PERCEPTION AND DISTRIBUTION OF EFFORT IN JOINT ACTION

VJERAN KERIĆ

SUPERVISOR: Natalie Sebanz

SECONDARY SUPERVISOR: Günther Knoblich

Thesis submitted for the degree of Doctor of Philosophy (PhD)

Department of Cognitive Science, Central European University

Submitted: Vienna, May 1, 2023

[this page intentionally left blank]

Declaration of Authorship

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or which have been accepted for the award of any other degree or diploma at Central European University or any other educational institution, except where due acknowledgment is made in the form of bibliographical reference. The present thesis includes work that appears in the following papers:

Keric, V., & Sebanz, N. (2021). Task construal influences estimations of the environment. *Frontiers in Human Neuroscience*, *15*, 625193.

Vjeran Kerić

Abstract

When it comes to effort, whether mental or physical, people and non-human animals tend to rationalize its expenditure. This is the case both when they act individually and when it comes to humans, cooperatively. The aim of this thesis was to investigate how effort expenditure reflects on collective action, action prediction and joint decision-making. Chapter 2 investigated how energy expenditure influences perceptual estimates of the environment. It showed that perception and effort might be linked but that the nature of this connection is social. More precisely, whether potential effort influences perception is tied to construal of a situation more than it is to energetic costs of an action. The third chapter studied whether distribution of mental effort is rational from a joint perspective, showing that mental effort distribution strategies are mostly focused on minimization of one's own effort rather than the total amount of effort expended by all participants of an interaction. These results demonstrate that compared to sharing of physical costs, collaborating in tasks involving mental effort are less likely to involve jointly rational strategies of effort distribution. The final empirical Chapter (Study 3) tested whether people expect jointly rational distribution of effort and if this informs prediction of ongoing actions. These experiments revealed that while explicit expectations were in line with joint rationality, online prediction of actions was more informed by immediate individual efficiency considerations. However, these considerations were equally pronounced for the acting agent and their partner, suggesting that prediction was informed, in equal measure, by the efficiency of an acting agent and their partner. Taken together, the findings of this thesis show that considerations of other's effort (motor or cognitive) guide both action prediction and decision-making but that in both cases neither decisions nor predictions are completely in line with joint rationality.

Acknowledgments

First and foremost, I want to thank my primary supervisor, supervisors, Natalie Sebanz and Günther Knoblich. Natalie, my primary supervisor, has guided me through this process with patience, understanding and an ever-present encouragement to engage with complex subjects. I am truly grateful and honored to have been able to learn from both of you.

This thesis includes a collaboration with Patric Bach and Katrina McDonough at the University of Aberdeen. Thank you Patric and Katrina for interesting discussions and for making Scotland a home away from home for a while. You were a pleasure to work with.

I also want to thank everyone at the SOMBY lab for the discussions we had, many of which taught me what science is and should be about. In particular, I would like to thank Arianna Curioni for a lot of helpful feedback on my work and Gina Török for sharing her insights, data and materials. Furthermore, I am very grateful to Andrea Jenei as well as the past and present lab managers: Fruzsina Kollányi, Vanda Derzsi, Dávid Csűrös, Fanni Takátsy and Candasch Acar.

I want to thank Reka for helping with numerous administrative issues. The level of intellectual curiosity, openness and support at the Department humbles me and I cannot be anything but grateful to have been a part of it. I want to thank everyone for creating this environment. I want to thank my friends and colleagues without which this experience would have been purely an academic exercise: Barbu, Levi, Tamas, Atsuko, Eszter, Gina, Laura, Nima, Francesca, Nazli, Luke, Martin and at least a few who I forgot to mention but hope they will forgive me.

For keeping me sane and entertained, thank you to my friends back home. Domagoj and the rest. You know who you are. And to friends at my new home, you know what you mean to me.

My parents deserve a big thank you for all the years they supported me throughout this process. Finally, I want to thank Dora, whose patience and support I cannot overstate. Your unquestioning belief in me is a very rare thing to have in one's life.

Table of Contents

1 Chapter 1: Introduction

Why do it the easy way when you can do it the hard way? The bit of humor in this question is that it turns our preference of how to go about things on its head. This accessible intuition can also be formulated in more technical terms: if two behaviors yield the same desired outcome but involve different degrees of effort, organisms will choose (or learn to choose) the least effortful one (Hull, 1947). A wide range of studies show that avoiding costs (not necessarily energetic ones, but also optimizing time and cognitive resources) is ubiquitous in the animal kingdom. When choosing a course of action, rodents and monkeys will decide based on the reward relative to the effort required to obtain it (Walton et al., 2006) and rats escaping a maze will prefer to press levers that require less force (Thompson, 1944). The behavior of many animals can be modelled as a tradeoff between time constraints and energy gains and losses (Hughes et al., 1992). For example, bison tend to minimize time needed to forage at the expense of energy (Bergman et al., 2001).

Much like other animals, humans also tend to minimize effort. In the motor domain in particular, it has been shown that action execution leads to biomechanically and energetically efficient paths (Anderson & Pandy, 2001; Cos et al., 2011; 2014; Todorov & Jordan, 2002; Todorov, 2004). These paths may deviate from optimality if they would instead afford comfortable end-states (Rosenbaum et al. 1990; Rosenbaum & Jorgensen, 1992). Minimization of costs is not limited to motor action. When choosing from two decks of cards, one containing more difficult cognitive tasks than the other, participants quickly started preferring to draw from the easier deck (Kool et al., 2010). This preference is not surprising given that engaging in mental effort has often been described as laborious (Dreisbach & Fischer, 2015; Inzlicht et al., 2018; Kurzban, 2016).

While we avoid both physical and cognitive costs, it can often happen that we have to incur one or the other to a certain degree. Rosenbaum and colleagues (2014) showed that when people have a choice of picking up a heavy object immediately and carry it across a distance or pick it up later along their path, they preferred to pick it up immediately. These findings are interpreted in terms of a trade-off between physical and cognitive resources – picking the object up later requires keeping in memory that it needs to be picked up. Along similar lines, Potts and colleagues (2018) presented people with a series of physical and cognitive tasks. After familiarizing themselves with the possible tasks, participants could choose which ones they wanted to complete. Results showed that while mental and physical effort played a role, choices were most strongly affected by subjective time – the time that participants judged it took to complete a task. This suggests that different types of costs are intrinsically linked, and people weigh one against the other when making decisions. The weighing process can reflect the amount of the cost, individual preferences, but also individual capabilities to complete the task at hand.

1.1 Effort and individual capabilities

When it comes to acting upon the environment, not every species, or individual organism within the same species, has the same capabilities. The visual system is adept at detecting affordances – actions that the environment enables us to do, given our physical constraints (Gibson, 1979). Evidence for visual perception priming action execution has been found in a wide range of studies (e.g., Craighero et al. 1996; Symes et al., 2005; Tucker & Ellis, 2001). For example, seeing a pan with its handle pointing towards the left will facilitate grasps with the left hand (Tucker & Ellis, 2001) but not if the handle is broken (Buccino et al., 2009) or the pan is out of reach (Cardellicchio et al., 2011). The relationship between perception and potential action goes both ways; action capabilities inform perceptual estimates of the environment. Proficient archers judge targets as larger than their less precise counterparts (Lee et al., 2009), shorter people judge the same flight of stairs as steeper compared to taller people (Warren, 1984) while judgments about the width of an opening correlate with one's shoulder width (Franchack et al., 2012; Warren & Whang, 1987). An influential proposal to account for these findings is that the perceptual system uses the body as a ruler with which it measures the environment (Proffitt 2006; Proffitt &

Linkenauger, 2013), informing the perceiver about action possibilities and the associated energetic cost. This has been supported by findings showing that hills are judged as steeper and distances as longer if the person judging was carrying a heavy object, reflecting the additional energetic cost required to climb the hill or traverse the distance (e.g., Bhalla |& Proffitt, 1999; Proffitt et al., 1995). Additional evidence comes from studies investigating how the visuomotor system updates percepts in light of incongruency between visual information and effort (Proffitt et al., 2003; Zadra & Proffitt, 2016). In these experiments, participants walked on a treadmill while equipped with VR goggles projecting either a stationary environment, or one which changed at a rate congruent with participants' walking speed. Participants who experienced a stationary image of the environment or a visual flow slower than their walking speed consistently overestimated distances in subsequent perceptual judgments. The authors proposed that walking coupled with a stationary visual flow recalibrated how the visuomotor system computes the relationship between effort and perceptual information. In other words, the motor system adapted to a state where more energy than normal is required in order for visual information to be consistent with standing in place. Taken together, these studies show that the extent of exerted effort varies based on individual capabilities and can influence how the environment is perceived. However, these findings were criticized as arising from participants guessing the purpose of the manipulation and adjusting their estimates in accordance with the hypothesis of the experiment (Durgin et al., 2009; 2012; Firestone, 2013; Firestone and Scholl, 2014). For example, when participants were told to ignore the backpack, Durgin and colleagues (2012) found that its weight did not influence perceptual estimates. A reply to this particular finding has been that instructing participants to ignore the backpack is a manipulation in itself and might interfere with perceptual judgments (Clore and Proffitt, 2016; Proffit, 2013; Witt, 2011; Witt et al., 2016). It remains an open question if changes in how the environment is perceived would be found using more subtle cover stories which do not interfere with the task during judgments but rather influence how people frame the task.

1.2 Joint action and joint effort minimization

Many of our daily actions are done with other people. The breadth of these interactions spans from simple interactions like passing a cup of coffee to much more complex (but still everyday) tasks such as a family coordinating moving apartments. When we engage in a task with others, we represent both our own part of the interaction and that of our partner. For example, when stimulus-response compatibility tasks are split between two participants, similar compatibility effects appear as if the task was done by a single participant (e.g., Atmaca et al., 2011; Sebanz et al., 2003; 2006). This shows that people represent aspects of the task of their co-actor even when it has no practical relevance for their own. Along similar lines, when a person is discriminating between local features of an object while their partner is doing the same for global features, a match between global and local features will lead to better performance while a mismatch will lead to slowdowns and more errors for both participants of the task. In contrast, when discriminating just one dimension of the features without a partner, such facilitation and interference does not occur (Böckler et al., 2011). This pattern of representing the partner's part of the task (*corepresentation*) has been demonstrated for different cognitive capacities, across a wide range of studies. When remembering a series of items in pairs, recall was enhanced for the partner's items even when they were irrelevant for the task of the individual participant (Elekes et al., 2016). In the same vein, people spontaneously take the visuospatial perspective of their co-actors regardless of whether the task at hand requires it.

Cooperating in a task also involves taking into account constraints of one's co-actors (Wel & Fu, 2015). A number of studies show that we can easily and consistently detect actions the environment affords to other people. For example, three-year-old children can differentiate reaching affordances for themselves and adults (Rochat, 1995). Stoffregen et al. (1999) demonstrated that observers can accurately judge the preferred sitting height of another person. Mark (2007) replicated these results and obtained similar findings for affordances of climbing, reaching, and stepping over gaps. In a study by Chang et al. (2009) participants walking next to a child judged the minimum width of an aperture needed for them to pass through. Estimates were scaled to the combined shoulder width of the adult and the child. This suggests that participants perceived affordances for the child and integrated them with their own. In a more direct test of representing constraints of partners in an interaction, Schmitz et al. (2017) tasked participants to coordinate reaching movements between two targets. In a condition where one of the participants had an obstacle placed between the targets, hand movement trajectories of the task partner were higher, mirroring the movements of their partner who had an obstacle in their way.

These studies show that we take into account our own affordances and those of our partners and integrate the constraints and the rules of our respective tasks when they are shared, forming a basis for reducing collective effort in interactions. For example, when asked to temporally synchronize their actions, participants started adopting strategies that helped their partner adapt; in particular, they reduced the variability of their actions to help their partner coordinate (Vesper et al., 2011). In another study, dyads were tasked, without being able to see each other, to jump together so that they would land at the same time. The individual who had to jump the shorter distance started to adjust their jump timing to their partner (Vesper et al., 2013). Similar findings showed that people tend to take additional effort upon themselves to decrease the effort of their partner(s) when passing objects (Constable et al., 2016; Gonzalez et al., 2011; Meyer et al., 2013) or holding doors open for others (Santamaria and Rosenbaum 2011). Similarly to the findings of potential effort influencing estimates of the environment (Proffitt 2006; Proffitt & Linkenauger, 2013), expecting to lift objects with others makes people judge them as lighter than when lifting them on their own (Doerrfeld et al., 2012). When it comes to sharing mental effort, people are more willing to take on the more difficult part of the task themselves if they believe that their partner is under a heavy cognitive load, less competent or would have to adopt a difficult visuospatial perspective (Duran et al., 2011; Galati et al. 2013; Mainwaring et al., 2009; Schober

2009).

