Counterfactual-based nudging and signaling promote more efficient coordination during group tasks

Alexander R. Hough , Kevin O'Neill & Ion Juvina

View supplementary material ⌴

Published online: 11 Jan 2021.

Submit your article to this journal ⌴

Article views: 420

View related articles ⌴

View Crossmark data ⌴

ARTICLE

Counterfactual-based nudging and signaling promote more efficient coordination during group tasks

Alexander R. Hough, Kevin O'Neill and Ion Juvina

Department of Psychology, Wright State University, Dayton, OH, USA

ABSTRACT

Groups of people often find it challenging to coordinate on a single choice or option. Even when coordination is achieved, it may be inefficient because better outcomes were possible. Numerous researchers attempted to address this coordination problem with various manipulations ranging in complexity and generalizability, but results were mixed. Here, we use a more parsimonious and generalizable method – counterfactuals – to nudge (i.e. indirectly guide and allow for free choice) individuals towards choosing options that are more likely to result in efficient coordination. We used a modified version of an existing coordination game, the minimum effort game (MEG), where we added actual effort (i.e. solving an arithmetic problem) and counterfactuals (i.e. statements highlighting the hypothetical outcomes had they or other players chosen differently). Based on previous literature and promising results from a pilot experiment using bidirectional counterfactuals (i.e. both upward and downward), we designed and preregistered a follow-up experiment to directly assess the effectiveness of counterfactuals. We replicated the pilot study with a bidirectional counterfactual condition, then added an upward, downward, and control (no counterfactuals) condition. We found weak evidence for counterfactual nudging and clear evidence that players can effectively nudge the group towards higher efficiency.

Introduction

Coordination is an interdisciplinary topic (e.g. biology, economics, computer science) that can be defined as "managing dependencies between activities" (Malone & Crowston, p. 4). These dependencies can take many forms, such as sharing resources, assigning tasks, and working around constraints, and in humans, coordination often involves emotion, motivation, incentives, and biases at the individual or group level (Malone & Crowston, 1994). Here, we are concerned with group coordination in humans and focus on a specific type where outcomes are a function of behavior at the individual and group level, resulting in unique or asymmetric outcomes for each individual. In this

type of asymmetric coordination situation, coordination is defined as the degree that individuals in a group settle on a single choice out of a set of choices, which can be approximated by calculating variance or the range of the choice distribution within the group (e.g. lower means better coordination). This coordination is "efficient" when individuals of the group coordinate on a choice resulting in good outcomes for all members of the group with respect to the costs and benefits (e.g. a better outcome than some of the other possible choices without negatively affecting or at the expense of another player). The degree of efficiency can be considered as a point on a continuum between the worst and best possible outcomes. Although coordination is possible, coordination failure is common and results from either failure to coordinate on any choice or failure to coordinate on an efficient choice (Camerer, 2003; Cooper et al., 1994; Van Huyck et al., 1990, 1991; R. W. Cooper et al., 1990).

In this paper, we simulate asymmetric coordination where individuals make choices simultaneously or without knowledge of other's choices, they cannot communicate beyond making choices, and the resulting outcomes are asymmetric or unique for each individual depending on their choice and an order statistic (e.g. mean, median, or minimum) based on the choices of the other members of the group. We use a controlled game called the minimum effort game (MEG; Van Huyck et al., 1990) where outcomes are determined by the "weak link" or minimum choice of the group. In the MEG, players typically converge towards the minimum choice resulting in inefficient outcomes for all members of the group. We chose this type of situation because it is very challenging to achieve coordination without an explicit coordination device or salient focal point (Blume et al., 1998; Mehta et al., 1994). In addition, other players' choices, particularly the lowest or weakest choices, can become a focal point and influence other players to coordinate on an inefficient choice (Brandts et al., 2015, 2014; Van Huyck et al., 1990). Several techniques were used for increasing coordination efficiency and some were successful to a degree; however, they involve changing the coordination situation, require substantial effort, and only work when everything is "just right." For instance, Brandts and colleagues (Brandts et al., 2015, 2014) found that allowing a leader to help other players could improve coordination efficiency; however, if this help was taken away too early, it was actually worse than no help at all. In addition, leaders often stopped helping because it incurred a cost that was higher than the benefit.

Our goal here is to evaluate the feasibility of achieving more efficient coordination in the MEG with minimal cost and intervention. If we are able to increase coordination efficiency in this particularly difficult setting, we believe this technique could be applied to some other asymmetric coordination situations as well. In order to achieve this goal, we propose using counterfactual-based nudges to encourage individuals to coordinate more efficiently while playing a MEG. Here, we define a counterfactual as a specific forgone outcome that could have occurred if a different choice was made (e.g. Byrne, 2016; Kahneman & Miller, 1986). We believe counterfactuals, specifically when they highlight a better hypothetical outcome than the current one, can nudge or indirectly guide individuals (Thaler et al., 2013) towards choices that could lead to more efficient coordination. Before discussing the pilot and preregistered study, we review some literature on coordination games and nudging techniques, then further explain how nudging techniques could be used in coordination games.

MEG and coordination

The MEG involves multiple players that simultaneously select a level of effort between 1 and 7, and all player payoffs are determined by individual choice and the lowest effort level or minimum in the group. The payoff matrix (see Table 1) is set up so that coordination on the minimum effort results in higher payoffs, particularly when coordinating at higher levels of effort.

The game structure provides seven coordination points, which are shown in bold and increase in efficiency (i.e. result in higher payoffs) diagonally from left to right (see Table 1). This structure provides two salient choice options: a risk dominant choice (1) that does not depend on other player's choices and always results in a guaranteed payoff (70) and a payoff dominant choice (7) that could result in both the highest payoff (130) or the lowest possible payoff (10). In general, risk increases for individuals as they choose higher effort from 1 to 7 on the y-axis (Table 1) and efficiency (i.e. payoff) increases along with higher minimum effort from 1 to 7 on the x-axis (Table 1). Van Huyck et al. (1990, 1991) demonstrated that players start the game using strategies such as the payoff dominant and risk dominant choices, but learning occurs during the game, which causes deviation from initial strategies. In the first round, players often choose somewhere between these two salient choices, resulting in coordination at an inefficient equilibrium lower than seven (Van Huyck et al., 1991). However, in repeated play, the minimum effort becomes a third salient focal point, since it determines the player's payoffs.

Van Huyck et al. (1990) used three different variations where players were given feedback about the minimum for each round and found players were sensitive to payoffs, group size, and the potential risk or cost of allocating more effort than the minimum. In the original variation described above, players quickly converge to one (risk dominant) over ten rounds. In a second treatment, payoffs were based solely on the minimum regardless of the individual's choice (diagonal in Table 1). There was no cost of choosing higher than the minimum and players converged to seven (i.e. payoff dominant). In a third condition, group size was reduced to two and players converged towards seven (i.e. payoff dominant). There is a potential issue with these variations because players were only given information about the minimum and were not able to determine if other players were choosing higher. To address this issue, Van Huyck et al. (1990) included a condition where players were given complete outcome information (i.e. the distribution of choices and the minimum); however, this had little effect and players still converged to one. This finding was later replicated and reported in Camerer and Ho (1998).

Table 1. MEG payoff matrix (Adapted from Van Huyck et al., 1991).

		Minimum Effort Choice in Group						
		1	2	3	4	5	6	7
Player's Effort Choice	1	70						
	2	60	80					
	3	50	70	90				
	4	40	60	80	100			
	5	30	50	70	90	110		
	6	20	40	60	80	100	120	
	7	10	30	50	70	90	110	130

Furthermore, Leng et al. (2018) found additional outcome information and continuous time treatments have no significant effects on coordination behaviors beyond only giving the minimum as feedback in discrete time. Devetag and Ortmann (2007) conducted an extensive literature review and found several other factors that affect coordination and its efficiency in several different contexts. For instance, the following manipulations can increase coordination efficiency: lowering the benefits of the risk dominant choice (i.e. secure option) relative to the riskier but more rewarding payoff dominant choice (Brandts & Cooper 2006), introducing pre-play communication that is costly (Van Huyck et al., 1993) or costless (Cooper et al., 1992), reducing ambiguity regarding potential losses or lower rewards based on choices involving risk (Cachon & Camerer, 1996; Rydval & Ortmann, 2005), and lowering the cost of experimentation or exploration of possible choices by increasing the number of rounds or modifying the payoff structure (Berninghaus & Ehrhart, 1998; Van Huyck et al., 2007).

Although coordination failure often occurs due to the lack of a coordinating device (Blume et al., 1998; Mehta et al., 1994) and the strong influence of other player's actions (Brandts et al., 2015, 2014), the coordination game literature provides three additional points. First, players' initial choice is very important since it is very difficult to improve efficiency once it has already converged on an inefficient equilibrium (Chaudhuri et al., 2009; Van Huyck et al., 1990). Typically, groups of players have a higher dispersion of first-round choices compared to later rounds (Camerer & Ho, 1998; Van Huyck et al., 1990), suggesting individual differences in strategies or selection principles. Players might use previous strategies that may not be appropriate to the current context and, even if they adapt to the current environment (Cooper & Van Huyck, 2018), that previous strategy or rule likely serves as a reference point (Costa-Gomes et al., 2009). General individual traits may also influence coordination behavior. Cachon and Camerer (1996) found some players display different degrees of loss avoidance, and Devetag and Ortmann (2007) suggested that both risk and trust are important components in coordination games, which are often not directly addressed.

