</sup> =10	19.6	22		University students	Frequency	1	X 20	NR	Golf-putting task	Sports
LaCourse_2004	C=12;	MS <sup>c</sup> =12	23.3	54		University students	Frequency	15	X 5	NR	Pointing task	Motor skills.
Lohasz_1997	C=15;	MS <sup>c</sup> =15	19-26	0		University athletes	Speed	1	X 3	0	Pointing task	Motor skills
Lukazewski_2012 Exp1	C=10;	MS <sup>c</sup> =10, MS <sup>f</sup> =10, MS <sup>g</sup> =10	19-26	50		University students	Speed	5	Once	0	Persistence	Motor skills

## MENTAL SIMULATION META-ANALYSIS

Lukazewski_2012 Exp2	C=32; MS <sup>b</sup> =32; MS <sup>d</sup> =32	20.4	53	“different educational backgrounds”(p.34)	NR	Speed	NR	Once	0	Persistence	Motor skills.
Lukazewski_2012 Exp4	C=30; MS <sup>d</sup> =30	26.0	52	“different educational backgrounds”(p.37)	NR	Speed	NR	Once	0	Persistence	Motor skills
Maring_1990	C=13; MS <sup>c</sup> =13	30.0	81	Students and employees	NR	Frequency	2	X 5	0	Throwing task	Motor skills
Marszał- Wiśniewska_2016 Exp2	C=18; MS <sup>h</sup> = 18; MS <sup>a</sup> = 18; MS <sup>f</sup> =17; MS <sup>g</sup> =18; MS <sup>e</sup> = 17	M=29 (SD= 19- 45)	100	Public visiting a dieting internet forum	No incentive provided (except potential weight loss)	Frequency	350	X 35	36	Frequency of log ins to an online weight management programme	Health
McNeil_2019	C=16; MS <sup>d</sup> =16	21.5	100	Athletes	NR	Speed	20	X 6	1	Reactive agility test	Motor Skills
Meleady_2013 Exp1	C=27; MS <sup>d</sup> =27	19.9	75	University Students	Stationary based on task performance	Accuracy	5	Once	0	Cooperation in prisoner's dilemma game	Social
Meleady2013 Exp2	C=24; MS <sup>d</sup> =24	20.9	100	University students	£3, one group received performance-based pay	Frequency	5	Once	0	Cooperation in prisoner's dilemma game	Social
Meleady2013 Exp3	C=27; MS <sup>d</sup> =27	18.5	89	Students (6 <sup>th</sup> form and university)	one group received performance-based pay	Frequency	5	Once	0	Cooperation in prisoner's dilemma game	Social
Meleady2013 Exp4	C=17; MS <sup>d</sup> =34	20.7	84	University students	Course credit	Frequency	5	Once	0	Cooperation in prisoner's dilemma game	Social



## MENTAL SIMULATION META-ANALYSIS

Mendoza1978	C=8; MS <sup>d</sup> =8	19-22	19	University students	NR	Frequency	15	X 6	1	Dart-throwing	Sports
Meslot2016 Exp1	C=30 MS <sup>a</sup> =21	M = 25.1	86	Undergraduate students	NR	Frequency	NR	Once	28	Physical activity	Health
Meslot2016 Exp2	C=18 MS <sup>a</sup> =20	Mean age = 27.44 years	72	Fitness centre members	NR	Frequency	NR	Once	133	Fitness center attendance	Health
Morin2000	C=10; MS <sup>d</sup> =10	46	7	Paper mill employees	NR	Frequency	30	X 4	30	Communication skills	Social
Murru2010	C=15; MS <sup>f</sup> =17; MS <sup>e</sup> =20	21.4	76	Students. White-Caucasian (48%), Asian, (20%), and South Asian (19%)	\$5	Frequency	3	Once	56	Duration of physical exercise (self-report, diary)	Health
O 2008	C = 24; MS = 72 <sup>c</sup>	18.1 YA	68	University student football players	NR	Speed	1	X 3	0	Dribbling a football	Sports
Overdorf2004	C=6; MS <sup>d</sup> =6	29.6	NR	University students	NR	Frequency	3	X 36	0	Errors on a reaction time task	Motor skills
Pascual-Leone1995 Exp2	C=9; MS <sup>c</sup> =9	32.0	40	NR	NR	Speed	120	X 5	0	Piano playing (also error)	Motor skills
Peluso2005	C=30; MS <sup>d</sup> =15	18-25	73	University students	Course credit	Frequency	1	X 10	0	Golf putting performance	Sports
Pham1999	C=21; MS <sup>f</sup> =21; MS <sup>a</sup> =21; MS <sup>h</sup> =21	18-28	73	University students. Anglo,	Course credit	Frequency	5	X 6	1	Examination performance	Occupational