Other studies investigated whether sharing of effort follows the norms of joint rationality (also termed "co-efficiency"). From a jointly rational perspective, whether one should decrease their own or their partner's effort in an interaction depends on the total reduction of effort resulting from the decision. This can mean reducing one's own effort but, also, increasing one's own effort if it would decrease the effort of the partner to a greater extent. Török and colleagues (2019) used a sequential task in which each participant had to move an object part of the way towards a final position. People had the choice to take the longer or shorter path for their part of the task. When taking the somewhat longer path would reduce the partner's path to a greater extent, people opted to take the longer path. Conversely, when the overall path (partner's + their own) was shorter if they would take the individually shorter path, they opted for the shorter path. A subsequent study showed that participants' decision-making process placed equal weight on their own and their partner's path length (Török et al., 2021). Using a similar setup to investigate the development of expectations about joint efficiency, Mascaro and Csibra (2022) showed that when observing an interaction such as this, infants are surprised if the agents do not take jointly efficient paths.

Taken together, these studies show that we integrate others' perspectives, task constraints, and lower a partner's effort. A question that remains open is whether people make co-efficient decisions when distributing cognitive effort. Studies in perspective taking and communication show that people are willing to take more effort upon themselves if their partner is not competent at a task or has a high cognitive load (Schober 2009; Mainwaring et al., 2009; Duran et al., 2012; Galati et al., 2013). However, assisting a collocutor does not necessarily mean that the interaction conforms to joint rationality. A more direct measure of sharing cognitive costs is needed to establish whether people employ strategies to minimize mental effort across the dyad with equal weight placed on one's own and their partner's costs.

1.3 Expectations of sharing effort and action prediction

The fact that we see affordances for others, and we minimize our own and others' effort informs our understanding of others' actions and generates expectations of how they might behave in a given situation (Dennett, 1987). We expect others to behave rationally, assuming they will take into account their abilities and minimize costs when trying to reach their goals. Expectations of efficient or rational action arise early in childhood (e.g., Csibra et al., 1999; Gergely et al., 1995; Liu & Spelke, 2017) and have been observed in other species (e.g., Rochat et al., 2008). These expectations have been shown to inform predictions of trajectories of ongoing actions. A series of studies have shown that when observing a moving hand, the location of the hand is represented as closer to efficient trajectories than inefficient ones (Hudson et al., 2018; McDonough et al., 2019; 2020; McDonough & Bach, 2022). The authors interpret these results in terms of expectations of efficiency informing prediction which in turn biases perceptual estimates towards efficient trajectories.

Many everyday actions require on-line prediction of our partners' movements. Previous research shows that expectations of individual efficiency guide online prediction, as well as that partners in interactions act in jointly efficient ways. It remains an open question whether observers have expectations of joint efficiency and whether such expectations inform predictions of actions in observed ongoing interactions.

1.4 Research questions

The aim of this dissertation was to investigate how people represent own and others' effort, how this affects their perception of the environment, the way they distribute tasks, and their expectations of joint efficiency.

Study 1 (Chapter 2) aimed to investigate whether, when engaging in an interaction, people represent the physical effort of their partner in a similar way that they co-represent a partner's visuo-spatial perspective, physical constraints, or response rules (Freundlieb 2016; 2017; Schmitz et al., 2017; Sebanz et al., 2003; 2006). If effort reflects energetic costs and is co-represented, perceptual estimates of the environment should reflect the energetic costs and abilities of both agents in an interaction (Bhalla & Proffitt, 1999; Proffitt et al., 1995; Proffitt 2006; Proffitt & Linkenauger, 2013). For example, if two people are engaged in a cooperative that requires traversing distances and one of the people has a heavy backpack, the person without the backpack would estimate the distance as longer than if their partner was unencumbered. As a first step, prior to investigating corepresentation of effort, three experiments were conducted to ensure that effects of effort on perception can reliably be observed in individuals expecting to perform instructed actions. Since a prominent criticism of these findings has been that participants guessed the purpose of the effort manipulation which might have influenced the results (e.g. Firestone, 2013; Firestone and Scholl, 2014), the current experiments used elaborate cover stories. Furthermore, to prevent cover stories from eliminating the potential effect of effort on perception (a possibility suggested by Clore and Proffitt, 2016; Proffit, 2013; Witt, 2011; Witt et al., 2016), the cover stories were employed before the experiment and did not require participants to attend to any feature of the cover story while making judgments. Experiments 1-3 used cover stories which successfully deceived participants about the main hypothesis and the results showed no influence of potential energetic costs on estimates of height or distance. Experiment 4 followed the same procedure as Experiment 3 except that no cover story was used. Debriefing confirmed that all the participants correctly guessed the purpose of the manipulation and the results showed that distances were judged as longer in the high-effort condition compared to the low-effort condition. Since the main difference between Experiments 3 and 4 was how participants construed the task, it seems likely that the results of Experiment 4 were due to task demands rather than potential energy expenditure influencing perception. These results cast doubt on whether effort can influence perceptual judgment in individuals, which makes the possibility of representing costs of a partner influencing one's own perception seems even more remote. For this reason, the study in Chapter 2 was not extended to the domain of joint action.

Study 2 (Chapter 3) investigated whether mental effort is divided between co-actors following principles of joint rationality. Previous research on joint action planning has approached co-efficiency from a physical or motor perspective (e.g., Török et al., 2019). On the other hand, studies on sharing of mental effort generally concentrated on whether one agent is willing to take more effort upon themselves in order to make their partner's task easier (e.g., Schober 2009), while not directly investigating if people tend to choose coefficient strategies. Chapter 3 reports six experiments that used multiple object tracking or working memory tasks to elicit cognitive costs. Participants had different choices as to how to distribute the associated effort between themselves and their partner. Results showed minimization of both own and the partner's effort, but participants did not place equal weight on both.

Study 3 (Chapter 4) tested if observers of interactions have expectations of joint efficiency and whether these expectations inform prediction of trajectories of ongoing actions. In three experiments, participants saw the layout of a game played by two players. The players had to lead a ball through a series of obstacles. The first player would bring the ball to the center of the screen after which the second player would pick it up and lead it to the end. The choice of route by the first player could be either individually or jointly efficient or inefficient. In some of the trials, the ball would disappear, and participants had to report the last seen position of the ball. Results showed explicit expectations of joint efficiency for both agents. These expectations guided predictions of where the first player was leading the ball, suggesting that action prediction is informed by expectations that agents will minimize both their own and their partner's costs.

However, this expectation led predictions of minimization of each of the agent's effort separately, rather than taking into account the difficulty of a given trial across both agents.

The final chapter (Chapter 5) discusses the findings as well as open questions for future research and draws conclusions concerning the role of mental and physical effort in decision-making and action prediction.

2 Chapter 2: Task Construal Influences Estimations of the Environment

Keric, V., & Sebanz, N.

2.1 Introduction

The way humans perceive, act, and think is shaped by properties of the body and its possibilities for interacting with the environment. Long before Embodied Cognition approaches became prominent (for reviews, see e.g., Wilson, 2002; Shapiro, 2011; Galetzka, 2017), Gibson (1979) argued that the visual system is geared toward perceiving properties of the world and objects within it that allow organisms to interact with the environment in a particular way ("affordances"). A roof affords hiding from the rain, a chair affords sitting, and a pencil affords grasping with a precision grip. The perception of these properties is based on the relations between a specific agent and the world; a steep cliff affords different actions to mountain goats and humans. People are adept at judging which actions are at their disposal. While Gibson focused on "direct perception," a large number of studies suggests that perceiving affordances is tightly coupled with action execution (e.g.,, Craighero et al., 1996; Tucker and Ellis, 1998, 2001; Buccino et al., 2009; Cardellicchio et al., 2011; Janyan and Slavcheva, 2012).

There is an increasing amount of evidence suggesting that the way we perceive the environment is influenced by our capabilities. For example, the same object will be judged as closer if one has a tool extending one's reach (Witt et al., 2005). Precise archers estimate targets as larger (Lee et al., 2012) compared to imprecise ones, proficient jumpers judge distances as shorter (Lessard et al., 2009) and participants trained in parkour see walls as shorter than those with little experience in it (Taylor et al., 2011). These results suggest that the properties and capabilities of our bodies are a "ruler" of sorts against which the environment is measured.

Why would this be the case? According to the interface theory of perception (Hoffman et al., 2015), natural selection has shaped perception so that it guides adaptive behavior. Relatedly, Proffitt (2006) and Proffitt and Linkenauger (2013) posit that a crucial role of perception is to inform efficient action. In this view, an action is “efficient” if it completes a task with the least energy expenditure. It is assumed that the perceptual system evolved in a way that gave our ancestors an advantage, and one of the greatest advantages of all is optimization of energy consumption. According to this view, an important role of perception is to integrate information and to provide an agent with a view of the world with all contextual considerations already factored in. The central prediction is that energetically expensive actions are coupled with percepts that overestimate the features of the environment that need to be overcome in order to execute these actions.

This prediction has been supported by a number of studies, mostly focusing on how metabolic states and energy expenditure influence perceptual judgments about distances and the steepness of surfaces. Slants of hills are judged as steeper when the participant making the judgment is encumbered by a heavy backpack, in poor physical shape, tired from previous physical activity or old rather than young (Proffitt et al., 1995, Proffitt et al., 2003; Bhalla and Proffitt, 1999). Hills that are very difficult to climb down from but manageable to climb up on seem steeper when viewed from the top as opposed to the bottom (Proffitt et al., 1995). Furthermore, if perception is tied to bodily states, metabolic changes should influence perceptual judgments. In a series of experiments, Schnall et al. (2010) gave one group of participants a caloric drink before they made slant judgments. Participants who consumed the caloric drink judged slopes as gentler compared to controls who drank a placebo. These findings were replicated for judgments of distance perception (Zadra et al., 2016).

While there is substantial evidence for energetic costs influencing estimates of distance and slope, including a recent meta-study confirming that effort influences distance estimates (Molto et al., 2020), others have criticized these findings, proposing that many of the described studies investigated biases in judgment rather than perception (Firestone and Scholl, 2016). For example, when estimating slants while carrying a heavy backpack, participants could have easily deduced why the backpack was introduced and then intentionally or unintentionally adjusted their judgment, so that the measured effect could have been due to task demands rather than increased effort biasing participants' estimates. In order to test for this possibility, Durgin et al. (2009) conducted a study in which carrying the backpack was embedded in an elaborate story. Participants were fitted with electrodes around their ankles and were told that the backpack contains electromyography equipment needed to measure muscle activation. The cover story eliminated overestimation and participants who were convinced by it judged slopes to be equally steep as participants not carrying a backpack. This suggests that the effects in other studies could have been due to participants inferring the goal of the backpack manipulation. In a later study, Durgin et al. (2012) found that if participants carrying a heavy backpack were informed about the role of the backpack and asked to ignore it, they did not estimate the slant of a hill differently than control participants not carrying a backpack.

One criticism of the experiments by Durgin et al. (2009, 2012) is that they introduced task demands of their own. More specifically, in Durgin et al. (2009) the cover story included carrying a noisy backpack, which might have drawn participants' attention to the backpack, while in Durgin et al. (2012) participants were explicitly told to ignore the backpack. It is possible that explicitly ignoring the weight of a backpack or attending to a noisy one interferes with the heuristic that scales distance and slope estimates with potential energy expenditure (Clore and Proffitt, 2016). Furthermore, the explicit instruction to ignore a heavy backpack might have biased participants' estimates in the opposite direction (Proffit, 2013; Witt, 2011; Witt et al., 2016).

In the present study, we investigated whether effort influences perceptual estimates when effective cover stories are employed that are neither distracting nor overly salient. In four experiments we investigated whether effort influences estimates of height (Experiment 1) and distance (Experiments 2–4). In Experiment 1, participants were asked to place an object on a shelf, and then estimate its height. Effort was manipulated by varying the height of the shelf as well as the weight of the object. The aim was to test if previous findings of effort influencing estimates also hold for vertical distance. Experiment 2 tested whether effort influences distance estimates in a novel task where effort was manipulated by varying the difficulty of locomotion. After giving estimates about the distance of a target, participants moved to the target by either walking ("low effort" condition) or by hopping on one leg ("high effort" condition). The goal was both to conceptually replicate Proffitt et al. (2003), as well as to see if a different kind of physical effort can bring about the effect. Since neither Experiment 1 nor Experiment 2 provided evidence for effects of effort on perception, we conducted two further experiments that included conditions closer to the ones studied by Proffitt et al. (2003). In Experiment 3, effort was manipulated by carrying a light or heavy backpack, followed by estimates of the traversed distance. This experiment still did not show the expected effects of anticipated effort on perception. However, anticipated effort modulated distance judgments in Experiment 4, where the cover story was dropped. We discuss these findings with regard to the role of social influences on perceptual judgments and consider implications for ongoing debates on the cognitive penetrability of perception.