Second, people may prefer to reduce the cognitive effort when possible and under-value the benefits of applying effort in a given situation. For instance, Westbrook et al. (2013) found many participants were willing to forgo higher monetary rewards to reduce cognitive effort (i.e. effort discounting) when given the option to choose the difficulty level of a task that corresponded to monetary value (e.g. more difficult tasks pay higher). This study and others (e.g. Juvina et al., 2018; Kool et al., 2010) suggest many people perceive the subjective cost of effort as higher and therefore undervalue the benefits associated with putting in effort. This "effort avoidance" was previously found to nega-tively correlate with personality traits that may have an effect on performance, such as attentional control, self-control, need for cognition, and abstract reasoning ability (Juvina et al., 2018; Kool et al., 2010). This research regarding effort avoidance is similar to steps of reasoning in game theory. Consider a penalty shot in the game of soccer. A goalie wants to correctly predict the location of the shot. In one step of reasoning, the goalie decides to dive to the left, because the shooter is left footed and this location is more likely. An additional step of reasoning may involve the goalie thinking that the shooter is thinking of shooting towards the other side in order to be less predictable. Based on this second step of reasoning, the goalie now decides to dive right. Engaging in multiple steps of reasoning can be taxing on memory and may be perceived as aversive or costly (Beard &

Beil, 1994; Duffy & Nagel, 1997; Ho et al., 1998; McKelvey & Palfrey, 1992; Nagel, 1995; Rubinstein, 1989; Schotter et al., 1994; Van Huyck et al., 2002). Therefore, individuals may prefer to take few steps of reasoning similar to the effort discounting described earlier (e.g. Juvina et al., 2018; Kool et al., 2010; Westbrook et al., 2013). In addition, one step of reasoning is cost-efficient and is often a good strategy as other player's behavior is difficult to predict and people are sensitive to the wasted effort (Camerer, 2003; Haruvy & Stahl, 2007; Ho & Weigelt, 1996). In the repeated play, players might apply little effort towards initial choices, in order to size up the group before applying too much effort.

Lastly, players might signal their willingness to take a risk and choose higher effort than other players in order to increase coordination efficiency over time, but this may depend on the effectiveness of signaling in that context (Charness et al., 2018). This is similar to Brandts et al.'s (2014, 2015) experiments where leaders can either choose first or help other players in order to improve coordination efficiency once it has already converged on an inefficient option; however, their experiments utilized communication or sequential choice. Signaling behavior involves a risky cost/benefit tradeoff that relates to intertemporal choice (e.g. Shenhav et al., 2017). There are short-term costs (e.g. lower payoffs) that may or may not lead to long-term benefits (e.g. improved coordination efficiency leading to higher payoffs). Players might focus on or be more sensitive to the short-term cost of lower payoffs and may perceive the delayed benefit as less attractive (e.g. Shenhav et al., 2017; E. U. Weber et al., 2007). In addition, players might see the risk in choosing higher effort than other players (e.g. Cachon & Camerer, 1996) and might negatively reciprocate when other player(s) choose a lower level of effort than they do (e.g. Offerman, 2002).

Although coordination failure is common, individuals often do coordinate over time, but have difficulty coordinating on efficient or high rewarding options (Camerer, 2003; Cooper et al., 1994; Riechmann & Weimann, 2008; Van Huyck et al., 1990; R. W. Cooper et al., 1990). Once coordination is achieved and stabilized, often on an inefficient option, it is very difficult to increase efficiency above that baseline (Brandts & Cooper, 2006; Brandts et al., 2015, 2014; Chaudhuri et al., 2009; Van Huyck et al., 1991). This is because it takes effort for players to settle on an equilibrium and bringing up efficiency would require a destabilization that breaks the coordination and additional effort is then required to settle on a new equilibrium. Efficiency may not improve enough for the benefits to offset the cost. For example, imagine a group of people spending time coordinating a social event and compromise on an available venue that is not very desirable; then a more desirable location becomes available. This situation involves "undoing" coordination and then coordinating a second time, which requires more effort that may or may not improve the level of efficiency achieved the first time. Since an individual's subjective cost of effort likely increases over time (e.g. Westbrook et al., 2013) they may now perceive the cost of effort as higher than the potential benefits of re-coordinating. Several techniques have been applied to this "re-coordination" problem with varying degrees of cost, effort, and effectiveness (Brandts et al., 2015, 2014; Sahin et al., 2015; Van Huyck et al., 1990, 1993, 1992; R. W. Cooper et al., 1990; R. Weber et al., 2001).

Nudging

Since our interest here is on parsimonious methods for improving coordination efficiency, we prefer to do as little as possible and allow individuals to make their own choices. As

Thaler et al. (2013) suggest, providing a nudge and giving people the option to choose is frequently more effective than giving more direct suggestions or mandating choices, as they are often perceived as aversive. In the MEG, payoffs are determined by the behavior of the group, which cannot be known ahead of time. In addition, providing direct suggestions would confound the dynamics of group coordination behavior and the generalizability to other coordination situations. Therefore, we propose providing nudges (i.e., information to indirectly guide individuals towards choices that could lead to better outcomes) in the form of counterfactuals (e.g. you could have earned the higher payoff of X if you had chosen the effort level of Y instead of Z). Previous research with the MEG suggests that just providing outcome information about the minimum and the distribution of player choices do not appear to be enough for players to coordinate efficiently (Camerer & Ho, 1998; Leng et al., 2018; Van Huyck et al., 1990). Therefore, we propose a novel approach by adding counterfactuals to outcome information to nudge individuals toward more efficient coordination. To the best of our knowledge, no MEG or coordination-focused study has used counterfactuals. The closest study we are aware of is Kray and Galinsky (2003) counterfactual priming experiment with group decision-making.

Counterfactuals

Counterfactual thinking involves considering forgone outcomes (Byrne, 2016; Kahneman & Miller, 1986), which is more likely after failures or shortcomings (Gilovich, 1983; Hur, 2001; Roese & Hur, 1997; Roese & Olson, 1997; Sanna & Turley, 1996; Sanna & Turley-Ames, 2000) and often involve correcting or improving upon previous behaviors (Markman et al., 1993; Roese, 1997; Roese et al., 1999). Epstude and Roese (2008) suggest that this may be dependent on the realization that there is a problem or goals are not sufficiently met, which is often signaled by negative affect (e.g. Lieberman et al., 2002; Schwarz, 1990; Schwarz & Clore, 1983; Taylor, 1991). Counterfactual thinking is typically targeted towards things that are easier to change, such as one's own choice compared to another person's choice (Kahneman & Miller, 1986). Situations that are harder to change or have less malleability may not benefit as much from counterfactuals (Smallman & Summerville, 2018), such as changing one's own behavior versus trying to change another person's behavior. They appear most helpful for events that repeat and there is an opportunity to apply changes in the future (Smallman & Summerville, 2018).

　　Upward counterfactuals involve identifying alternative actions that could have led to more positive outcomes, are often associated with negative affect (e.g. regret), and are more likely when individuals are trying to improve performance (Rim & Summerville, 2014; Roese, 1994; Roese, 1997; White & Lehman, 2005). Behaviors to improve future outcomes might be achieved through goal-oriented reasoning (Epstude & Roese, 2008; Roese & Epstude, 2017) or by increases in motivation, persistence, and performance (Dyczewski & Markman, 2012; Markman et al., 2008). Upward counterfactuals can increase motivation or effort (Markman & McMullen, 2003; Markman et al., 2008), particularly when improvement is believed to be possible (Dyczewski & Markman, 2012), and improve future performance (Morris & Moore, 2000; Nasco & Marsh, 1999). When change is less attainable, upward counterfactuals could still help to improve performance, for instance, by encouraging players to pay a temporary cost (e.g. payoff and risk) and signal the willingness to choose higher effort.

On the other hand, downward counterfactuals focus on how situations or outcomes could have been worse (Rim & Summerville, 2014), and can often lead to rationalization rather than improvement of future performance (Roese, 1994; Smallman & Summerville, 2018). However, they can be more beneficial for maintaining performance when improvement is or is believed to be less attainable (Dyczewski & Markman, 2012). To make things more complex, Markman and McMullen (2003) suggest that motivation to improve performance is a function of both counterfactual direction and whether they involve reflective or evaluative thinking. For instance, after getting a B on a test, an individual could generate an upward reflective (I almost got an A) that is associated with positive affect or an upward evaluative (I failed to get an A) counterfactual that likely generates negative affect. According to the "affect as information" theory (Schwarz & Clore, 1983), negative affect tends to be more motivating. That being said, both upward and downward counterfactuals could have motivating effects on performance improvements. There are, however, some results regarding individual problem-solving tasks that suggest the benefits of counterfactuals are not dependent on their direction or emotion elicitation, but rather whether they influence mental simulation or consideration of other possible choices (Galinksy & Moskowitz, 2000; Galinksy et al., 2000).

According to the counterfactual literature, upward counterfactuals in the MEG (e.g. you could have earned a higher payoff had you chosen X) should be more motivating than downward (e.g. you could have earned a lower payoff had you chosen X), especially when outcomes can be easily changed (e.g. changing one's own behavior). Therefore, in the MEG, we expect upward counterfactuals (compared to downward) to encourage individuals to choose higher effort, earn a higher payoff, and coordinate more efficiency if they and others choose higher levels of effort. In the context of the MEG, counterfactual thinking is constrained to considering what could have happened if one chose differently or the minimum was different and is therefore similar to self or team reflection. However, the way the counterfactual information is presented in our setup can trigger both reflective and evaluative thinking (see Markman et al., 2008): highlighting specific alternative choices may facilitate reflection, whereas the differences in payoff between alternatives may trigger evaluative thinking. We only provide a set of counterfactuals and allow individuals to engage in either reflective or evaluative processing and use any self-regulatory strategies they may be prone to (e.g. promotion- vs. prevention-focused self-regulation; Florack et al., 2013). We found some support for the idea that counterfactuals could improve performance in the MEG from Kray and Galinsky (2003). They found a counterfactual priming manipulation that involves generating counterfactuals prior to a decision-making task increased group decision-making and there was a positive relationship between the number of counterfactuals generated and group decision-making accuracy.