## MENTAL SIMULATION META-ANALYSIS

				44.6%								
				Asian,								
				10.9%								
				Latino,								
				5.9%								
				African								
				American,								
				and 8.9%								
				other.								
Post_2018	C=14; MS <sup>d</sup> =14	20.6 YA	100	University students	NR	Speed	1	X 54	1	Reaction time task	Motor skills	
Raison_2018	C=17; MS <sup>d</sup> =20	23.55 YA	40	Junior doctors/ medical students	NR	Frequency	30	X 1	0	Simulated surgical skills	Occupational	
Rawlings1972 Exp1	C=8; MS <sup>c</sup> =8	NR	100	University students	NR	Frequency	1	X 8	1	Rotary pursuit task	Motor skills	
Rawlings1972 Exp2	C=10; MS <sup>c</sup> =10	NR	0	University students	NR	Frequency	13	X 8	1	Rotary pursuit task	Motor skills	
Renner_2019	C=48; MS <sup>d</sup> =24	36.35	68	Communit y volunteer panel	Monetary payment	Frequency	12	X 15	0	Completing daily activities	Health	
Ross_1985	C=6; MS <sup>c</sup> =6	22.0	3	Music students specializin g in trombone	NR	Frequency	NR	X 3	0	Trombone playing	Motor skills	
Rozand_2016	C = 19; MS <sup>d</sup> = 12	M=24 .3, SD=2 .5	45	NR	NR	Speed	1	X 100	0	Pointing task	Motor skills	
Ruffino_2017	C = 10; MS <sup>d</sup> =10	M = 27	50	NR	NR	Speed	1	X 30	0	Nine Hole Peg Test	Motor Skills	
Ryan_1986	C=20; MS <sup>c</sup> =20	NR	72	University students	NR	Frequency	1	X24	0	Throwing task	Motor skills	

## MENTAL SIMULATION META-ANALYSIS

Saimpont_2013	C=10; MS <sup>c</sup> =10	27.4	NR	NR	NR	Speed	1	X 720	206 days	Motor sequence with left foot	Motor skills.
Sanders_2008	C=31; MS <sup>c</sup> =31	NR	NR	Medical students	NR	Frequency	30	X 2	10	Surgical procedure	Occupational
Seif-Barghi2012	C=17; MS <sup>d</sup> =17	19.4 (13- 32)	0	Elite football players; four age categories, U19, U21, over 21	NR	Frequency	15	X 7	0	Football passing performance	Sports.
Shackell2007	C=10, MS <sup>d</sup> =10	20.1	0	Student football, basketball and rugby players	One protein bar per 15 min session	Frequency	15	X 15	2	Strength (hip flexor task)	Motor skills
Shanks2000	C=10; MS <sup>d</sup> = 10	NR	52	University students	NR	Speed	7	Once	0	Sequential reaction time task	Motor skills
Sheeran2013	C=42; MS <sup>h</sup> =42	53.9	0	Angling club members. Mostly working class 69% employed; 29.80% retired	Prize draw of annual membership (2 available)	Frequency	3	Once	210	Physical activity (self-report)	Health
Smith2003	C=6; MS <sup>d</sup> =6	29.3	0	Staff and postgradua te	NR	Frequency	3	X 8	2	Strength (finger)	Motor skills
Smyth1975 Exp1	C=10; MS <sup>c</sup> =10	NR	41	University Students	No payment	Speed & Frequency	NR	X 5	0	Mirror-drawing task	Motor skills
Smyth1975 Exp2	C=10; MS <sup>c</sup> =10	NR	44	University Students	No payment	Frequency	4	Once	0	Pursuit rotor task	Motor skills