2.2 Experiment 1: lifting heavy objects

The aim of the first experiment was to provide a conceptual replication of previous findings showing that distance estimates are modulated by effort. Participants were asked to lift objects of different weight and place them on shelves of various heights. The prediction that effort should influence height estimates follows from previous research. If effort changes height estimates in a similar way as those of distance (Proffitt et al., 2003; Josa et al., 2019) and slant (Bhalla and Proffitt, 1999), lifting heavier weights should increase perceived height. Furthermore, it is possible that placing weights on higher surfaces, requiring more effort, would influence estimates more strongly than placing them on lower ones. The experiment employed both an explicit verbal measure and a non-verbal one in which participants were asked to mark the height of the surface on a schematic drawing representing all the possible heights. The reasoning behind this was that similar results across different measures would provide stronger evidence that effort influenced estimates. In order to ensure that potential differences in estimates were not due to task demands, a cover story was employed. A final consideration relates to anticipating to act rather than giving estimates per se. It has been suggested that the intention to execute an action may be a precondition for effort influencing estimates (Witt et al., 2004). Therefore, in this experiment participants knew they would execute the actions after having given their estimates.

2.2.1 Materials and Methods

Participants

Nineteen right-handed participants (mean age = 26.76, SD = 2.56, 14 female) signed up for this study and received gift vouchers (1500 HUF) for their participation. All participants had normal or corrected-to-normal vision, were naive to the purpose of the study, signed a consent form before testing began and were debriefed at the conclusion of the experiment. The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and was conducted in accordance with the Declaration of Helsinki (1991).

Design and Apparatus

The experiment manipulated effort by varying the weight of an object and the height of a shelf on which it had to be placed. The three weights (see Figure 1) used were (1) an empty base of an exercise weight, (2) the base with an added 1.7 kg, and (3) the base with an added 3.4 kg (coded as light, medium and heavy). A shelf of adjustable height (in 5 cm increments) was used for the height manipulation (see Figure 1). The heights were divided into three ranges: low, medium and high. This resulted in a 3 × 3 within-subject design with variables of Weight (light, medium and heavy) and Height (low, medium and high).

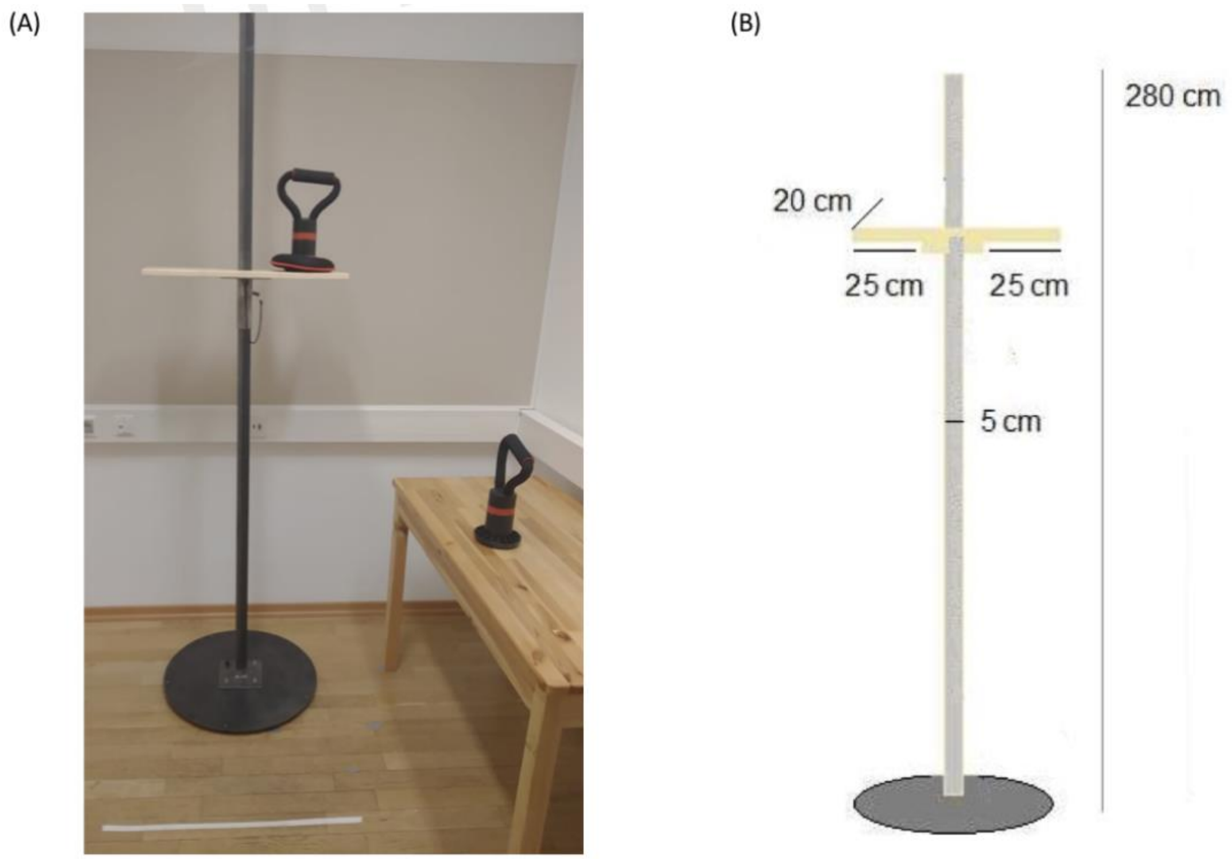

Figure 1. Experimental setup with two example weights (light and medium) marked with horizontal red stripes where participants were asked to place their thumb while grasping the weight (A). Dimensions of the shelf and pole (B).

Two dependent measurements were taken; verbal and non- verbal. Verbal estimates were reported in terms of shelf height in centimeters. For the non-verbal measure, participants marked the height of the shelf on a schematic representation of the pole that held the shelves.

Procedure

Before testing began, participants read an information sheet and signed consent forms. As part of the cover story, they were told that the apparatus would be used in future experiments and that the purpose of the present study was to test shelf stability under different weights and whether the shelves allowed for precise height estimates.

Participants were positioned in front of the apparatus so that the pole holding the shelves was aligned to their body midline. They were instructed to stand between 10 and 40 cm from the pole, depending on where they felt most comfortable to place objects on the shelf. Once a comfortable distance was determined, it was kept constant throughout the experiment. A trial started with the participant giving a verbal estimate of the height of the shelf and by marking the height on a line representing the height of the apparatus (shelf to line scale was 1:14). Participants then took a weight from a desk on their right-hand side, placed it on the shelf, lowered their hand in a resting position next to their body, and then returned the weight to the desk. In order to make lifting the weights more difficult, participants were instructed to grasp the body of the weight rather than the handle (see Figure 1). While the participant was returning the weight to its original place, the experimenter adjusted the shelf for the next trial. The experiment took 20–25 min to complete.

Before testing, participants went through a few (3–5) practice trials. Once they indicated that they understood the procedure, the testing phase consisting of 27 trials began. Trials were blocked so that each participant consecutively went through 9 trials per weight, and the blocks were counterbalanced across participants, creating 6 counterbalancing orders. Between blocks, the experimenter asked the participant to grasp the weight that will be used in the next block and lift it over their shoulder. This was done in order for the participants to get a sense of the weight, but they were told that the experimenter was making sure they were holding the new weight correctly. The height of the shelf was divided into three ranges; low (80–100 cm), medium (10–125 cm), and high (130–150 cm) with participants going through 9 trials in each range. The exact heights in each range and the presentation order were randomized. At the conclusion of testing, participants' maximum reaching height was measured; they were asked to stand so that their toes were touching the base of the pole and to place their palm on the shelf. To check if participants guessed the purpose of the weight manipulation, they were asked whether they had thought about why we used different weights. This was followed by a debriefing where the purpose of the study and the weight manipulation were explained. Finally, participants were asked if during testing it had occurred to them that the weights were intended to influence their height estimates. If they answered affirmatively to the question or mentioned the relationship between weights and height estimates in the previous, open-ended question, they were excluded from the analysis.

2.2.2 Results

As shelf heights were selected randomly from predefined ranges, a preliminary analysis (one-way ANOVA) was run to confirm that there were no differences in assigned heights between the three weight conditions ($p = 0.994$), counterbalancing groupings ($p = 0.997$) or individual participants ($p = 0.999$). One participant correctly guessed the purpose of the study and was excluded from the analysis.

Verbal Responses

Correlational analysis suggested a good overall performance on the task (see Supplementary Materials). A 3 × 3 × 6 mixed ANOVA was conducted with the within-subject factors of Weight (light, medium, heavy) and Height (low, medium, high) and counterbalancing order as a between-subject factor. The counterbalancing order was included in the analysis in order to control for carry-over effects. Previous studies showed that fatigue can influence estimates (Bhalla and Proffitt, 1999; Hunt et al., 2017) which could have created a situation in which participants' estimates were influenced by a previous block of trials (e.g.,, overestimating heights in a "light" condition due to fatigue from previously completing "heavy" trials). Participants' maximum reach was mean-centered and included in the analysis as a covariate. As Mauchly's test revealed that a violation of the sphericity assumption occurred [$\chi2(2) = 6.86$, $p < 0.05$], degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.668$). The results showed an expected main effect of height [$F(1.34,14.7) = 30.06$, $p < 0.001$, $\eta2p = 0.732$] and no significant effect of weight or interactions (ps > 0.05). Counterbalancing order did not have a significant effect or significant interactions (ps > 0.05) and maximum reach was not a significant covariate ($p > 0.05$). For the distribution of verbal responses, see Figure 2A.

Non-verbal Responses

Correlational analysis suggested good performance, while also indicating that the non-verbal task was possibly more difficult than the verbal one (see Supplementary Materials). A 3 × 3 × 6 mixed ANOVA was performed with the within-subject factors of Weight (light, medium, heavy) and Height (low, medium, high), the between-subject factor of counterbalancing order, and maximum reaching height as a covariate. Greenhouse-Geisser correction was applied ($\varepsilon = 0.719$) because Mauchly's test revealed that the sphericity assumption was violated [$\chi2(2) = 7.91$, $p < 0.05$]. The results showed a significant main effect of height [$F(1.44,14.29) = 5.6$, $p < 0.05$, $\eta2p = 0.338$] and no other significant main effects or interactions (all ps > 0.05). For the distribution of non-verbal responses, see Figure 2B.