Pilot study

There were three goals for the pilot study: 1) test our novel implementation of the MEG that includes actual effort and counterfactuals, 2) assess the effectiveness of counterfactuals to nudge individuals towards more efficient coordination and compare to previous MEG findings, and 3) assess the relationships with coordination behavior and effort preferences. The rationale, method, and results from the pilot study provides justification

for the preregistered experiment. Both experiments share the same rationale and methods, however, the pilot study used bidirectional counterfactuals and the preregistered experiment used separate counterfactual treatments (i.e. bidirectional, upward, downward, and a no counterfactual control).

Previous research with the MEG or related coordination games typically uses either stated or actual effort, which is not directly comparable (Bortolotti et al., 2016; Charness et al., 2018). Here, we developed a modified MEG, hereafter referred to as the real-effort MEG or REMEG, that incorporates actual effort in the form of arithmetic problems that correspond in difficulty to the stated level of effort. A recent study addressed this effort issue by creating a MEG with actual effort (Bortolotti et al., 2016), based on the original MEG (Van Huyck et al., 1990). However, this MEG variation involves a slightly different type of coordination situation since the effort level was not directly selected by participants but determined post hoc based on the participants' exerted effort. This is in contrast with other MEG experiments that used explicitly stated effort (i.e. selecting a level of effort; Camerer & Ho, 1998; Leng et al., 2018; Van Huyck et al., 1990). Bortolotti et al. (2016) used performance on a difficult coin sorting and counting task to represent effort and the lowest performing member of the group determined the minimum. They identified weak link players as the source of coordination failure, found coordination at higher levels than previously observed, and information availability had little effect on effort and weak links. There were two potential limitations regarding generalizability to other MEG studies. There was no effort selection and players may have perceived the coin sorting task as a more of an individual task than a group task, compared to typical MEG experiments using stated effort. In the pilot study, we address the potential limitations to Bortolotti et al. (2016) by including both stated and actual effort related to coordination for each participant. The effort-eliciting task was designed to be effortful, while not necessarily difficult, and involved mental addition with single-digit numbers (ranging from two to eight digits depending on effort level). However, it is possible that participants with greater mathematical skill would perceive these problems as less effortful, which could potentially skew this measure of effort. Instead of using errors as a measure of effort, incorrect answers result in deductions from an individual player's payoff and do not affect the group. This is somewhat related to less skilled players in the Bortolotti et al. (2016) study in which the players paid for more time to complete the coin task, which was a personal loss but was beneficial for the group payoff.

In addition to actual effort, the REMEG has some other modifications. In a typical MEG, players receive instructions about game structure and payoffs, are not allowed to communicate, and are given outcome information about the minimum for each round. Players usually converge to the least efficient or risk-dominant option over time, even when they are given the distribution of other player choices (e.g. Camerer & Ho, 1998; Leng et al., 2018; Van Huyck et al., 1990). As previously mentioned, there are several ways to increase coordination; however, the concern here is parsimony and minimal interaction between individuals. For this reason, we only manipulated information availability. Players are given access to a payoff matrix at all times and are given the distribution of player choices and counterfactuals (i.e. forgone payoffs based on different choices and different minimums) after each round. These modifications ensured that participants understood how to play, and having access to the distribution of payoffs allowed players to see the variation in player choices and potentially use choices as signals to bring up the minimum.

Since it is difficult to determine underlying motivations, covert intentions of individuals, and whether or not players were attempting to signal, we included some basic questions at the end of the MEG to provide additional information. After the last round, players were asked six questions about their preferences during the game. Participants rated their strength of preference for each question on a 7-point Likert scale (e.g. 1 – lowest preference and 7 – highest preference). The questions were: "I didn't want to have to do a hard math problem," "I wanted a challenging math problem," "I wanted to maximize my own payoff," "I didn't want to seem self-centered," "I wanted to be a team player," and "I wanted to see what would happen." In addition, effort-related preferences and traits are measured to better understand individual behavior in coordination. These include: Need For Cognition (Cacioppo & Petty, 1982), Tolerance Of Mental Effort (Dornic et al., 1991), Preference For And Tolerance Of The Intensity Of Exercise (Ekkekakis et al., 2005), Industriousness (Jackson et al., 2010), risk avoidance, the Brief Self-Control Scale (Tangney et al., 2004), and Trait Trust (Collins et al., 2016). Effort-related traits were included to explore potential relationships with coordination behavior and might indicate which individuals are more likely to coordinate efficiently. As previously mentioned, several studies (e.g. Juvina et al., 2018; Kool et al., 2010) reported that avoiding effort was negatively correlated with several effort-related traits, including self-control and the need for cognition. In addition, risk and trust were added based on Devetag and Ortmann (2007) literature review and findings from Bosworth (2013) and Engelmann & Normann (2010) suggesting that trust encourages more efficient coordination.

We hypothesized that 1) players will coordinate more efficiently compared to previous findings where players converged towards the least efficient option and 2) effort-related trait measures correlate with effort selections in the REMEG.

Method

Participants

Participants were recruited by advertisements on Wright State University's campus in Dayton Ohio. Data were collected from a total of 32 participants (54% female) with a mean age of 22.8 ($SD = 6.8$). All participants were compensated by a base pay of 10 USD and the opportunity to earn an additional 10 USD bonus contingent on performance during the experiment. Performance pay was calculated based on following instructions and performance on the REMEG (i.e. cumulative payoff and deductions for incorrect arithmetic problem solutions).

Materials

The entire experiment was run on computers, which were organized in rows of four booths with barriers between each computer booth. The REMEG and surveys were programmed and run using the O-tree platform (Chen et a., 2016). These materials are available upon request.

Design and procedure

Prior to the start of the experiment, each participant was assigned a computer station in one of the four booths. Participants read and signed a consent form, received instructions, and were informed about their compensation. Participants were instructed that

communication of any kind was not allowed and were told the REMEG was a group task that includes all participants. The REMEG and effort-related trait questionnaires took approximately 45 minutes to complete.

REMEG. In the REMEG (O-tree platform; Chen et al., 2016), participants were given instructions prior to starting the game, including: 1) a coordination scenario to provide some context to the MEG where each individual has a separate task to complete which affects the overall group performance and the resulting outcomes, 2) instructions about the structure of the game, 3) how effort selections determine the difficulty of the corresponding arithmetic problem and incorrect solutions result in deductions from the payoff earned during the current round, and 4) instructions and examples of player choice distributions and counterfactuals given after each round. In the actual game, each round begins by choosing how much effort to allocate towards the group task. Participants will then be required to solve an arithmetic problem that corresponds in difficulty to the level of effort chosen during a 90 second delay. For instance, choosing the effort level of one involves adding two single-digit numbers and each increment in effort adds one single-digit number to the problem. After choosing a level of effort and completing the arithmetic problem, they are shown the responses of all group members, the minimum effort of the group, and their own individual payoff. After these results are shown, participants are then shown bidirectional counterfactuals that highlight how players could have done better or worse if they had chosen differently. For instance, if the player chose two and the minimum was one, they would receive a payoff of 60. An upward counterfactual could state how they would have earned 70 if they had chosen one and coordinated on the minimum or 80 if the minimum was two. It is important to note each player may receive unique counterfactuals since counterfactuals are generated based on their choice and the minimum for a given round. This was a potential confound for evaluating the impact of counterfactuals, which will be further addressed in the preregistered experiment. To ensure participants are actually attempting to accurately solve these problems, there was an additional cost for incorrect answers. Participants were informed that points earned during the game and their performance on arithmetic problems will affect their performance bonus.

Effort-related questionnaires. Upon completion of the minimum effort game, participants filled out the effort-related trait scales.

Pilot study results

The results are separated into three sections. Round level analyses consolidated all participants to look at average effort across all 20 rounds. Group-level analyses focused on coordination efficiency. Individual-level analyses investigated relationships between REMEG variables, and between the REMEG and effort-related traits for each participant. All data analyses were completed using MatLab version R2018a.

Round level

Participants' effort was averaged for each of the 20 rounds and compared to data from Van Huyck et al. (1990) and Leng et al. (2018) for 10 rounds. However, it is important to

note that only our study involved an effortful task corresponding to effort choices. Our REMEG data consists of 8 four-person groups (*N* = 32). In Van Huyck et al. (1990) there were seven 16-person groups (*N* = 112) and participants were only given the minimum choice after each round. In Leng et al. (2018), 10 six-participant groups (*N* = 60) were given complete and continuous outcome information after each round (i.e. effort selection, minimum effort, and cumulative payoffs). Participants in our REMEG had higher average effort (*M* = 4.84, *SD* = 0.43) across 20 rounds compared to the average effort (*M* = 2.78, *SD* = 1.18) for the Van Huyck et al. (1990) data across 10 rounds. The difference in effort was significant, $t(28) = 7$, $p < .0001$, $d = 2.32$. We did not have access to the complete data from Leng et al. (2018), so we were unable to perform a *t* test; however, we used the average first and last round effort to approximate a trend line for comparison (see Appendix B in supplementary document).

Group level

Average group effort and effort distribution for each group were calculated by averaging effort and effort distribution (i.e. range) per round and then getting a complete average for each group. Average group effort represents efficiency and effort distribution reveals the degree that players coordinated. A comparison between average group effort and effort distribution revealed a non-significant negative relationship, $r(6) = -0.69$, $p = .057$. Similarly, there was also a non-significant negative relationship between average effort and effort distribution, $r(5) = -0.25$, $p = 0.59$, for the Van Huyck et al. (1990) data which is interpreted as a complete lack of relationship. Although non-significant, the negative trend in our data suggests that when players coordinated (low effort distribution) it was more efficient (higher effort).