## MENTAL SIMULATION META-ANALYSIS

Strachan 2017	C = 86; MS <sup>f</sup> = 81; MS <sup>h</sup> = 77	29.6	82	Community sample	\$10 gift token	Frequency	5	Once	240	Physical activity (self-report)	Health
Theeuwes_2018_Exp2	C = 20; MS <sup>c</sup> = 20	NR	NR	University students	10 euros cash	Speed & Accuracy	1	X 15	0	Typing task	Motor skills
Vogt1995 Exp1	C=16; MS <sup>c</sup> =16	20-39	50	University Students	25 DM, extra 25 DM for top performer	Frequency	1	X 24	24	Motor task	Motor skills
Vogt1995 Exp2	C=16; MS <sup>c</sup> =16	18-36	50	University Students	20 DM	Frequency	1	X 60	0	Motor task	Motor skills
Wohldmann2007 Exp1	C=20; MS <sup>c</sup> =20	NR	NR	University students	Course credit	Speed	1	X 320	0	Typing number strings	Motor skills
Wohldmann2007 Exp2	C=27; MS <sup>c</sup> =27	NR	NR	University students	Course credit	Speed	1	X 640	90	Typing number strings	Motor skills
Wollman1985 Exp1	C=7; MS <sup>c</sup> =7	NR	0	University cross-country track team	NR	Speed	NR	NR	NR	Running	Sports
Woolfolk1985	C = 8; MS <sup>b</sup> ; MS <sup>c</sup> ; MS <sup>d</sup> ; MS <sup>e</sup> ; MS <sup>c</sup> , (n=8 in each group)	NR	0	University students	Course credit	Frequency	7	X 4	0	Golf putting performance	Sports.
Yue1992	C=10; MS <sup>d</sup> = 10	21-29	NR	NR	NR	Frequency	1	X 300	NR	Finger strength on a laboratory task	Motor skills
Zecker1982	C=10; MS <sup>c</sup> = 10	NR	50	University students	Course credit	Accuracy	1	X 40	0	Throwing to target	Motor skills
Zinatelli1990	C=10; MS <sup>c</sup> =10; MS <sup>d</sup> =10	19-25	0	University students	Cash (\$5 each)	Speed	3	X 8	0	Pursuit Rotor Task	Motor skills.

Note: MS duration = The duration of one trial of mental simulation practice. Delay: Any delay within the same day = 0, 1 = 1 day delay and so on; C=Control Group; MS = mental simulation / experimental group; a = process; b = inferior performance; c = neutral performance; d =

## MENTAL SIMULATION META-ANALYSIS

superior performance; e = inferior outcome; f = superior outcome; g = inferior outcome + process; h = superior outcome + process.; Sex (% females / ♀). PP = physical practice. Delay (Immediate  $\leq$  5 mins). Duration: To give a conservative estimate of MS duration, where low and upper bounds were provided, the maximum was used for analysis. DM = Deutschmark. Similarly, where numbers were not integers we rounded up, and 'a few seconds' was assigned 1 minute. 'A few' was assigned 3 minutes in MS duration and MS frequency analyses. In contrast to other studies, in Lukazewski2012 Exp2 participants had no physical practice/demonstration of the task in question before imagining.