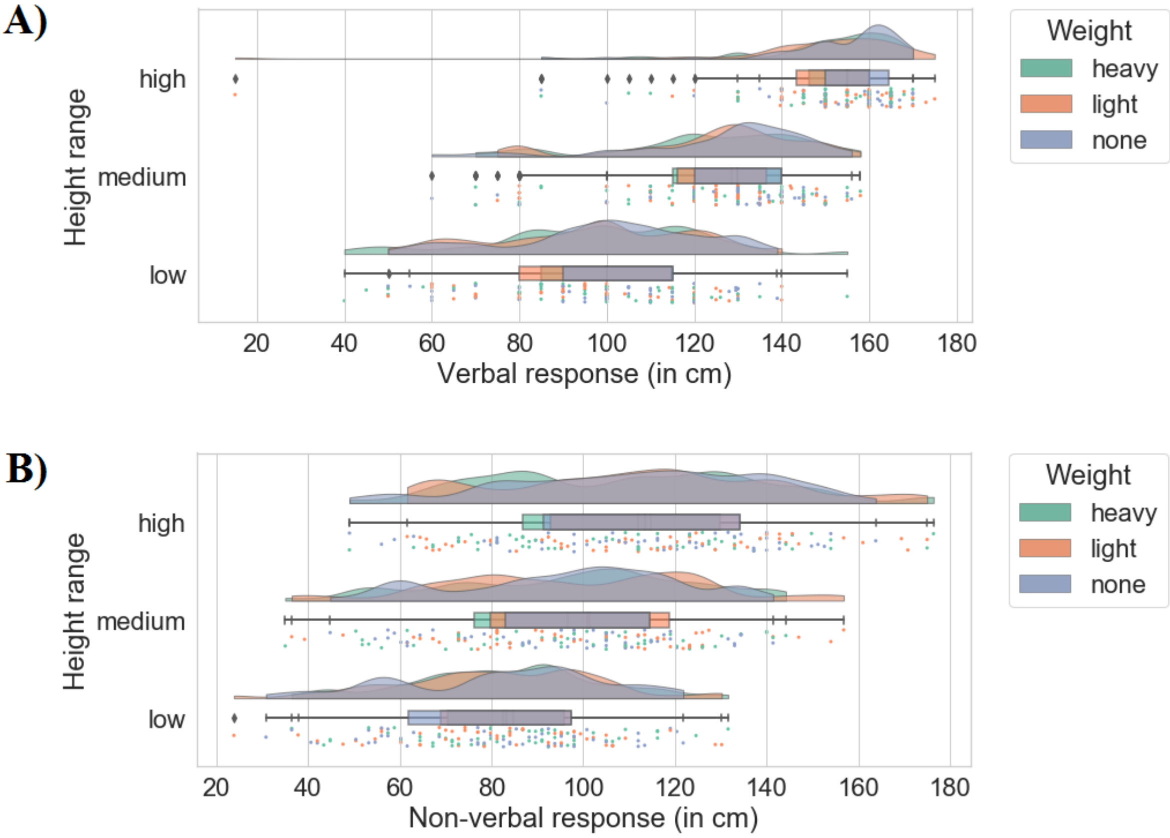

Figure 2. Participants' verbal (A). and non-verbal responses (B) in Experiment 1.Superimposed distributions show responses in different weight conditions under which individual data points are vertically jittered. Boxes indicate interquartile range with median, whiskers show 1.5 interquartile range (modified from Allen et al., 2019)

Additional analyses

In order to test whether the data support the null hypothesis that weight does not influence judgment estimates, the data were broken down by weight and a series of Bayesian paired-samples t-tests were conducted using a built-in function of JASP. In all tests, H0 stated that the effect size is $\delta = 0$ while H1 assigned effect size a Cauchy prior centered on 0 with the interquartile range of $r = 0.707$. For verbal responses, comparing the influence of heavy and medium weight on participants' estimates showed strong support for the null hypothesis (BF01 = 11.32). Comparisons between heavy and light weight received moderate support (BF01 = 7.82) as did those of light and medium weight (BF01 = 7.26). For non-verbal responses, the null hypothesis was strongly supported in all three comparisons showing BF01 ≈ 11 (see Supplementary Material for a more detailed analysis).

2.2.3 Discussion

Manipulating effort did not seem to influence participants' height estimates, regardless of the height range of the shelf they placed the objects on. For both verbal and non-verbal estimates, this conclusion is further supported by supplementary Bayesian analyses showing the data moderately to-strongly favor the null hypothesis with Bayes factors in favor of H0 ranging from BF01 ≈ 7 for verbal estimates to BF01 ≈ 11 for non- verbal ones. One possibility could be that the heights were too easy to judge. Perhaps computing effort may not have been necessary for estimations because participants were using their own height and reach to judge height. However, looking at the standard deviations in each condition (see Supplementary Table 1) and taking into account that participants misjudged the actual shelf height by 10 cm in verbal responses (Mactual = 114, Mestimated = 124.8), this does not seem very likely.

One open question is whether the effort manipulation should have been implemented taking into account participants' physical characteristics. In the current study all participants, regardless of height, weight or fitness lifted the same weights. Most reported fatigue during the experiment or the debriefing but several said they thought the weights were meant to make it easier to place the weight holder on the shelf. While based on these reports it seems unlikely that the task was not sufficiently effortful, without adjusting the weights for each participant this possibility cannot be completely discounted.

Finally, it is possible that the results speak in favor of perceived task demands rather than effort influencing estimates. The cover story could have eliminated otherwise more obvious task demands, leading to physical effort having no effect on perceptual estimates of height. Taken together, the results of Experiment 1 did not provide evidence for the claim that energy expenditure influences estimates with regard to judgments of height, contrary to previous work reporting effects of effort on judgments of distance (Proffitt, 2006; Proffitt and Linkenauger, 2013).

2.3 Experiment 2: walking and hopping across distances

This experiment followed a similar logic as Experiment 1, with two key differences. Firstly, instead of estimating heights, participants were asked to estimate distances. The reasoning was that previous studies have established that effort influences distance estimates whereas there is no available data on whether this is the case for heights. Secondly, instead of varying weight, the effort manipulation was operationalized by asking participants to cross distances by walking vs. hopping on one leg. Since effects of weight on distance estimates have been reported, we wanted to test whether other effort manipulations would yield similar effects. The main prediction was that the additional effort involved in hopping to a certain location compared to walking would increase estimates of its distance.

2.3.1 Materials and Methods

Participants

Eighteen right-handed participants (mean age = 25.53, SD = 6.68, 13 female) signed up for this study and received gift vouchers (1500 HUF) for their participation. All participants had normal or corrected-to-normal vision, were naive to the purpose of the study, signed a consent form before testing began and were debriefed at the conclusion of the experiment.

Design

The experiment used a within-subject design. Participants traversed fourteen distances (1 m apart) while effort was manipulated by means of locomotion (walking vs. hopping). Dependent measures were participants' verbal estimates of distance and the time it took for them to traverse a certain distance.

Apparatus and Procedure

The study was run in a large hall with a 14m long and 1.5m wide cardboard track marking the testing area (see Figure 3A). Participants read an information sheet, signed consent forms and the experimenter measured their height. The cover story was that the study is testing the relationship between the ability to balance one's body and objects in one's hand. The name of the study ("Balancing act") served as an initial misdirection to the purpose of the experiment. To reinforce the story, before testing participants were asked to balance on one foot with their eyes closed while touching their nose with their index finger and then to walk along the track while balancing a ping-pong ball on a table tennis racket. Following the balancing tasks the experimenter said that the next part of the experiment is about maintaining body balance while walking and hopping and that the last part will return to balancing objects.

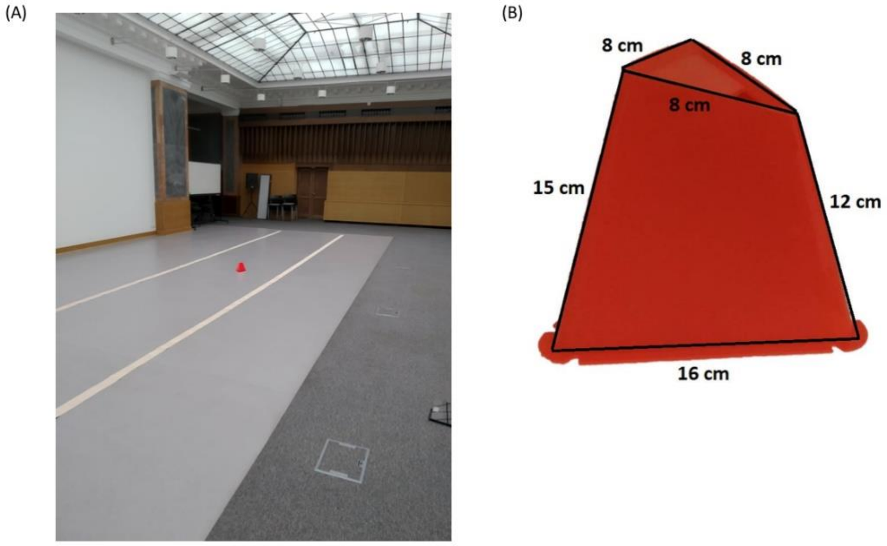

Figure 3. Setup in experiments 2 and 3 with the cone placed at 10 meters from the participant (A). Dimensions of the target cone (B)

A trial began with the participant standing with their back turned to the track. The experimenter placed a cone (see Figure 3B) on the track (between 1 and 14 m away from the participant) and asked the participant to turn around, give a verbal estimate of the cone's distance and then move to it and touch it with their dominant hand. The time from when the participant started moving to when they touched the cone was measured. After touching the cone, the participant walked back to the beginning of the track and remained with their back turned while the experimenter set the cone at the next distance. Each participant went through a "walking" and a "hopping" block with 14 trials per block. In "hopping" trials, participants' legs were positioned such that the ankle of one leg was touching the knee of the other, while in "walking" trials they were instructed to move to the cone at a comfortable pace. Before testing, there were 4 practice trials (2 of each movement type). The order of distance presentation was random and the blocks were counterbalanced across participants. The experiment took between 25 and 35 min to complete. At the conclusion of testing, participants were informed that the study was not about testing balance and prompted to say what they thought the study was about. Finally, the aim of the study and the purpose of the manipulations was explained and they were asked if it had occurred to them during testing that hopping was intended to increase their distance estimates. Only participants who did not answer affirmatively to this question and had not guessed the purpose of the study during the open-ended question were included in the analysis.

2.3.2 Results

One participant correctly guessed the purpose of the study and was excluded from the analysis, resulting in N = 17.

Verbal Responses

Correlational analysis suggested good overall performance on the task (see Supplementary Materials). A paired samples t-test revealed no significant difference between estimates in the hopping and walking conditions ($p = 0.185$). To test for order effects and the possibility that estimates were influenced by effort only at certain distances, a 2 × 14 × 2 mixed ANOVA was performed with the within-subject factors of movement (hopping, walking) and distance (1–14 m) and the between subject factor of counterbalancing order (walking first/hopping first). Participants' height was mean-centered and included in the analysis as a covariate. Greenhouse-Geisser correction was applied because Mauchly's test indicated a violation of the sphericity assumption [$\chi 2(90) = 257.02$, $p < 0.001$] for the main effect of distance [$F(12.03,28.363) = 98.88$, $p < 0.001$, $\eta 2p = 0.876$]. The results showed no other significant main effects or interactions (all ps > 0.05). For the distribution of estimates, see Figure 4A.

Movement Time Analysis

A 2 × 14 × 2 mixed ANOVA was performed with the within- subject factors of movement (hopping, walking) and distance (1–14 m) and the between subject factor of counterbalancing order (walking first/hopping first). Participants' height was mean- centered and included in the analysis as a covariate. Mauchly's test was significant for distance [$\chi2(90) = 232.96$, $p < 0.001$]. The effect persisted after applying Greenhouse-Geisser correction [$F(2.25,33.88) = 131.44$, $p < 0.001$, $\eta2p = 0.898$]. The main effect of movement was also significant [$F(1, 16) = 17.67$, $p < 0.001$, $\eta2p = 0.522$] with participants moving faster in the hopping (Mhopping = 5.13 s, SD = 2.52) than in the walking condition (Mwalking = 5.79 s, SD = 2.54). No other main effect or interaction reached significance.

Additional analyses

To establish if the data supports that the hopping and walking did not influence distance estimates differently, a Bayesian paired samples t-test was performed. As in the analysis in Experiment 1, the prior was Cauchy (0, 0.707). The analysis revealed weak evidence that the data support the null hypothesis BF01 = 1.36).

2.3.3 Discussion

The results of Experiment 2 did not confirm our predictions. There was no significant main effect of the effort manipulation nor did it interact with distance. Participants' reports during the experiment suggested that the effort manipulation was successful. At debriefing they overwhelmingly confirmed that it was more difficult to hop than to walk. Furthermore, during the procedure four participants asked if they could change the leg they were hopping on and an additional five asked how many trials were left because they were becoming tired.

Given that the results were non-significant and that the Bayesian analysis suggests that support for the null hypothesis was only anecdotal it is somewhat difficult to draw firm conclusions. It could have been the case that, apart from visual cues and effort, time needed to traverse the distances also served as a cue for estimation. If this was the case, then it is difficult to tell if the effort manipulation was successful since participants were faster in the hopping than in the walking trials. It is possible that effort increased distance estimates but that this effect was negated because traversing distances faster in the hopping condition made them appear shorter. A follow-up study controlling for time it takes to cross each distance could help disambiguate between these possibilities. Finally, as was the case in the previous experiment, it is possible that the cover story kept participants from perceiving task demands that could otherwise have affected distance estimates.