Individual level

The difference between individual effort selection and the group minimum was calculated for each round and then averaged across all 20 rounds for each player, producing a mean distance from the minimum (*M* = 0.98, *SD* = 0.55). This method was more appropriate for looking at differences between individuals than just comparing effort since it accounts for the fact that individuals are nested within groups and each group is dynamic and unique. The average distance from the minimum had a non-significant negative relationship with both first-round effort ($r(30) = -0.09$, $p = 0.63$) and average effort ($r(30) = -0.24$, $p = 0.18$), and had a significant negative relationship with average group effort ($r(30) = -0.56$, $p < 0.001$). This suggests that coordination (i.e. low average distance from minimum) tended to occur more often at high levels of effort and coordination on low levels of effort was often disrupted by players' attempts to choose higher levels of effort (we will later refer to this behavior as signaling). Players were ranked according to their mean distance from the minimum to look at weak links; however, no significant results were found and players classified as weak links were not consistently the weak link across all rounds. There were no significant or notable relationships found between REMEG variables and effort-related trait measures.

Preregistered experiment

In most MEG experiments, players converge towards the most inefficient or lowest payoff option within a few rounds (e.g. Van Huyck et al., 1990) or more gradually (e.g. Leng et al., 2018). However, in the pilot study, we saw more efficient coordination (i.e. higher effort) and more stable average effort across 20 rounds, compared to these previous findings. This higher effort and stability suggest that when players coordinate, it is more likely at higher effort. These results were promising regarding improving coordination efficiency; however, we can only speculate as to why this improvement was observed. For instance, behavior could have been influenced by the real-effort task, the additional outcome information (i.e. distribution of all choices and payoffs), the counterfactuals, or the small group size (i.e. smaller groups tend to coordinate more efficiently). However, a previous study (Bortolotti, et al. 2016) already addressed the potential influence of effort and Van Huyck et al. (1990), Camerer and Ho (1998), and Leng et al. (2018) included a condition with player choice distributions, which did not do much to improve coordination efficiency over the minimum only condition(s). In addition, Leng et al. (2018) found higher levels of effort with their smaller groups of six and complete outcome information, but effort gradually declined over time showing a less extreme, yet similar pattern as previous studies. Considering player choice distributions had little effect on coordination efficiency in previous studies (e.g. Camerer & Ho, 1999; Leng et al., 2018; Van Huyck et al., 1990), it is assumed the greater efficiency observed here more likely resulted from the counterfactuals. However, there is currently no direct supporting evidence for this conclusion and there could be confounding variables (e.g. group size). In addition, each player received unique counterfactuals (e.g. number of counterfactuals and the difference between actual and forgone payoffs) that were dependent on their choice relative to the minimum. To address these issues, we investigated the extent that counterfactuals improved coordination efficiency and included additional measurements (e.g. the number of counterfactuals and the difference between counterfactual payoffs and the actual payoff) to address the potential confound of participants receiving unique counterfactuals. Since we focus on improving coordination efficiency, we will investigate the effectiveness of counterfactuals to nudge individuals towards choosing higher effort and coordinating more efficiently. The same design from the pilot study was used except rather than one condition, there were four counterfactual conditions: bidirectional, upward, downward, and no counterfactuals (control). The bidirectional condition served as a replication of the pilot study and the no counterfactual control group served as a useful comparison for the counterfactual conditions and previous studies without counterfactuals. The upward and downward counterfactual conditions served as a separate assessment of counterfactual direction and were compared to the bidirectional and control conditions. In general, upward counterfactuals are more likely to lead to behavioral change (Markman & McMullen, 2003; Markman et al., 2008) and improved future performance (Morris & Moore, 2000; Nasco & Marsh, 1999), while downward counterfactuals are less likely to lead to improvement due to rationalization (Smallman & Summerville, 2018) and focus on how outcomes could have been worse (Rim & Summerville, 2014). Therefore, we expected coordination to be most efficient in the upward counterfactual condition, followed by the bidirectional and lastly, the downward condition. Since the control condition had no counterfactuals and did not include other player's choices, we expected it to have the lowest efficiency. This

result, in combination with the replication of the bidirectional counterfactual condition, could support the assumption from the pilot study that counterfactuals nudged participants towards choosing higher effort and coordinating more efficiently. Then, differences between upward and downward counterfactual conditions could further elaborate on why this may have occurred. The control condition would allow us to unequivocally attribute any observed differences to the counterfactual manipulation. We also included an instruction quiz and manipulation check. Participants completed a brief quiz after the REMEG instructions to ensure they had some understanding of how the REMEG works and what the outcome information means. A debrief questionnaire was included after the REMEG as a manipulation check to indicate whether participants were aware that the counterfactual information consistently indicated how outcomes could have been better (upward) or worse (downward). However, we did not expect the benefits of counterfactuals to be contingent upon one's awareness.

Method

We used the same REMEG and survey set up as the pilot study with four counterfactual treatments, an instruction quiz, and a manipulation check. Also, based on null findings with the survey and the long experiment run time, we excluded all trait measures except for industriousness and trait trust. We originally planned on running four-person groups, but this was not feasible due to a very high dropout rate, long wait times, and a substantial increase in cost. Our solution was to run the experiment with one human and three agents that were previously developed and implemented in our pilot study. To reduce noise in agent behavior and allow for more clear-cut comparisons, we used the same agents for each condition. Furthermore, we did not have appropriate data to estimate the influence of counterfactuals, so agents did not receive counterfactuals and instead only received information about other players' choices and the minimum for each round. This means that only humans receive counterfactuals and since all agents are the same, any differences between conditions can be mainly attributed to the influence of counterfactuals on human behavior. We acknowledge that this change needs to be considered when framing hypotheses and interpreting the results. More details about this design change and agents were documented in an amendment to the original preregistered manuscript and can be found in the supplementary document (Appendix C).

Participants

We previously performed a power analysis and approximated effect size using the average of effect sizes (.2) observed in Markman et al. (2008) between upward and downward counterfactuals (η = .15 and η = .32). The power analysis was performed for a fixed-effects, omnibus, one-way ANOVA using G*power software (Faul et al., 2009, 2007). We used an alpha of .05, power of .8, used a small-to-medium effect size of .2, and assumed that the treatment groups will have the same number of participants. The power analysis resulted in a required sample size of 280.

We recruited 278 participants through Mturk. However, we had to remove the first 16 participants as we fine-tuned the methodology and two participants who repeated the study. Therefore, our analyses included 260 participants. The sample was predominately male (65%) and the mean age was 34.9 (SD = 9.6). Participants (N = 260) were randomly

assigned to control ($n = 63$), upward ($n = 66$), downward ($n = 68$), and bidirectional conditions ($n = 63$).

Inclusion/exclusion

Participants were eligible to participate as long as they were over 18 years of age, proficient in English, had a normal or corrected vision, and did not have serious or unstable medical or mental illnesses. Participants were required to have an MTurk study completion rate of 90% or higher and were only allowed to participate once, including the pilot studies, regardless of condition. Since the experiment is online and takes approximately 20 minutes, compensation was adjusted to a base pay of 3 USD and the opportunity to earn an additional 2 USD bonus contingent on performance during the experiment. Just like the pilot study, performance pay was calculated based on following instructions and performance on the REMEG. Participants were informed about the performance pay in the instructions.

Materials

The REMEG and trait scales will run on the O-tree platform (Chen et al., 2016) on MTurk.

Design and procedure

Prior to the start of the experiment, participants were given a consent form, received instructions, and were informed about their compensation. Participants completed the REMEG and two trait scales, which lasted approximately 20 minutes.

REMEG. The same REMEG (O-tree platform; Chen et al., 2016) from the pilot study was used with the exception of separate counterfactual treatments and group composition (i.e. one human and three agents). In addition, we will add a brief quiz after the MEG instructions and a manipulation check to assess participants' awareness of the counterfactual manipulation.

Effort-related questionnaires. We used the same industriousness and trait trust questionnaires from the pilot study.

Hypotheses

We expected to see more stable and efficient coordination in counterfactual conditions (not control) compared to previous studies (e.g. Leng et al., 2018; Van Huyck et al., 1990), just as we did in the pilot study. However, we expected to see a significant difference in average effort across 20 rounds between the counterfactual conditions. Specifically, we expected those in the upward condition to coordinate more efficiently (i.e. higher effort) compared to those in the downward and control conditions. In addition, we expected the control to have significantly lower effort than all other treatment conditions since it was the most similar to previous work with less efficient coordination. If our hypotheses are supported, it would suggest counterfactuals were likely responsible for the improved coordination efficiency observed in the pilot study and upward counterfactuals are more likely to lead to increases in coordination efficiency. Furthermore, these results would suggest that nudging individuals towards more efficient coordination by means of upward counterfactuals is more effective than just giving them more information (e.g.

distribution of choices or bidirectional counterfactuals). Lastly, in the pilot study, the relationships between effort-related traits and behavior in MEG were inconclusive. However, we decided to include industriousness and trait trust in the preregistered study as well, because they may prove useful in characterizing the real-effort version of MEG and they may moderate the effect of the counterfactual intervention. The general hypothesis regarding these measures is that individuals who are generally more prone to expend effort or trust others also tend to select higher levels of effort in REMEG (at least in the initial rounds) and possibly respond more positively to the counterfactuals. Since we used three agents and one human in the experiment, the hypotheses needed to be put in the proper context. Since only the human receives counterfactuals, we expected the effects of counterfactuals to be weaker compared to four human players all receiving counterfactuals. Therefore, prior to collecting data, we felt it was appropriate to run simulations for 100 groups composed of one simulated human and three agents for each condition to place our hypotheses in the appropriate context (see Figure 1 and an additional figure in Appendix D of the supplementary document).

The agents are identical to those used in the study and the human player is simulated by providing them with condition-specific counterfactuals and assuming they are weighted half as much as the actual outcome of each round. In addition, due to the differences between humans and agents, we performed exploratory analyses to compare them; however, we did not have preregistered hypotheses for this comparison.
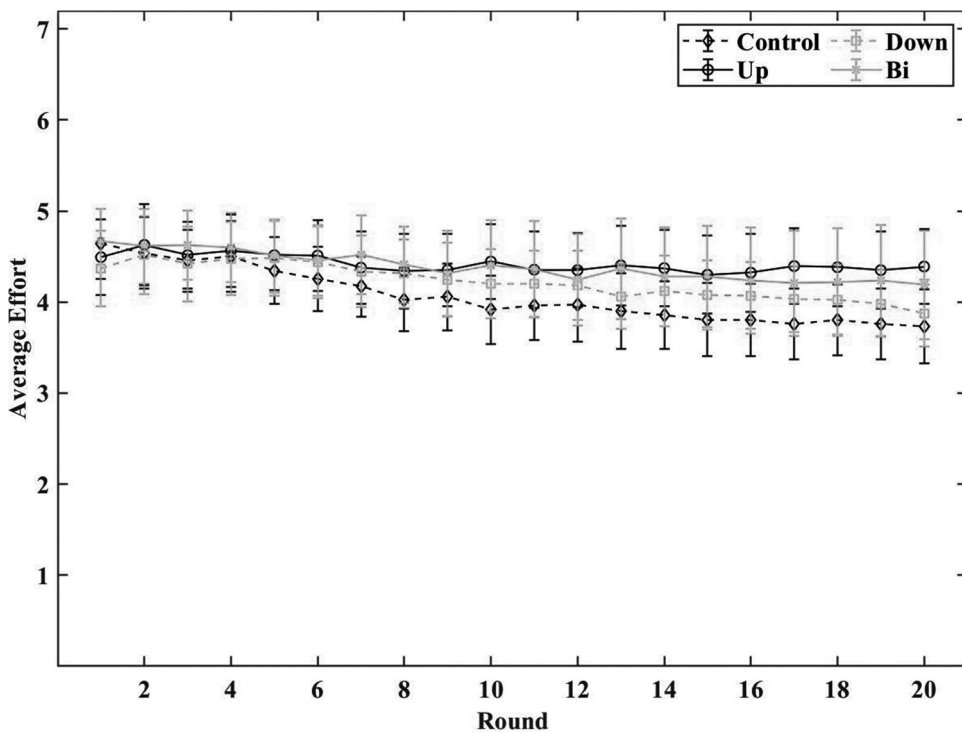

Figure 1. Line plot showing average effort per round for 100 simulated groups (simulated human exposed to counterfactuals and three agents) in each counterfactual condition.

Results: preregistered

We performed nested ANOVAs and correlational analyses using MatLab version R2020a and all linear effects modeling was conducted using R version 4.0.3 with the lmerTest (cran.r-project.org/web/packages/lmerTest) and r2glmm statistical packages (github.com/bcjaeger/r2glmm).

Instruction quiz and manipulation check

Before running analyses, the instruction quiz (IQ) and manipulation check (MC) were explored (see Appendix E in the supplementary document for additional information). Participants took the IQ immediately after instructions and were given feedback. In addition, participants were shown the payoff matrix at the start and end of each round. Participants passed if they answered at least two of the three questions correctly. More than half of the participants failed the IQ. However, a linear mixed-effects model for average human effort with condition and IQ as factors, and group as a random variable, revealed no main effect of IQ or interaction with condition. The MC results were less straightforward. Participants passed the MC if their answer corresponded to their experimental condition and about 70% of participants failed. However, the same linear mixed-effects model for MC did not reveal any effects or interactions. Furthermore, participants gave similar answers for the MC and a Chi-squared test revealed there were no differences in answers between conditions, $\chi^2(12, N = 259) = 17.12$, $p = .15$. It is not clear why participants failed the MC; however, the benefits of counterfactuals were not expected to be contingent on awareness. We also attempted to identify participants who were not paying attention based on time spent on results and counterfactual pages, but this was inconclusive. Therefore, it seemed reasonable to include all of the data for the preregistered analysis. We did perform and include exploratory analyses for those who passed and failed the IQ (Appendix F), MC (Appendix G), and for those who passed both (Appendix H) in the supplementary document.

Unique counterfactuals

Since each player received a unique set of counterfactuals, we wanted to rule out this possible confound. Unique counterfactuals were quantified for each participant by subtracting the actual payoff from each counterfactual payoff and the sum for each round was calculated. We ran a linear mixed-effects model for counterfactual payoff difference with condition and round as fixed effects and group as a random variable. The model revealed effects for the downward, $\beta = -57.55$, $t(364) = 10.61$, $p = 2e\text{-}16$, *partial R^2* = .15, and bidirectional, $\beta = -24.64$, $t(364) = -7.85$, $p = 4.72e\text{-}14$, *partial R^2* = .031, conditions indicating differences between all conditions with the downward condition having the lowest counterfactual payoff difference, followed by the bidirectional condition. In addition, there was an effect for round, $\beta = -.54$, $t(3740) = -4.92$, $p = 9.19e\text{-}07$, *partial R^2* = .004, indicating a negative trend for all conditions for counterfactual payoff difference. At the individual level, there was no relationship between the player's average counterfactual payoff difference and average effort, suggesting it was not a confound (see Appendix I in the supplementary document).

Preregistered analysis

To test the hypotheses, we used a nested ANOVA with effort selection as the dependent variable, condition as the independent variable, and players nested within a group as a random variable (see Appendix J in the supplementary document). The hypotheses predicted the upward condition to have the highest effort and the downward condition to have higher effort than the control condition. The nested ANOVA was significant, $F(3,1036) = 4.42$, $p = .004$, $\eta2 = .01$, and post hoc Tukey HSD tests revealed higher effort in the upward condition ($M = 4.47$, $SD = 1.10$) compared to downward ($M = 4.13$, $SD = 1.05$), $t(1036) = 3.61$, $p = .0001$, $d = .31$. The same ANOVA with only human data was not significant.

We had expectations to see trends across 20 rounds as the first study and model simulations suggested no differences in first-round choices and differences emerging later around round 5. Here, a linear mixed-effects model was used because it has less restrictive assumptions about variance and repeated measures (e.g. Krueger & Tian, 2004). Since only human players are unique and differences between conditions were only present with all player data, player type was added as an interaction term. We used the following formula: Effort ~ Condition * Round * Player Type + (1 | Group/Player), which corresponds to effort as the dependent variable, an interaction effect for condition, round, and player type, and players nested within a group as a random effect (see Figure 2). This model serves as the standard for all subsequent analyses. The model revealed a significant effect for round, $\beta = -.03$, $t(19,750) = -4.89$, $p = 1.03e\text{-}06$, partial $R^2 = .001$, indicating
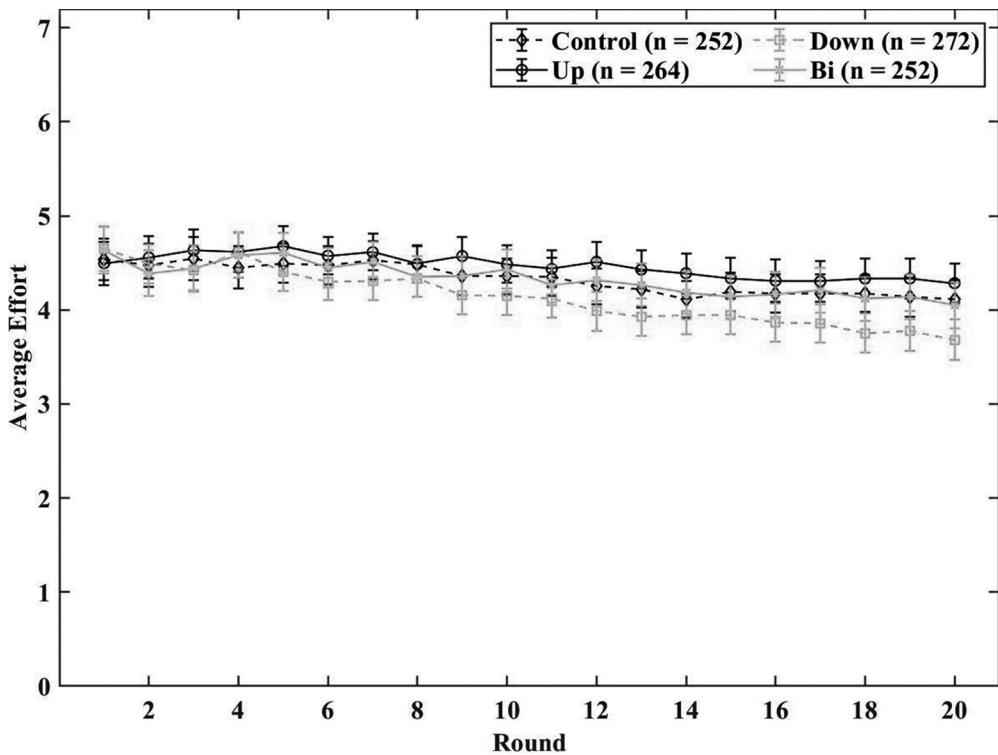

Figure 2. Average effort per round for all conditions. Error bars are 95% confidence intervals.

a negative trend across rounds for effort. In addition, we found some significant interaction effects. There was an interaction effect for agents and the upward condition, $\beta = .422$, $t(3724) = 3.02$, $p = .003$, partial $R^2 = .000$, indicating that, on average, agents had higher effort than humans in the upward condition. There was an interaction effect for the upward condition and round, $\beta = .02$, $t(19,750) = 2.37$, $p = .02$, partial $R^2 = .000$, and the downward condition and round, $\beta = -.02$, $t(19,750) = -2.02$, $p = .04$, partial $R^2 = .000$, indicating the weakest negative trend across rounds for the upward condition and the strongest negative trend for the downward condition.

We also performed these analyses for payoff and included them in the supplementary document (Appendix K). Lastly, we explored trait trust and industriousness and failed to find any relationships with any variables in the REMEG.