2.4 Experiment 3: carrying weight across distances

Given that neither Experiment 1 nor Experiment 2 found effects of effort on perceptual estimates, the aim of the third experiment was to replicate previous findings using a more well-established paradigm (Proffitt et al., 2003). Furthermore, in Experiment 2, effort was manipulated by using different types of movement which left open the possibility that the faster speed in the more effortful condition mitigated a potential bias arising from increased effort. Experiment 3 followed the procedure by Proffitt et al. (2003) by manipulating effort by having participants carry a heavy or light backpack and used the same dependent measure (distance estimates). Unlike in the study by Proffitt et al., participants were instructed to walk to the target cones after giving estimates.

2.4.1 Materials and Methods

Participants

Twenty participants (18 right-handed, mean age = 26.15, SD = 4.29, 11 female) signed up for this study and received gift vouchers (1500 HUF) for their participation. All participants had normal or corrected-to-normal vision, were naive to the purpose of the study, signed a consent form before testing began and were debriefed at the conclusion of the experiment.

Design, Apparatus, and Procedure

The design, apparatus and procedure were almost identical to Experiment 2. The design differed with respect to how the effort manipulation was implemented: instead of walking or hopping, participants walked while fitted with a heavy or empty backpack. As in previous experiments, participants read an information sheet, signed consent forms and the experimenter measured their height and weight before testing. The weight of the backpack was individually adjusted so that it was 20% (± 1.5 kg) of participants' body weight. The cover story was that the study aims to investigate preferred walking speeds while carrying different items.

2.4.2 Results

Two participants correctly guessed the purpose of the study and were excluded from the analysis. An additional participant was excluded due to making attempts to measure the length of the track while walking.

Verbal Responses

Correlational analysis suggested good overall performance on the task (see Supplementary Materials). A paired samples t-test revealed no significant differences in estimates in the two backpack conditions ($p = 0.885$). To test for the possibility that there was an effect of effort only at certain distances, an additional analysis was performed. The data was analyzed with a 2 × 14 × 2 mixed ANOVA with the within-subject factors of backpack (empty, full) and distance (1–14) and counterbalancing order (empty backpack first, full backpack first) as a between-subject factor. Participants' height was mean-centered and included in the analysis as a covariate. Mauchly's test of sphericity revealed a sphericity violation for distance [$\chi2(90) = 318.125$, $p < 0.001$]. After applying Greenhouse-Geisser correction, the main effect of distance remained significant [$F(1.39,19.47) = 57.94$, $p < 0.001$, $\eta 2p = 0.805$]. There were no other significant main effects or interactions. For the distribution of estimates, see Figure 4B.

Movement Time Analysis

A 2 × 14 × 2 mixed ANOVA was performed on time participants took to reach the target with the within-subject factors of backpack (empty, full) and distance (1–14 m) and the between-subject factor of counterbalancing order (empty backpack first/full backpack first). Participants' height was mean-centered and included in the analysis as a covariate. Mauchly's test indicated a violation of the sphericity assumption [$\chi 2(90) = 148.8$, $p < 0.001$] so degrees of freedom were adjusted using the Greenhouse-Geisser correction ($\varepsilon = 0.312$). The main effect of distance was significant, with longer distances leading to longer movement times [$F(4.05,56.76) = 639.23$, $p < 0.001$, $\eta 2p = 0.979$]. Furthermore, there was a significant main effect of backpack weight [$F(1,16) = 17.4$, $p < 0.001$, $\eta 2p = 0.554$], with participants walking faster when the backpack was empty ($M = 6.12$ s, $SD = 0.58$) as opposed to full ($M = 6.5$ s, $SD = 0.45$).

Additional analyses

A Bayesian equivalent of a paired-samples t-test was conducted using a built-in function in JASP. The analysis used a Cauchy prior (0, 0.707). Results suggested moderate support for the null hypothesis ($BF_{01} = 3.88$).

2.4.3 Discussion

During debriefing participants reported that the backpack was heavy and that they needed to put in more effort to walk in the full backpack condition. Nevertheless, the findings did not show an effect of effort on distance estimates. The lack of a main effect of backpack was unexpected, especially given that the procedure closely resembled paradigms where such a manipulation modulated distance and slant estimates (Bhalla and Proffitt, 1999; Proffitt et al., 2003). Furthermore, a Bayesian paired-samples t-test showed that the data provide moderate support that there was no difference in estimates between the conditions. Interestingly, participants walked slower in the full than the empty backpack condition, but their estimates did not differ. This speaks against our speculation about the results in Experiment 2, where we suggested that traversing distances faster in effortful conditions might have mitigated the effect of increased effort on distance estimation and raises the question of which aspects of prior studies were missing from Experiment 3 to replicate prior results.

2.5 Experiment 4: standing in place with a heavy or empty backpack

In light of the series of null results in the previous experiments, the aim of Experiment 4 was to try to more closely replicate the original findings showing that carrying weight influences distance estimates. As was the case in earlier studies (e.g.,, Bhalla and Proffitt, 1999; Proffitt et al., 2003) participants were standing in place when giving estimates, no cover story was employed, and the effort manipulation was implemented with a heavy or light backpack.

2.5.1 Materials and Methods

Participants

Twenty participants (all right-handed, mean age = 24.18, SD = 2.8, X, 6 female) signed up for this study and received gift vouchers (1500 HUF) for their participation. All participants had normal or corrected-to-normal vision, were naive to the purpose of the study, signed a consent form before testing began and were debriefed at the conclusion of the experiment.

Design, Apparatus and Procedure

The design, apparatus and procedure were almost identical to Experiment 3. However, the procedure differed in two key aspects. After giving estimates, participants were not asked to walk to the target cone. Instead, they turned away from the track and a new trial began with placing the cone at a different distance. Secondly, unlike in the previous experiments, there was no cover story. Participants were not told anything about the purpose of the backpack nor how it related to their estimates. After the experiment, participants were asked an open-ended question about their opinion on what the experiment was about. This was done to see if they were aware that the backpack weight manipulation was intended to influence their distance estimates.

2.5.2 Results

All participants except one correctly guessed the purpose of the weight manipulation. More precisely, when asked "what the experiment was about," they answered that the purpose of the weight was to make them judge the target cone as being farther away. While explanations for the underlying mechanism differed across participants (e.g.,, two reported the backpack was intended to make thinking more difficult which would lead to higher estimates), all but one guessed the purpose of the manipulation. The participant who did not correctly guess the purpose of the manipulation was excluded on different grounds. During debriefing this participant reported that (s)he gave completely random estimates so that they can complete the experiment as quickly as possible. This was reflected in the data as their mean estimated distance was 167 m while the actual mean distance across the trials was 7.5 m. Correlational analysis suggested good overall performance on the task (see Supplementary Materials). A paired-samples t-test showed a significant difference between the full and empty backpack conditions [$t(18) = 2.66$, $p = 0.016$, $d = 0.61$)]. For the distribution of estimates, see Figure 4C.

Additional analyses

A Bayesian equivalent of a paired-samples t-test was conducted using a built-in function in JASP. The analysis used a Cauchy prior (0, 0.707). Results suggested moderate support for the alternative hypothesis (BF10 = 3.53).

2.5.3 Discussion

The results of Experiment 4 suggest that potential effort of walking across a distance with a heavy backpack increased participants' estimates of that distance. Given that the two main differences from the previous experiment are that participants stood in place and that there was no cover story, each of these could have made a difference. In particular, it could be argued that in Experiments 2, 3, participants traversed distances and that this allowed for better estimates. In such a case, moving across the distance would be a better basis for estimation and possibly make computations of effort unnecessary. However, this possibility seems unlikely, given that the additional analyses from the previous two experiments did not show any improvement in estimates as trials progressed (see Supplementary Materials). The more plausible explanation is that the results in Experiment 4 were due to perceived task demands (Durgin et al., 2012; Firestone, 2013; Firestone and Scholl, 2014). Removing the cover story made the purpose of the experiment apparent to the participants, which likely influenced their estimates.

2.6 Summary and conclusion

In the present study, we investigated whether effort influences perceptual estimates when the cover stories employed are both effective and not distracting. Four experiments were conducted to test how widely and robustly effort influences perceptual judgments of the environment.

Experiment 1 utilized a novel task in which participants estimated height while handling objects of different weights. Experiments 2, 3 investigated whether effortful locomotion and carrying weight increases distance estimates. Several key differences between these experiments and previous studies that found evidence for effort influencing estimates should be pointed out. Firstly, Experiment 1 used estimation of height rather than slope or distance as a dependent variable. The lack of significant effects of effort on estimates of height could be due to employing a successful cover story. However, it is worth noting that physical effort (in terms of height and weight of objects) was not controlled across participants. A potential follow-up would be to scale weights and shelf heights for each participant based on their height, reach, and fitness and repeating the experiment with and without a convincing cover story. Secondly, in Experiments 2, 3, participants traversed the judged distances while in previous studies they made estimates while standing in place (Proffitt et al., 1995, Proffitt et al., 2003; Bhalla and Proffitt, 1999; Meagher and Marsh, 2014; Josa et al., 2019). It might be the case that walking across the distances informed estimates better than effort which was consequentially disregarded. However, this seems unlikely for two reasons. On the one hand, additional analyses showed that participants' performance did not improve in later trials compared to earlier ones. On the other hand, looking at the mean estimates of participants in each of the experiments, there seems to be no systematic indication that traversing the distance improved estimates or that the precision of the estimate was related to effort[1].

Given the results of Experiment 4, where effort seemed to affect distance estimates in the absence of a cover story, we need to consider the possibility that the effort manipulations in

[1] The real mean distance was 7.5 m. Participants were the most precise in Experiment 2 (M = 7.74) in which they traversed distances but there was no significant effect of effort. This was followed by Experiment 4 (M = 8.59) where participants were stationary and there was a significant effect of effort.

Experiments 1–3 did not influence estimates due to the effectiveness of the employed cover stories. The rationale behind employing elaborate cover stories was based on critiques of previous findings proposing that estimates could have been influenced by perceived task demands rather than increase in effort (Durgin et al., 2012; Firestone, 2013; Firestone and Scholl, 2014). For example, Durgin et al. (2009) found that slants appeared steeper only to participants who guessed that they were fitted with a heavy backpack in order to manipulate their estimates. Taking into account criticism of Durgin et al. (2009, 2012) stories as being intrusive or biasing participants in the other direction (Proffit, 2013; Witt et al., 2016), the cover stories employed here were more subtle. Considering that our experiments used cover stories successfully, the possibility that findings of some of the previous studies are due to task demands rather than effort cannot be discounted. This possibility is supported by the fact that Experiment 4 did not use a cover story and the effect of effort was significant. Unfortunately, due to the low number of participants excluded based on correctly assessing the aim of the experiments (four across Experiments 1– 3), statistical analysis of their estimates would not be informative. Taking the results of the four experiments together, they provide little evidence that effort influences the way physical properties of the world are estimated. Furthermore, Bayesian analyses suggest that evidence moderately leans toward effort having no effect when it comes to estimates, at least in Experiments 1, 3.

It should be pointed out that the literature suggesting that energy expenditure influences estimation is much broader and not all of the evidence relies on cover stories. Direct energy manipulations such as consummation of a caloric drink (Schnall et al., 2010; Zadra et al., 2016) or action-based measures or perceptual matching (e.g.,, Witt et al., 2004) might be more robust and independent from task demands. The same could be true for studies manipulating energy expenditure and visual flow (Proffitt et al., 2003; Experiments 2, 3; Zadra and Proffitt, 2016, Experiment 3). These (and other) studies provide converging evidence for effort influencing perception.

However, a growing number of experiments has put even this evidence into question. For example, Woods et al. (2009) failed to replicate effort-based effects using both verbal and action-based measures. Similarly, Shaffer et al. (2013) showed no effect of glucose on participants' slant estimates. In that experiment, participants who consumed a placebo but believed it was a caloric drink reported the hill as less steep than those consuming a caloric drink while blind to the experimental manipulation. Taken together, studies by Durgin et al. (2009, 2012) and Shaffer et al. (2013) suggest that many of the effects interpreted as effort affecting perception may be a product of task demands. The present findings do not directly address the debate concerning energetic effects and certainly do not warrant the conclusion that all reported effects of effort were due to task demands. However, they provide evidence that estimates can be affected by how a task is framed or construed and should prompt further investigation to dissociate effects of effort from effects of task construal (which are interesting in their own right as a reflection of social influence).