Results: exploratory

Comparing humans and agents

Groups were comprised of one human and three agents, and we found an effect for player type where agents appeared to choose higher effort than humans. To better understand this unexpected result, we explored signaling behavior. Signaling behavior was quantified by subtracting each player's choice by the minimum to get the distance from the minimum (min-dist) for each round, which is informative about instances of signaling and its strength (i.e. costliness). The standard linear mixed-effects model with min-dist as the dependent variable revealed an effect for round, $\beta = .064$, $t(19,750) = -10.49$, $p = $ 2e-16, *partial $R^2 = .005$*, indicating a negative trend across rounds. In addition, we found an effect for the upward condition, $\beta = -.35$, $t(3018) = -2.71$, $p = .007$, *partial $R^2 = .001$*, and an interaction effect for the upward condition and agents, $\beta = .42$, $t(3622) = 3.04$, $p = .002$, *partial $R^2 = .001$*, suggesting the upward condition effect was driven by the counterfactual manipulation and the effect of human signaling on the agents' behavior in the upward condition (see Appendix L in the supplementary document). These results suggest that signaling behavior is not independent of condition and the agent's signaling behavior may be more frequent than humans. However, agents were by design, not capable of signaling like humans. Signaling is effective if it influences other players to choose higher effort, which brings up the minimum and therefore potential payoffs for all players. Signaling is often more effective when it is known to be costly to the signaler and persists over time as it has delayed effects on other players. Effective signaling requires short-term costs (i.e. lower payoffs), which are offset by higher payoffs over time as the minimum increases. Agents only best respond to other players' choices, specifically the minimum and mode(s), and do not consider the short-term cost of signaling with long-term rewards of higher payoffs. In addition, the actual payoff received is weighted higher than all forgone payoffs, which can make the previous round choice more attractive than forgone choices. Therefore, an agent's signaling-like behavior is an artifact of their "stickiness" to a mode or their previous choice. We performed additional linear mixed-effects models to better understand the differences between human signaling and the signaling-like behavior of agents. Players were identified as signalers if their min-dist was above zero for at least four rounds in a row. We choose 4 after exploration of a number of criteria (e.g. self-report measures), correspondence with previous literature (e.g. Brandts et al., 2015, 2014), and it provided the most even split between signalers and non-signalers. We ran the standard linear mixed-effects models for

effort and payoff with signalers replacing condition (i.e. Effort ~ Round * Signaler * Player type + (1 | Group/Player)). Since the effects for round and player type, and their interaction effect were already discussed for effort and payoff (in supplementary document), they are left out here. We found several interaction effects for effort and included them in a table in the supplementary materials (Appendix M). For simplicity, we only report the three-way interaction for agent, round, and signaler, $\beta = -.10$, $t(19,760) = -14.52$, $p = 2e\text{-}16$, *partial* $R^2 = .006$. This interaction reveals differences between signalers and non-signalers, however, that difference is greater for humans compared to agents (see Appendix M in supplementary document). Similarly, the model for payoff revealed several effects and interaction effects (see table in Appendix N of the supplementary document) and the same three-way interaction for agent, round, and signaler, $\beta = -1.48$, $t(20,530) = -14.07$, $p = 2e\text{-}16$, *partial* $R^2 = .006$. This interaction reveals differences between signalers and non-signalers for both humans and agents (Figure 3). Human signalers showed the expected pattern of short-term costs (i.e. lower payoffs) with the long-term benefit of higher payoffs over time by surpassing that of non-signalers. However, differences between agent "signalers" and "non-signalers" highlighted the artificiality of their signaling-like behavior as "non-signalers" earned higher payoffs than "signalers," likely due to their choice stickiness.

Agent's signaling was artificial and human signalers showed the expected cost of signaling with the long-term benefits of higher payoffs. In addition, humans were the only unique players in the group. Therefore, we explored how human signaling affected group effort (Appendix O in supplementary document) and payoff (Figure 4).
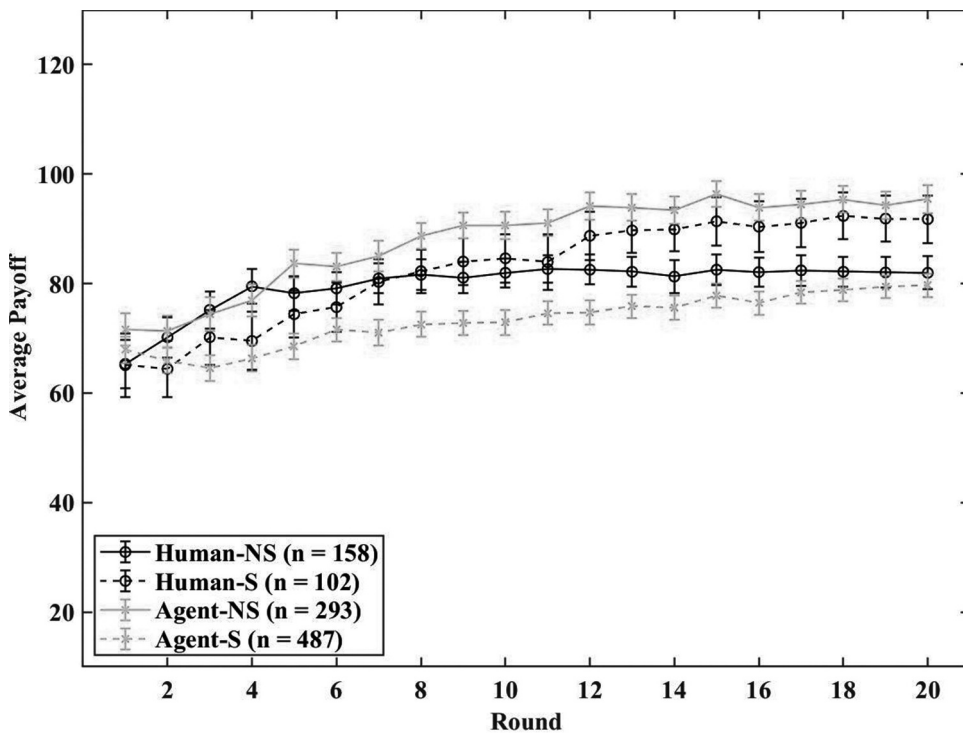

Figure 3. Average payoff per round for humans (black) and agents (gray) that were classified as signalers (dotted line) and non-signalers (solid line). Error bars are 95% confidence intervals.
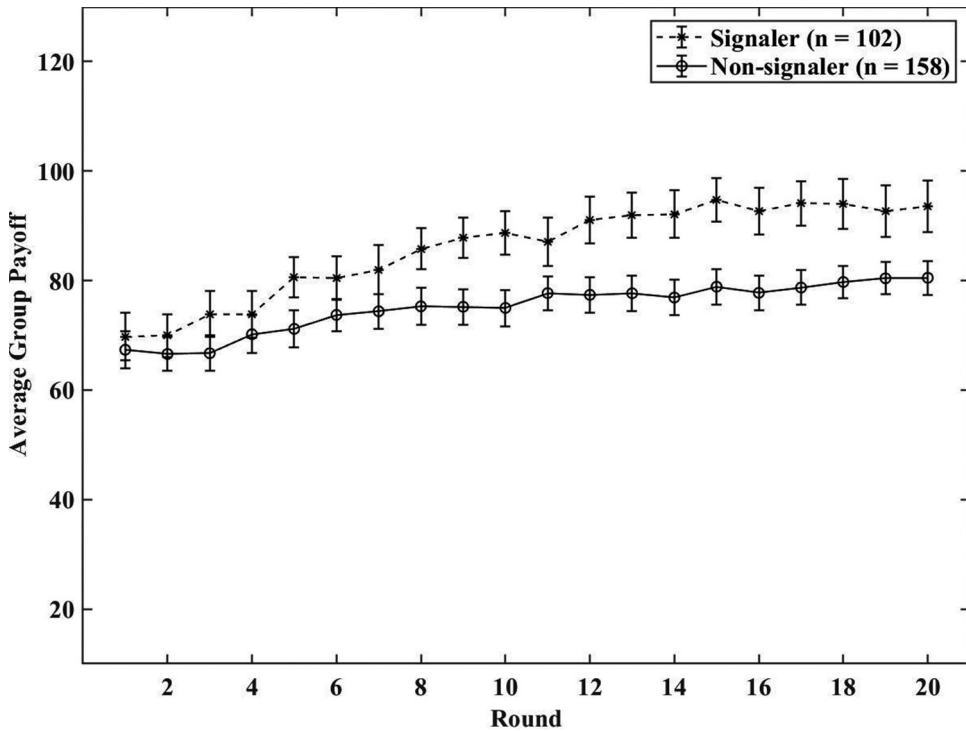
Figure 4. Average payoff per round for groups with human signalers (solid line) and non-signalers (dotted line). Error bars are 95% confidence intervals.

We ran linear mixed-effects models for average group effort and payoff with round and signaler as factors, and group as a random variable (i.e. Average group effort ~ Round * Signaler + (1 | Group)). For average group effort, there were effects for round, $\beta = -.045$, $t(4938) = -10.3$, $p = 2e\text{-}16$, *partial* $R^2 = .013$, and signaler, $\beta = 1.38$, $t(403) = 8.53$, $p = 3.1e\text{-}16$, *partial* $R^2 = .032$, indicating an average negative linear trend for average group effort across rounds and higher average group effort for groups with human signalers. There was also an interaction effect for round and signaler, $\beta = .043$, $t(4938) = 6.16$, $p = 7.8e\text{-}10$, partial $R^2 = .005$, indicating a positive linear trend for average group effort for signalers, and a negative trend for non-signalers.

We found the same effects and interacting effects for average group payoff. The effects for round, $\beta = .56$, $t(4938) = 10.57$, $p = 2e\text{-}16$, *partial* $R^2 = .016$, and signaler, $\beta = -7.25$, $t(519) = -4.47$, $p = 9.7e\text{-}6$, *partial* $R^2 = .007$, indicated an average positive linear trend for average group payoff across rounds and higher average group payoff for groups with human signalers. The interaction effect for round and signaler, $\beta = .96$, $t(4938) = 11.38$, $p = 2e\text{-}16$, partial $R^2 = .018$, indicated a stronger positive linear trend for average group effort for signalers, compared to non-signalers. These results suggest that on average the human was able to influence the group to choose higher effort by signaling, and the human signaling resulted in higher payoffs over time for the group.

General discussion

Preregistered analyses

In the preregistered analyses, we found evidence that counterfactual nudging can increase coordination efficiency. We expected nudging in the form of counterfactuals to influence coordination efficiency with the upward condition having the highest effort and the control the lowest. The nested ANOVA only revealed differences between the upward and downward conditions. The linear mixed-effects model extended the nested ANOVA findings by revealing different linear trends across rounds for the upward and downward conditions. In addition, agents had higher effort than humans in the upward condition. This provides evidence that counterfactual nudging influenced coordination efficiency and provides evidence that humans and agents behaved differently, however, the effect sizes were small. Prior to collecting data, we expressed concern (method section and amendment in Appendix C of supplementary document) that the effects of counterfactual nudging would be weakened since only the human received the manipulation. Although we found supporting evidence for counterfactual nudging, the small effect sizes support that concern. Furthermore, although we included all the data, the fact that 70% of the participants failed the MC and about half failed the IQ was still concerning. This could be due to participants not taking the experiment seriously, not paying attention to the counterfactuals, getting confused by the counterfactual manipulation (suggested by similar answers across conditions), or generating their own counterfactuals and ignoring the ones provided. In addition to counterfactual conditions, we saw a higher average effort than expected in the control condition. There were no counterfactuals or additional outcome information beyond the minimum; however, the control condition was almost identical to the bidirectional condition and had higher effort (not significant) than the downward condition.

Exploratory analyses

In the exploratory analyses, we found evidence that (1) humans and agents behaved differently in respect to effort, payoff, and signaling, and (2) one human signaler, in a group of three equivalent agents, can increase coordination efficiency of the group. The linear mixed-effects model with min-dist revealed a weak relationship with condition since there was only one interaction effect (i.e. upward condition and agents). Signaling could be seen as individuals nudging each other to choose higher effort and may have been stronger or at least additive to the counterfactual nudges. This could be due to counterfactuals lacking context beyond the minimum since all player choices were shown on the results page prior to counterfactuals. Signaling behavior could help explain why the control condition performed better than expected. Although there were no counterfactuals, signaling was still possible and, without information of all player choices, players could not properly assess signaling effectiveness and if or when to give it up. Furthermore, agents in the control condition had access to all player choices, which could have helped keep the effort higher and more stable.

We found player type differences in our preregistered analysis with effort and our exploratory analysis with min-dist, where agents had higher average effort and min-dist than humans. To better understand these differences, we followed up these findings with

exploratory analyses by classifying signalers and including it as a factor in separate linear mixed-effects models with effort and payoff as dependent variables. For effort, we found "signaling" and "non-signaling" agent's effort choices were similar, whereas signaling human's effort was about 1.5 effort units higher than non-signaling humans. For payoff, we found that signaling humans increased their payoff over time, compared to non-signaling humans who flattened out after about five rounds. On the other hand, "signaling" agents had the lowest payoffs and "non-signaling" agents had the highest. Also, there were more agent "signalers" (62%) compared to humans (39%). The differences between signaling and non-signaling humans are what we would expect with effective signaling behavior. However, the behavior of signaling and non-signaling agents is unusual. As previously mentioned, we believe agents signaling-like behavior is merely an artifact of the influence of mode(s) and previous choices. That being said, the agent's choices could often be maladaptive in terms of their own payoff and the coordination efficiency of the group.

The linear mixed-effects models with signaling as a factor revealed differences between signaling and non-signaling humans and agents for both effort and payoff. However, since only humans are capable of signaling, we performed a subsequent exploratory analysis to explore the behavior of groups with signaling and non-signaling humans. The linear mixed-effects models revealed groups with signaling humans had significantly higher effort and payoffs compared to groups with non-signaling humans. This result is promising because it suggests that the single human in the group had awareness of the need for signaling persistence and succeeded in influencing the behavior of three equivalent agents, regardless of counterfactual nudges.

Overall summary

The preregistered and exploratory analysis results are somewhat promising for counterfactual nudging and more so for signaling. The cards were stacked against the human since agents did not receive counterfactuals, were not capable of signaling and may be hard to influence. However, we still found some evidence that upward counterfactuals nudged humans towards choosing higher effort and this had an influence on the group. Furthermore, signaling humans were able to nudge the behavior of the group towards more efficient choices regardless of experimental condition. A follow-up study is necessary to rule out potential confounds by (1) developing more human-like agents capable of signaling and giving them counterfactuals, or using four-human groups, (2) making sure all players in the control condition receive the same information, (3) providing a clearer counterfactual manipulation by explicitly relating them to group behavior (e.g. if other players had chosen one effort unit higher, you would have earned X), and (4) finding a more appropriate way to assess whether participants were paying attention and following instructions.

Acknowledgments

References

Beard, T. R., & Beil, R. O. (1994). Do people rely on the self-interested maximization of others? An experimental test. *Management Science*, *40*(2), 252–262. https://doi.org/10.1287/mnsc.40.2.252

Berninghaus, S. K., & Ehrhart, K.-M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization*, *37*(2), 231–248. https://doi.org/10.1016/S0167-2681(98)00086-9

Blume, A., DeJong, D. V., Kim, Y. G., & Sprinkle, G. B. (1998). Experimental evidence on the evolution of meaning of messages in sender-receiver games. *The American Economic Review*, *88*(5), 1323–1340. https://www.jstor.org/stable/116874

Bortolotti, S., Devetag, G., & Ortmann, A. (2016). Group incentives or individual incentives? A real-effort weak-link experiment. *Journal of Economic Psychology*, *56*, 60–73. https://doi.org/10.1016/j.joep.2016.05.004

Bosworth, S. J. (2013). Social capital and equilibrium selection in Stag Hunt games. *Journal of Economic Psychology*, *39*, 11–20. https://doi.org/10.1016/j.joep.2013.06.004

Brandts, J., & Cooper, D. J. (2006). A change would do you good …. *An Experimental Study on How to Overcome Coordination Failure in Organizations*. *American Economic Review*, *96*(3), 669–693. https://doi.org/10.1257/aer.96.3.669

Brandts, J., Cooper, D. J., Fatas, E., & Qi, S. (2015). Stand by me—experiments on help and commitment in coordination games. *Management Science*, *62*(10), 2916–2936. https://doi.org/10.1287/mnsc.2015.2269

Brandts, J., Cooper, D. J., & Weber, R. A. (2014). Legitimacy, communication, and leadership in the turnaround game. *Management Science*, *61*(11), 2627–2645. https://doi.org/10.1287/mnsc.2014.2021

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*(1), 135–157. https://doi.org/10.1146/annurev-psych-122414-033249

Cachon, G. P., & Camerer, C. F. (1996). Loss-avoidance and forward induction in experimental coordination games. *The Quarterly Journal of Economics*, *111*(1), 165–194. https://doi.org/10.2307/2946661

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

Camerer, C., & Ho, T. H. (1998). Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation. *Journal of Mathematical Psychology*, *42*(2–3), 305–326. https://doi.org/10.1006/jmps.1998.1217

Camerer, C., & Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*(4), 827–874. https://doi.org/10.1111/1468-0262.00054

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Russell Sage Foundation.

Charness, G., Gneezy, U., & Henderson, A. (2018). Experimental methods: Measuring effort in economics experiments. *Journal of Economic Behavior & Organization*, *149*, 74–87. https://doi.org/10.1016/j.jebo.2018.02.024

Chaudhuri, A., Schotter, A., & Sopher, B. (2009). Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. *The Economic Journal*, *119*(534), 91–122. https://doi.org/10.1111/j.1468-0297.2008.02207.x

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Collins, M. G., Juvina, I., & Gluck, K. A. (2016). Cognitive model of trust dynamics predicts human behavior within and between two games of strategic interaction with computerized confederate agents. *Frontiers in Psychology*, *7*, 49. https://doi.org/10.3389/fpsyg.2016.00049

Cooper, D. J., & Van Huyck, J. (2018). Coordination and transfer. *Experimental Economics*, *21*(3), 487–512. https://doi.org/10.1007/s10683-017-9521-8

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1994). Alternative institutions for resolving coordination problems: Experimental evidence on forward induction and preplaycommunication. In J. W. Friedman (Ed.), *Problems of coordination in economic activity* (pp. 129–146). Springer.

Cooper, R., De Jong, D., Forsythe, R., & Ross, T. (1992). Communication in coordination games. *Quarterly Journal of Economics*, *107*(2), 739–771. https://doi.org/10.2307/2118488

Cooper, R. W., DeJong, D. V., Forsythe, R., & Ross, T. W. (1990). Selection criteria in coordination games: Some experimental results. *The American Economic Review*, *80*(1), 218–233. https://www.jstor.org/stable/2006744

Costa-Gomes, M. A., Crawford, V. P., & Iriberri, N. (2009). Comparing models of strategic thinking in Van Huyck, Battalio, and Beil's coordination games. *Journal of the European Economic Association*, *7*(2–3), 365–376. https://doi.org/10.1162/JEEA.2009.7.2-3.365

Devetag, G., & Ortmann, A. (2007). When and why? A critical survey on coordination failure in the laboratory. *Experimental Economics*, *10*(3), 331–344. https://doi.org/10.1007/s10683-007-9178-9

Dornic, S., Ekehammar, B., & Laaksonen, T. (1991). Tolerance for mental effort: Self-ratings related to perception, performance and personality. *Personality and Individual Differences*, *12*(3), 313–319. https://doi.org/10.1016/0191-8869(91)90118-U

Duffy, J., & Nagel, R. (1997). On the robustness of behaviour in experimental 'beauty contest'games. *The Economic Journal*, *107*(445), 1684–1700. https://doi.org/10.1111/j.1468-0297.1997.tb00075.x

Dyczewski, E. A., & Markman, K. D. (2012). General attainability beliefs moderate the motivational effects of counterfactual thinking. *Journal of Experimental Social Psychology*, *48*(5), 1217–1220. https://doi.org/10.1016/j.jesp.2012.04.016

Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005). Some like It vigorous: Measuring individual differences in the preference for and tolerance of exercise Intensity. *Journal of Sport and Exercise Psychology*, *27*(3), 350–374. https://doi.org/10.1123/jsep.27.3.350

Engelmann, D., & Normann, H. T. (2010). Maximum effort in the minimum-effort game. *Experimental Economics*, *13*(3), 249–259. https://doi.org/10.1007/s10683-010-9239-3

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*(2), 168–192. https://doi.org/10.1177/1088868308316091

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Florack, A., Keller, J., & Palcu, J. (2013). Regulatory focus in economic contexts. *Journal Of Economic Psychology*, *38*(C), 127–137. https://doi.org/10.1016/j.joep.2013.06.001

Galinksy, A. D., & Moskowitz, G. B. (2000). Counterfactuals as behavior primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology*, *36*(4), 257–383. https://doi.org/10.1006/jesp.1999.1409

Galinksy, A. D., Moskowitz, G. B., & Skurnik, I. (2000). Counterfactuals as self-generated primes: The effect of prior counterfactual activation on person perception judgments. *Social Cognition*, *18*(3), 252–280. https://doi.org/10.1521/soco.2000.18.3.252

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology*, *44*(6), 1110. https://doi.org/10.1037/0022-3514.44.6.1110

Haruvy, E., & Stahl, D. O. (2007). Equilibrium selection and bounded rationality in symmetric normal-form games. *Journal of Economic Behavior & Organization*, *62*(1), 98–119. https://doi.org/10.1016/j.jebo.2005.05.002

Ho, T. H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental" p-beauty contests". *The American Economic Review*, *88*(4), 947–969. https://www.jstor.org/stable/117013

Ho, T. H., & Weigelt, K. (1996). Task complexity, equilibrium selection, and learning: An experimental study. *Management Science*, *42*(5), 659–679. https://doi.org/10.1287/mnsc.42.5.659

Hur, T. (2001). The role of regulatory focus in activation of counterfactual thinking. *Korean Journal of Social and Personality Psychology*, *15*, 159–171.

Jackson, J. J., Wood, D., Bogg, T., Walton, K. E., Harms, P. D., & Roberts, B. W. (2010). What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC). *Journal of Research in Personality*, *44*(4), 501–511. https://doi.org/10.1016/j.jrp.2010.06.005

Juvina, I., Nador, J., Larue, O., Green, R., Harel, A., & Minnery, B. (2018). Measuring individual differences in cognitive effort avoidance. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, Madison, WI.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136. https://doi.org/10.1037/0033-295X.93.2.136

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

Kray, L. J., & Galinsky, A. D. (2003). The debiasing effect of counterfactual mind-set: Increasing the search for disconfirmatory information in group decisions. *Organizational Behavior and Human Decision Processes*, *91*(1), 69–81. https://doi.org/10.1016/S0749-5978(02)00534-4

Krueger, C., & Tian, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biological Research for Nursing*, *6*(2), 151–157. https://doi.org/10.1177/1099800404267682

Leng, A., Friesen, L., Kalayci, K., & Man, P. (2018). A minimum effort coordination game experiment in continuous time. *Experimental Economics*, *21*(3), 549–572. https://doi.org/10.1007/s10683-017-9550-3

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). Academic Press.

Malone, T. W., & Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)*, *26*(1), 87–119. https://doi.org/10.1145/174666.174668

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, *29*(1), 87–109. https://doi.org/10.1006/jesp.1993.1005

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality and Social Psychology Review*, *7*(3), 244–267. https://doi.org/10.1207/S15327957PSPR0703_04

Markman, K. D., McMullen, M. N., & Elizaga, R. A. (2008). Counterfactual thinking, persistence, and performance: A test of the reflection and evaluation model. *Journal of Experimental Social Psychology*, *44*(2), 421–428. https://doi.org/10.1016/j.jesp.2007.01.001

McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, *60*(4), 803–836. https://doi.org/10.2307/2951567

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review*, *84*(3), 658–673. https://www.jstor.org/stable/2118074

Morris, M., & Moore, P. C. (2000). The lessons we (don't) learn: Counterfactual thinking and organizational accountability after a close call. *Administrative Science Quarterly*, *45*, 737–765.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, *85*(5), 1313–1326. https://doi.org/10.2307/2667018

Nasco, S. A., & Marsh, K. L. (1999). Gaining control through counterfactual thinking. *Personality & Social Psychology Bulletin*, *25*(5), 556–568. https://doi.org/10.1177/0146167299025005002

Offerman, T. (2002). Hurting hurts more than helping helps. *European Economic Review*, *46*(8), 1423–1437. https://doi.org/10.1016/S0014-2921(01)00176-3

Riechmann, T., & Weimann, J. (2008). Competition as a coordination device: Experimental evidence from a minimum effort coordination game. *European Journal of Political Economy*, *24*(2), 437–454. https://doi.org/10.1016/j.ejpoleco.2007.09.004

Rim, S., & Summerville, A. (2014). How far to the road not taken? The effect of psychological distance on counterfactual direction. *Personality & Social Psychology Bulletin*, *40*(3), 391–401. https://doi.org/10.1177/0146167213513304

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*(5), 805–818. https://doi.org/10.1037/0022-3514.66.5.805

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*(1), 133–148. https://doi.org/10.1037/0033-2909.121.1.133

Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 1–59). Academic Press.

Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. In J. M. Olson (Ed.), *Advances in experimental social psychology* (Vol. 56, pp. 1–79). Academic Press.

Roese, N. J., & Hur, T. (1997). Affective determinants in counterfactual thinking. *Social Cognition*, *15*(4), 274–290. https://doi.org/10.1521/soco.1997.15.4.274

Roese, N. J., Hur, T., & Pennington, G. L. (1999). Counterfactual thinking and regulatory focus: Implications for action versus inaction and sufficiency versus necessity. *Journal of Personality and Social Psychology*, *77*(6), 1109. https://doi.org/10.1037/0022-3514.77.6.1109

Rubinstein, A. (1989). The electronic mail game: Strategic behavior under" Almost common knowledge". *The American Economic Review*, *79*(3), 385–391.

Rydval, O., & Ortmann, A. (2005). Loss avoidance as selection principle: Evidence from simple stag-hunt games. *Economics Letters*, *88*(1), 101–107. https://doi.org/10.1016/j.econlet.2004.12.027

Sahin, S. G., Eckel, C., & Komai, M. (2015). An experimental study of leadership institutions in collective action games. *Journal of the Economic Science Association*, *1*(1), 100–113. https://doi.org/10.1007/s40881-015-0010-6

Sanna, L. J., & Turley, K. J. (1996). Antecedents to spontaneous counterfactual thinking: Effects of expectancy violation and outcome valence. *Personality & Social Psychology Bulletin*, *22*(9), 906–919. https://doi.org/10.1177/0146167296229005

Sanna, L. J., & Turley-Ames, K. J. (2000). Counterfactual intensity. *European Journal of Social Psychology*, *30*(2), 273–296. https://doi.org/10.1002/(SICI)1099-0992(200003/04)30:2<273::AID-EJSP993>3.0.CO;2-Y

Schotter, A., Weigelt, K., & Wilson, C. (1994). A laboratory investigation of multiperson rationality and presentation effects. *Games and Economic Behavior*, *6*(3), 445–468. https://doi.org/10.1006/game.1994.1026

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527–561). Guilford Press.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. https://doi.org/10.1037/0022-3514.45.3.513

Shenhav, A., Rand, D. G., & Greene, J. D. (2017). The relationship between intertemporal choice and following the path of least resistance across choices, preferences, and beliefs. *Judgment and Decision Making*, *12*(1), 1–18. http://dx.doi.org/10.2139/ssrn.2724547

Smallman, R., & Summerville, A. (2018). Counterfactual thought in reasoning and performance. *Social and Personality Psychology Compass*, *12*(4), e12376. https://doi.org/10.1111/spc3.12376

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. Journal of Personality, *72*(2), 271–324. https://doi.org/10.1111/j.0022-3506.2004.00263.x

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*,(1), 67–85. https://psycnet.apa.org/buy/1991-32481-001

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2013). Choice architecture. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 428–439). Princeton University Press.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, *80*(1), 234–248. https://www.jstor.org/stable/2006745

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, *106*(3), 885–910. https://doi.org/10.2307/2937932

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1993). Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication. *Games and Economic Behavior*, *5*(3), 485–504. https://doi.org/10.1006/game.1993.1026

Van Huyck, J. B., Battalio, R. C., & Rankin, F. W. (2007). Evidence on learning in coordination games. *Experimental Economics*, *10*(3), 205–220. https://doi.org/10.1007/s10683-007-9175-z

Van Huyck, J. B., Gillette, A. B., & Battalio, R. C. (1992). Credible assignments in coordination games. *Games and Economic Behavior*, *4*(4), 606–626. https://doi.org/10.1016/0899-8256(92)90040-Y

Van Huyck, J. B., Wildenthal, J. M., & Battalio, R. C. (2002). Tacit cooperation, strategic uncertainty, and coordination failure: Evidence from repeated dominance solvable games. *Games and Economic Behavior*, *38*(1), 156–175. https://doi.org/10.1006/game.2001.0860

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*(6), 516–523. https://doi.org/10.1111/j.1467-9280.2007.01932.x

Weber, R., Camerer, C., Rottenstreich, Y., & Knez, M. (2001). The illusion of leadership: Misattribution of cause in coordination games. *Organization Science*, *12*(5), 582–598. https://doi.org/10.1287/orsc.12.5.582.10090

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *Plos One*, *9*(7), e68210.

White, K., & Lehman, D. R. (2005). Looking on the bright side: Downward counterfactual thinking in response to negative life events. *Personality and Social Psychology Bulletin*, *31*(10), 1413–1424. https://doi.org/10.1177/0146167205276064