A broader question our results touch upon is whether perception is susceptible to top down influences. There is a growing literature suggesting that perception is permeable, with inputs from memory, emotions and action seeping in. For example, language knowledge seems to influence color discrimination (Winawer et al., 2007), knowledge of object colors influences perception of grayscale objects (Witzel et al., 2011), desirable objects are judged as closer (Balcetis and Dunning, 2010) and the room participants are sitting in is judged as darker after reading about an immoral compared to a moral act (Banerjee et al., 2012). Following this line of reasoning, the results of the current experiments could suggest that how a situation is framed directly influences perception. On the other hand, serious arguments have been raised against top-down influences on perception (Firestone and Scholl, 2016). A problem specific to manipulation of effort is whether the measured effects reflect changes in judgments or perception. To put the question simply: do we really “see” the hill as steeper if we are tired or do we simply report it as such? We know from other areas of judgment and decision-making that framing can have a strong influence on decisions. In moral reasoning, seemingly unimportant phrasing differences can sway participants’ decisions (Petrinovich and O’Neill, 1996; Cao et al., 2017) as is the case in decision-making (Wang and Johnston, 1995; LeBoeuf and Shafir, 2003). An interesting approach for future study could be to try to manipulate only the type and framing of the cover story to see whether the effects on estimation would be as pronounced as the framing effects reported in the decision-making literature. Furthermore, given the social nature of experiments (Roepstorff and Frith, 2004), another interesting direction would be to investigate whether the way information about the task is structured and communicated influences participants’ estimates.

In conclusion, our findings raise challenges for the interpretation of effects of effort on perceptual estimates and suggest that how participants interpret the task might play a strong role in modulating their estimates of the physical environment.

2.7 Supplementary Materials

2.7.1 Means and standard deviations per condition (exp. 1-4)

Experiment 1: Verbal estimates

	low shelf	medium shelf	high shelf
light weight	99 (17)	126 (19)	153 (10)
medium weight	96 (21)	124 (17)	150 (14)
heavy weight	97 (21)	123 (20)	151 (12)

Table 1. Mean verbal height estimates (standard deviations) per condition

Experiment 1: Non-verbal estimates

	low shelf	medium shelf	high shelf
light weight	82 (21)	96 (22)	112 (23)
medium weight	81 (20)	97 (24)	112 (26)
heavy weight	83 (19)	97 (23)	109 (27)

Table 2. Mean non-verbal height estimates (standard deviations) per condition

Experiment 2: Verbal estimates

	MOVEMENT	
DISTANCE	**Walking**	**Hopping**
1	0.95 (0.43)	0.95 (0.37)
2	1.92 (0.63)	1.7 (0.57)
3	3.14 (1.36)	2.64 (1.19)
4	3.72 (1.17)	3.8 (1.4)
5	5.2 (1.76)	5.07 (2.12)
6	6.43 (1.69)	6.1 (2.90)
7	7.03 (2.18)	7.08 (2.27)
8	8.97 (4.12)	8.24 (3.3)
9	9.83 (3.47)	9.82 (3.98)
10	11.09 (4.75)	10.39 (3.55)
11	11.74 (4.68)	10.95 (4.02)
12	12.09 (4.41)	12.44 (4.34)
13	13.78 (4.62)	13.01 (4.33)
14	14.39 (5.01)	14.16 (4.72)

Table 4. Mean (SDs) verbal estimates of distance per condition. Distances are expressed in meters.

EXPERIMENT 2: Time

	MOVEMENT	
DISTANCE	**Walking**	**Hopping**
1	2.19 (2.09)	1.95 (1.01)
2	2.77 (1.63)	2.18 (0.46)
3	3.05 (0.6)	2.81 (0.8)
4	3.64 (0.94)	3.21 (0.59)
5	4.59 (1.14)	3.85 (0.82)
6	4.62 (0.95)	4.66 (1.42)
7	5.56 (0.78)	4.7 (0.81)
8	6.41 (1.22)	5.19 (1.06)
9	6.28 (0.85)	5.89 (1.15)

10	7.44 (1.78)	5.98 (1.23)
11	7.92 (1.09)	7.27 (1.75)
12	7.84 (1.91)	7.41 (1.24)
13	9.19 (1.09)	7.87 (1.53)
14	9.52 (1.51)	8.87 (2.12)

Table 5. Mean time (SDs) in seconds participants took to cross each distance.

EXPERIMENT 3: Verbal estimates

	BACKPACK	
DISTANCE	**Empty**	**Full**
1	1.06 (0.46)	0.86 (0.39)
2	1.94 (0.81)	2.58 (2.79)
3	3.16 (1.3)	3.17 (1.38)
4	4.31 (1.69)	4.32 (1.75)
5	5.29 (1.93)	5.68 (2.42)
6	6.63 (2.86)	7.09 (3.27)
7	8.29 (3.41)	8.31 (3.64)
8	9.4 (4.17)	9.23 (3.45)
9	10.91 (5.04)	10.83 (4.71)
10	11.07 (4.61)	11.85 (4.62)
11	12.76 (5.49)	13.15 (5.36)
12	14.09 (6.38)	14.09 (6.13)
13	14.83 (6.15)	13.89 (6.63)
14	16.28 (7.07)	15.77 (6.9)

Table 6. Mean (SDs) verbal estimates of distance per condition. Distances are expressed in meters.

EXPERIMENT 3: Time

	BACKPACK	
DISTANCE	**Empty**	**Full**
1	1.59 (0.31)	1.71 (0.41)

2	2.08 (0.41)	2.8 (1.72)
3	2.85 (0.42)	3.19 (0.37)
4	3.57 (0.58)	3.82 (0.68)
5	4.27 (0.46)	4.56 (0.49)
6	4.93 (0.61)	5.2 (0.58)
7	5.86 (0.71)	6.32 (0.84)
8	6.48 (0.7)	6.98 (0.84)
9	7.16 (0.71)	7.81 (0.77)
10	8.14 (0.95)	8.12 (1)
11	8.39 (1.83)	9.12 (0.84)
12	9.03 (1.19)	9.51 (0.87)
13	9.9 (1.29)	10.23 (1.07)
14	10.69 (1.16)	10.95 (1.15)

Table 7. Mean time (SDs) in seconds participants took to cross each distance.

2.7.2 Participants' performance analysis

EXPERIMENT 1

Pearson correlation showed that participants' verbal estimates of height were highly correlated with actual shelf height ($r(486) = .768$, $p < .001$). The correlation between estimates and actual distances was computed for each participant, converted by a Fisher Z-transformation and a single sample t-test with a test value of 0 was run ($t(17) = 14.38$, $p < 0.001$, $d = 3.39$) which confirmed a good overall performance on the task. The same analyses were performed for non-verbal responses resulting in a high Pearson correlation ($r(486) = .453$, $p < .001$), additionally confirmed by a t-test ($t(17) = 6.72$, $p < 0.001$, $d = 1.58$).

EXPERIMENTS 2-4

The same analyses were conducted for participants' distance estimates in Experiments 2-4 showing a high correlation between estimates and actual distances. For Experiment 2, the analyses revealed a high Pearson correlation ($r(476) = .805$, $p < .001$) and a significant single sample t-test ($t(16) = 28.29$, $p < 0.001$, $d = 6.86$). The same was the case for Experiment 3 ($r(476) = .75$, $p < .001$ and $t(19) = 19.01$, $p < 0.001$, $d = 4.26$) and Experiment 4 ($r(532) = .799$, $p < .001$ and $t(18) = 24.75$, $p < 0.001$, $d = 5.68$).

2.7.3 Bayesian analyses (experiments 1-4)

EXPERIMENT 1

In order to test whether the data support the null hypothesis that weight does not influence judgment estimates, the data were broken down by weight and a series of Bayesian paired-samples t-tests were conducted using a built-in function of JASP. Before the analyses the data were checked to confirm that there were no outliers (see Figure 4.). Shapiro-Wilks tests confirmed that the normality assumption was not violated for comparisons between light and medium ($p = 0.32$), heavy and light ($p = 0.44$) and heavy and medium weights ($p = 0.8$) for verbal responses. Furthermore, no violation was detected for non-verbal responses for comparisons between light and medium ($p = 0.45$), heavy and light ($p = 0.53$) and heavy and medium weights ($p = 0.52$).

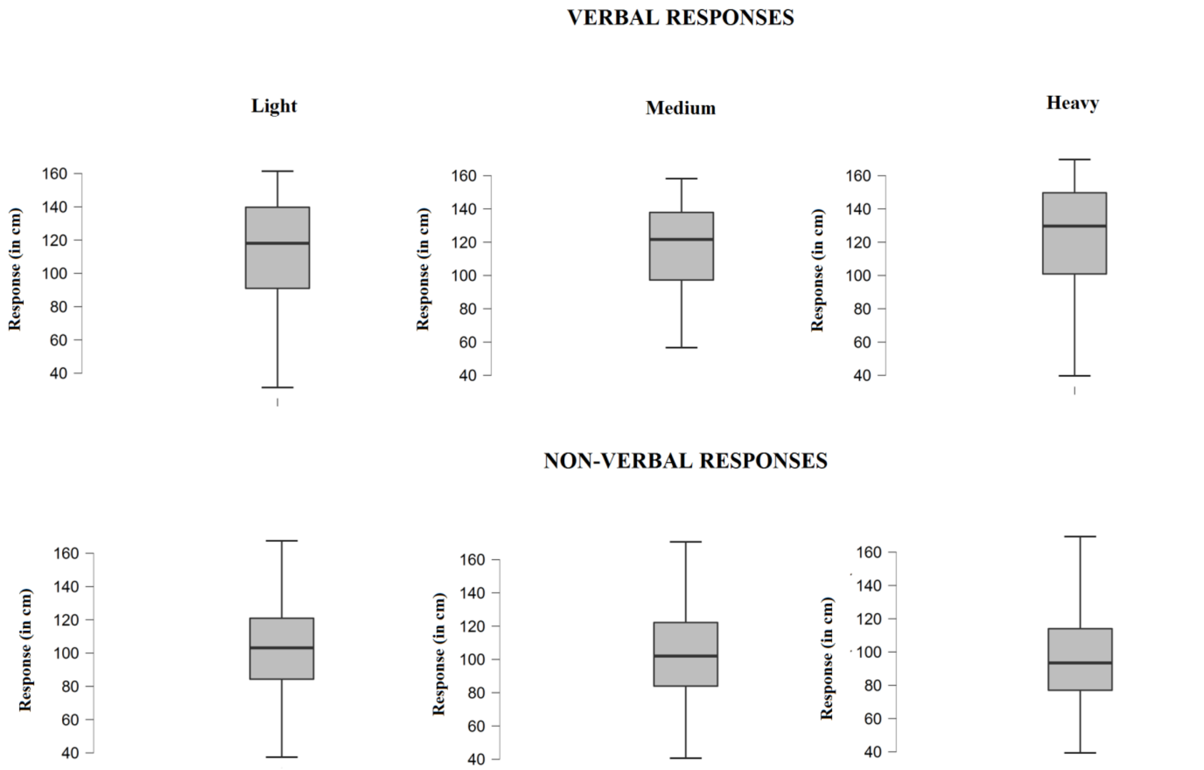

Figure 4. Boxplots of verbal and non-verbal responses for each weight.

In all tests, H_0 stated that the effect size is $\delta = 0$ while H_1 assigned effect size a Cauchy prior centered on 0 with the interquartile range of r = .707. For verbal responses, comparing the influence of heavy and medium weight (see Figure 4) on participants' estimates, showed strong support for the null hypothesis (BF_{01} = 11.32). Comparisons between heavy weight and light weight received moderate support (BF_{01}= 7.82) as did those of light and no mediumt (BF_{01}= 7.26). For non-verbal responses, the null hypothesis was strongly supported in all three comparisons showing $BF_{01} \approx 11$ (see Figure 4 for exact values).

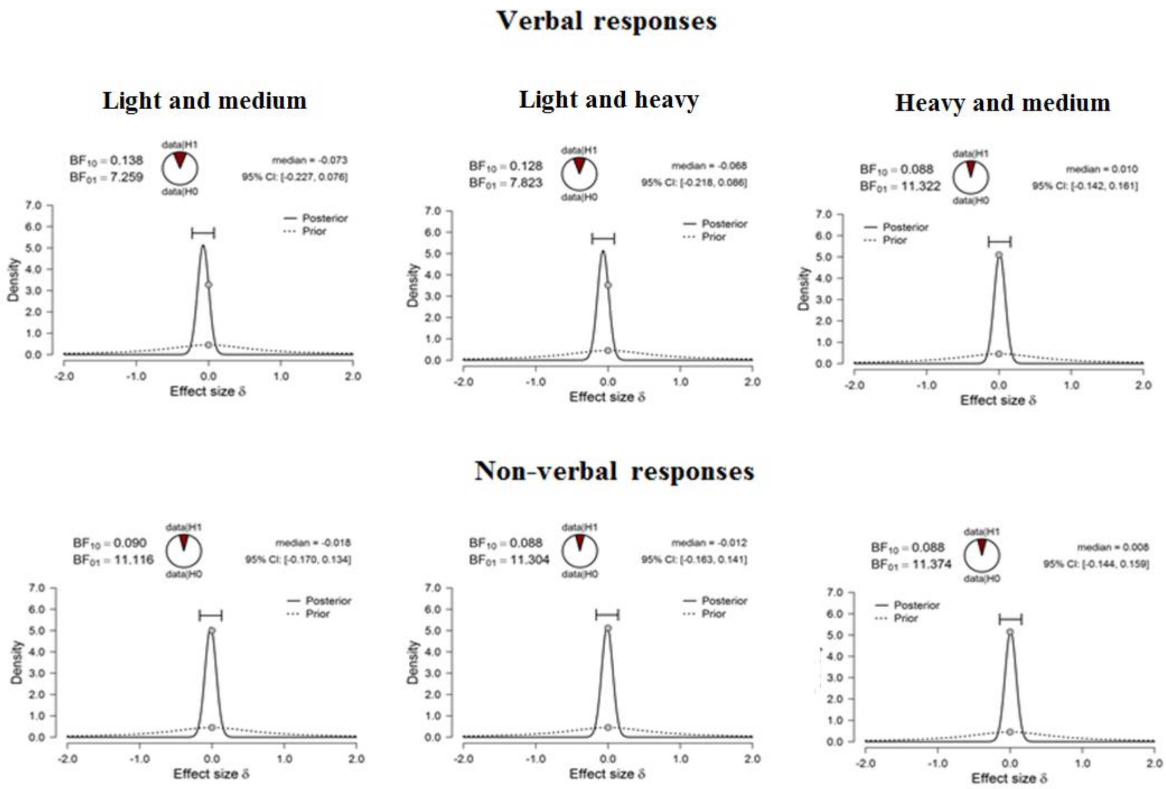

Figure 5. Pairwise comparisons of verbal and non-verbal estimates in different weight conditions. Horizontal bars represent 95% credibility intervals.

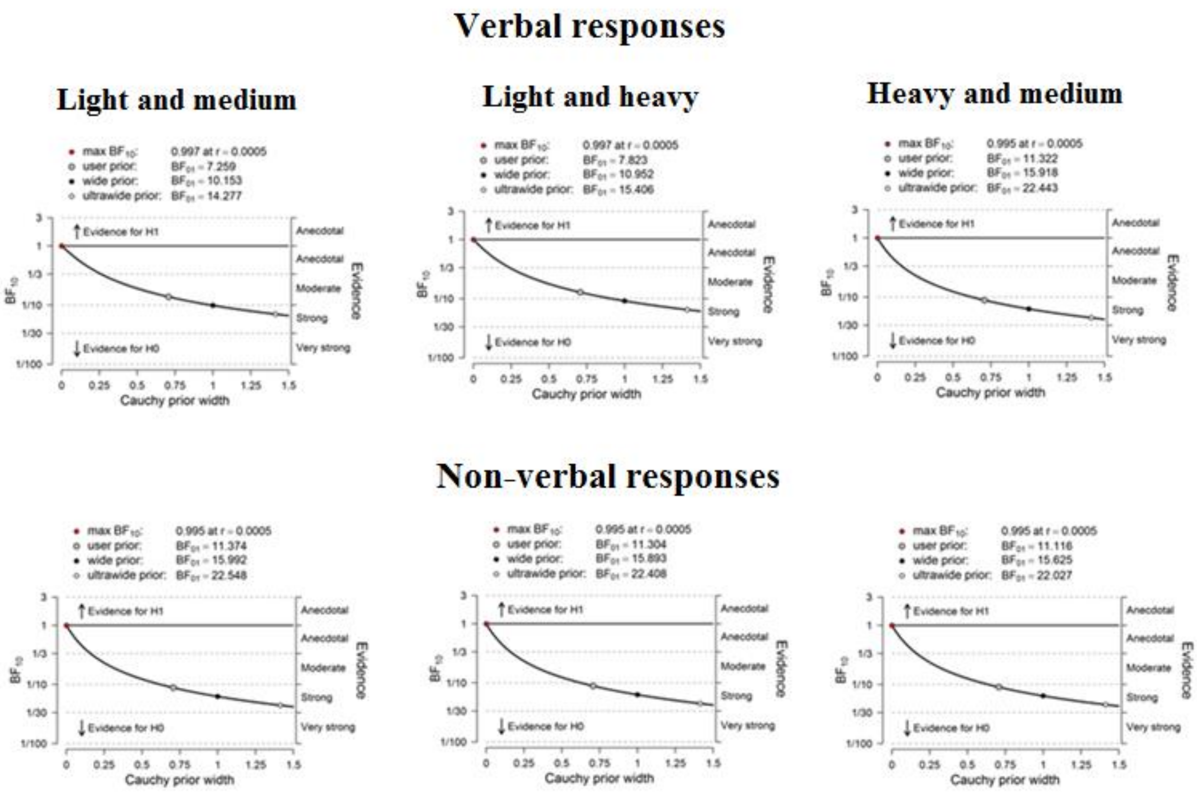

Figure 6. Comparing the prior set in the experiment (r = 0.707) to a wide (r = 1) and ultrawide prior (r = 1.41).

A Bayes factor robustness analysis was conducted using a number of alternative priors (see Figure 5) to see if the chosen prior was a reasonable one. For both verbal and non-verbal responses, it showed that different priors would not change the qualitative conclusion of the data supporting the null hypothesis. Under wider priors than the ones chosen for the analysis, the Bayes factor would give stronger support for the null. Taking into account the robustness and magnitude of the Bayes factor, the analysis shows that the data moderately to strongly suggest that weight did not have an effect on verbal and non-verbal height estimates.

EXPERIMENT 2

To establish if the data supports that the hopping and walking did not influence distance estimates differently, a Bayesian paired samples t-test was performed. Before the analysis, the data was checked for outliers (see Figure 7) and the normality assumption was confirmed with a Shapiro-Wilk test (p = 0.51) As in the analysis in Experiment 1, the prior was Cauchy (0, .707).

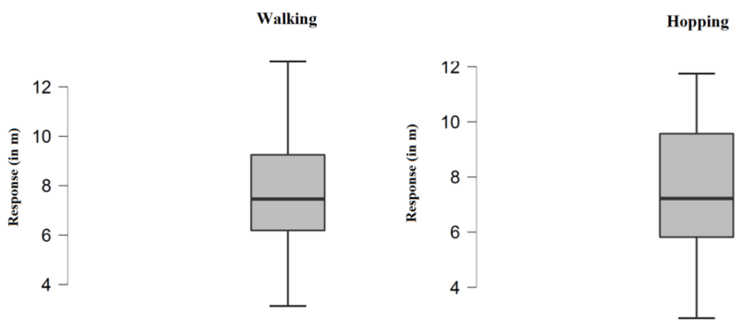

Figure 7. Boxplots of verbal and non-verbal responses for the two effort manipulations.

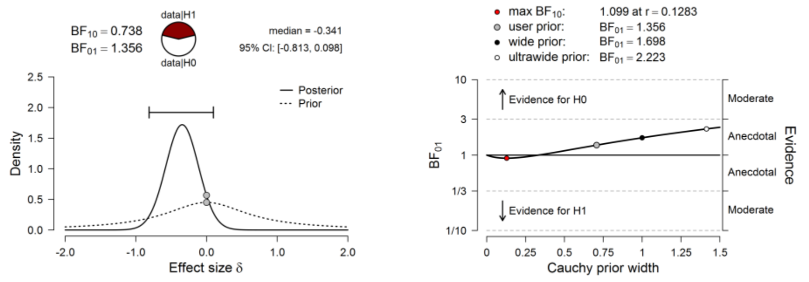

Figure 8. Bayes factor and robustness check for influence of hopping and walking on distance estimates.

The analysis revealed weak evidence that the data support the null hypothesis. A robustness check further suggested that while wider priors would have increased the relative evidence for H_0, it would still remain anecdotal.

EXPERIMENT 3

A Bayesian equivalent of a paired-samples t-test was conducted using a built-in function in JASP. The analysis used a Cauchy prior (0, .707). Shapiro-Wilk test confirmed that the normality assumption was not violated (p = 0.9). Removing the single outlier did not change the results so it was kept in the analysis for the sake of completeness.

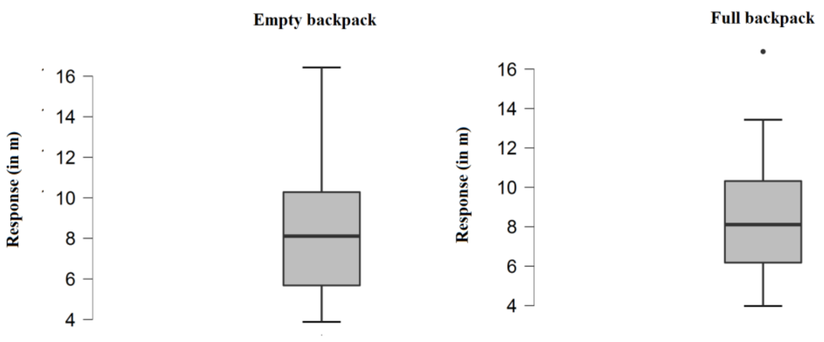

Figure 9. Boxplots of verbal and non-verbal responses for the two effort manipulations.

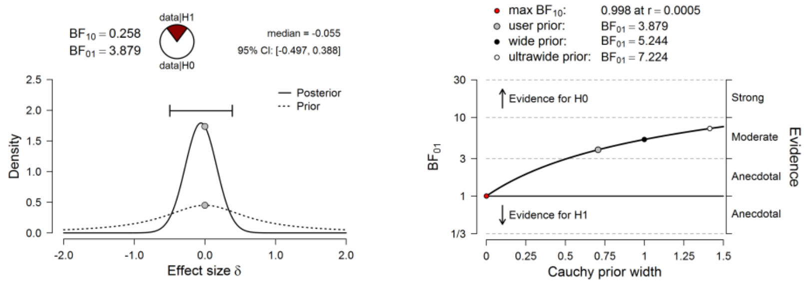

Figure 10. Bayes factor and robustness check for influence of wearing a heavy or empty backpack on distance estimates.

Results suggested moderate support for the null (BF_{01} = 3.88). Robustness analysis showed that wider priors would provide stronger evidence for H_0 but support would remain in the moderate range.

EXPERIMENT 4

A Bayesian equivalent of a paired-samples t-test was conducted using a built-in function in JASP. The analysis used a Cauchy prior (0, .707). Shapiro-Wilk test confirmed that the normality assumption was not violated (p = 0.99).

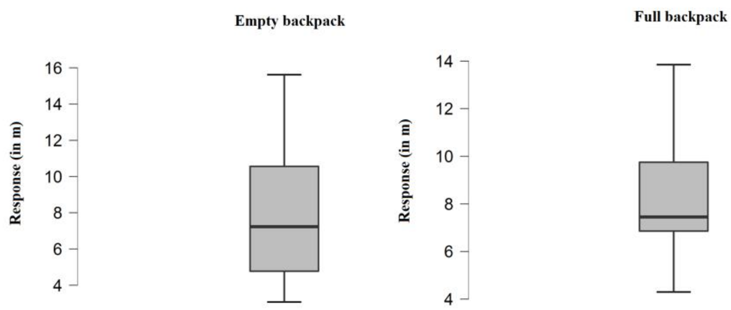

Figure 11. Boxplots of verbal and non-verbal responses for the two effort manipulations.

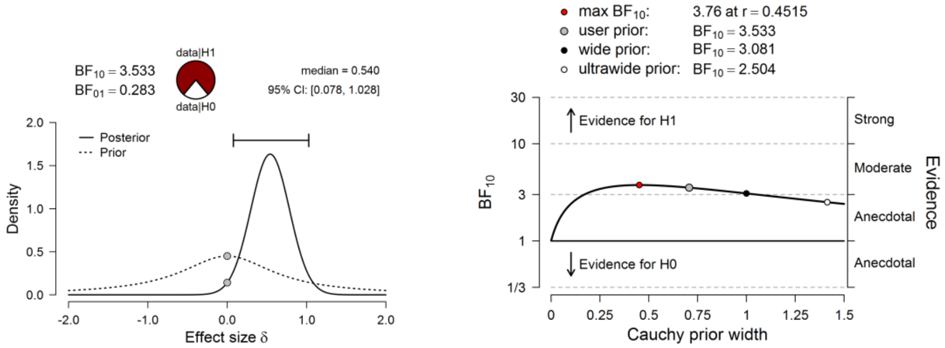

Figure 12. Bayes factor and robustness check for influence of wearing a heavy or empty backpack on distance estimates.

The analysis revealed moderate support for the alternative hypothesis. A robustness check further suggested that while wider priors would have decreased the evidence for H_0 with evidence becoming anecdotal if r = √2.

2.7.4 Analysis of effects of learning and movement

One key difference between Experiment 4 and the previous two experiments should be addressed. In Experiments 2 and 3, participants traversed the distance rather than standing in place. It is possible that moving across the distance served as a strong cue for estimation and therefore participants did not have to resort to effort-based heuristics. To account for this possibility, data from this experiment as well as the two previous ones were re-analyzed to look at estimates over time. Trials were split into two bins: the first 14 and the second 14 trials. If moving across the distance informed estimates, then participants' estimates should improve over the course of the experiment. A paired-sample t-test from the data of Experiment 2 (walking/hopping) did not show significant differences $t(17) = -0.116$, $p = 0.909$). The same analysis for Experiment 3 (walking with empty/heavy backpack) revealed a significant difference $t(17) = -3.03$, $p = 0.008$, $d = -0.73$). Participants judged the distances closer to the actual mean distance of 7.5 in the first half of the experiment ($M = 8.37$, $SD = 3.29$) than in the second half ($M = 8.84$, $SD = 3.47$). This result shows that estimates did not improve over the course of the experiment. Additional paired-samples t-tests showed that participants' performance did not improve during the second part of Experiment 1 ($t(17) = -0.278$, $p = 0.784$) and Experiment 4 ($t(18) = 1.47$, $p = 0.16$).

3 Chapter 3: Mental effort sharing

3.1 Introduction

It has long been suggested in psychology that we tend to maximize benefits while minimizing effort (Hull, 1943). Motor control, for example, can be seen as solving complex problems while minimizing associated effort (Wolpert & Landy, 2012; Shadmehr et al., 2016; Todorov & Jordan, 2002). That observing others acting creates expectations regarding effort minimization has been demonstrated in infants (Southgate & Csibra, 2009; Liu et al., 2019; Liu & Spelke 2017; Jara-Ettinger et al., 2015) as well as other species (Rochat et al., 2008). Furthermore, the motor system uses efficiency as one of the sources to constrain action predictions. Namely, under uncertain locations of a moving stimulus, the predicted future location of the stimulus is influenced by the optimal path for it to reach a known end-point (Hudson et al. 2018; McDonough et al., 2019).

Previous research in joint action suggests that costs of actions are represented by both task partners who tend to minimize the costs for self and other. When handing objects, people rotate them so that they are easier for their partners to grasp (Ray & Welsh, 2011). In a study by Meyer and colleagues (2013), participants were handing objects to their partner who placed them on a shelf. Participants chose to hand the objects in a way that allowed the partner to comfortably manipulate the objects and place them at the correct location. Agents in joint action also take into account the difficulty of the partner's task. Ray and colleagues (2017) devised a task in which one participant had to place a cylindrical object on a line between two marked locations. The object diameter and size of the locations varied such that the object could be placed easily on one of the locations but not on the other. When the object was placed on the line, one of the locations was randomly chosen and the second participant had to place the cylinder on it. Results showed that the participants initially placing the object positioned it closer to the difficult target location, suggesting that they wanted to minimize potential effort of their partner in the event that they had to move it to the more difficult location.

A more direct investigation of cost sharing comes from a study by Török and colleagues (2019). In a series of experiments, participants jointly moved objects on a touch screen. Each agent controlled the object for one part of the path to the destination. The participant that acted first had a choice to take a longer or shorter route. Participants preferred to take longer routes for themselves if that meant that the total length of the route was shorter. In a subsequent paper, Török and colleagues (2020) showed that participants applied the same weights to their own and their partners effort in controlling the object along its path. While the outcome of this strategy was a decrease in individual efficiency, it minimized the total amount of effort across the dyad (co-efficiency).

The sharing of effort in joint action has so far mostly been investigated in the motor domain. However, many cooperative tasks also require mental effort. Both simple tasks like going shopping for a common household to complex ones like operating a sailboat require efficiently distributing memory and attentional resources. This raises the questions whether task partners would consider the costs for self and other when sharing mental effort, similar to the way they have been shown to share physical effort.

Much like with physical effort and motor planning, people tend to optimize individual mental effort expenditure and avoid unnecessary exertion. Kool et al. (2010) presented participants with repeated choices to draw cards from two decks. A card drawn from the deck presented a task of lower of higher difficulty with one of the decks being stacked more heavily with difficult trials. Difficulty was manipulated by having participants switch between different tasks or perform simple or complex arithmetic problems. The results showed that participants preferred to draw from easier decks as opposed to the difficult ones. This gave support to a previously untested proposal that agents prefer to expend the least amount of mental effort assuming the potential payoffs do not differ.

Direct evidence of avoiding mental effort is relatively recent, but it is not very surprising given that there are significant energetic costs associated with mental effort (Fairclough & Houston 2003; Trujillo, 2019) and that the same tendency to avoid mental effort has been demonstrated in other species (e.g. Cocker et al., 2012). In a study by Suzuki and Matuszawa (1997) chimpanzees performed a matching-to-sample task with various levels of rewards offered as reinforcement. In one of the experiments (Experiment 4), when the reward was kept constant, chimpanzees consistently chose a variation of the task where the time between presenting the two comparison stimuli was shortest. Furthermore, when it comes to human performance, mental effort is often reported to bring about a distinct phenomenal experience, feeling laborious and unpleasant (Kurzban, 2016; Inzlicht et al., 2018; Dreisbach & Fischer, 2015).

Evidence of mental effort sharing mostly comes from studies in communication. The typical structure of those experiments is that one participant (the director) has access to a visual scene and attempts to successfully communicate it to the other. Often, there are several frames of reference that the director might adopt, some of which are easier for their partner to grasp. These studies show that the competence of a partner (Schober 2009), estimated cognitive load (Mainwaring et al., 2009) and capability of equal contribution to scene disambiguation (Duran et al., 2011) all affect the level of perspective-taking when describing a scene. Furthermore, this tendency is enhanced if the partner's task is judged too difficult without adopting their perspective (Galati et al. 2013). While there is evidence that agents share mental effort in these tasks, it is less clear whether these strategies result in co-efficiency. The goal in the described experiments is to be understood by one's partner, which may or may not imply an equal distribution of effort. While mentally rotating a visual scene to take a partner's perspective requires mental effort, the extent of effort involved and its amount relative to the partner's effort remains unclear. Furthermore, not all joint tasks have successful communication as their objective.

Given that many joint actions recruit attentional or memory resources (e.g. Kourtis et al., 2014; Clark & Krych, 2004; Shintel & Keysar, 2009), and that those resources are limited and expending them is aversive, a jointly rational strategy for dividing mental effort would be beneficial to achieving joint goals. The first question addressed by the present study is whether people divide mental effort co-efficiently. However, co-efficiency does not necessarily have to mean a completely even distribution of effort, or even a fair one (Strachan & Török, 2020). A rational strategy from a joint perspective might take into account the relative competences of the partners. An underperforming agent might judge that the same task is more difficult for them than their partner. Therefore, our second question was whether competence at a task influences how effort is distributed. Given that participants might not have direct access to their competence, objective performance on the task might not necessarily inform their decisions on effort distribution. As a complementary measure, the experiments tested if self-assessed competence influences effort distribution strategies.

3.2 Experiments overview

Six experiments were conducted, four online experiments using a multiple-object tracking paradigm (MOT) and a further two using a working memory task. MOT tasks draw heavily on limited attentional resources (Tombu & Seiffert, 2008) and elicit mental effort (Alnæs et al., 2014, Wahn et al., 2016). Furthermore, performance in MOT highly correlates with several other paradigms used in attention research (Huang et al., 2012), suggesting it is a good paradigm for generalizing over related, but separate, attentional mechanisms. The purpose of the first experiment was to establish a reliable way to manipulate tracking task difficulty and confirm that participants can tell the difference between different task difficulties. The second experiment aimed to confirm that participants preferred easier tasks over more difficult ones. The third experiment manipulated the difficulty of the task for the participant and their partner (a computer script) to measure participants' preferences of effort distribution between themselves and the partner. Experiment 4 modified the procedure of Experiment 3 so that participants believed that only their preferences (and not their partner's) would influence effort distribution. Experiment 5 followed the same procedure as Experiments 3 and 4 with two key differences: the experiment was conducted live and used a memory task instead of multiple object tracking. The aim of Experiment 5 was to test if findings from Experiments 3 and 4 generalize on a different task. Finally, Experiment 6 used the same procedure as Experiment 5 except that it manipulated the competence of the supposed partner. More precisely, feedback on performance was given in the aggregate ("you and your partner solved X trials correctly") while the partner's performance was modelled on 2SD lower than the worst performers in Experiment 4. The goal was to test if being paired by a poorly performing partner would influence how people distribute mental effort.

3.3 Experiment 1

Before testing how people distribute mental effort, a suitable experimental setup needs to be established. The aim of the first experiment was to pilot several difficulties of a multiple object tracking task and check (i) whether the manipulation of effort is successful in terms of participants' performance on the task and (ii) that participants can accurately judge the differences between the difficulties.

3.3.1 Materials and methods

Participants

Thirty participants (ages 18-34, Mean age = 23.87, SD = 4.14, 12 female) were recruited using the online platform Prolific (www.prolific.co). Participants were filtered for language proficiency (fluent in English) and age (18-35 years). Participants were filtered by age because older age groups show decreased performance on multiple object tracking and memory-based tasks (Fabiani et al., 2006). Compensation for participation slightly varied based on time taken to complete the study (Mean payment = 8.21 GBP per hour). All participants had normal or corrected-to-normal vision, were naive to the purpose of the study, indicated their consent before testing began and were debriefed at the conclusion of the experiment. The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and was conducted in accordance with the Declaration of Helsinki (1991).

Design

The experiment included two tasks: a tracking task and a rating task. In the tracking task, participants had to keep track of several targets among distractors. In the rating task, participants were shown a configuration of targets and distractors and asked to indicate how difficult it would be to track the targets. Both tasks had five possible difficulties, with the total number of circles being the same (12) in each of difficulties. The difficulties varied based on the proportion of targets and distractors: *very easy* (2 targets, 10 distractors), *easy* (3 targets, 9 distractors), *medium* (4 targets, 8 distractors), *difficult* (5 targets, 7 distractors) and *very difficult* (6 targets, 6 distractors).

Procedure

The experiment was programmed in Psychopy3 (Peirce et al., 2019) and conducted online using the Pavlovia (pavlovia.org) platform. Participants were told that in the experiment, they will be helping a child called Alex collect lost marbles. The story was that Alex is walking home carrying a bag of marbles but that the bag has a hole. The marbles keep falling out and mixing with the marbles of other children. Participants' task was to keep track of the marbles and collect only those belonging to Alex, without taking marbles from other children. In addition, participants were informed that between marble tracking trials, they will be presented with an image of targets and distractors and would need to rate how difficult they think it would be to successfully complete that task. In the last part of the instructions, both speed and accuracy were emphasized, and participants were asked not to use external aids to help them track the marbles. A trial (see Figure 1) began with a presentation of a gray square in the middle of the screen with a light-gray fixation cross at the center. The size of the square was defined in normalized units with each side being equal to half of the total screen width. After two seconds, the cross disappeared and a number of blue (target) and black (distractor) circles were presented. The circles size was 3/280 of total screen width. The participants' mouse cursor was frozen, and they were not able to see it. The targets remained blue for three seconds, after which they turned black, and all the circles started moving in random directions. The circles moved at 120 pixel-width of the screen per second. When the position of the circles was detected to be at the boundary of the gray square, the circle would bounce off the edge. The new angle of direction was calculated as $\pi - \alpha$ when bouncing off vertical walls and $2\pi - \alpha$ when bouncing off horizontal ones, with α being the previous angle of direction relative to a horizontal line. After 10 seconds of movement, participants' mouse cursor became visible again and they were asked to identify the initial blue circles by mouse-clicking on them and press the Space Bar key when satisfied with their choice. They were free to de-select a previously selected target by clicking on it again. The maximum number of allowed selections was equal to the number of targets in a particular trial.

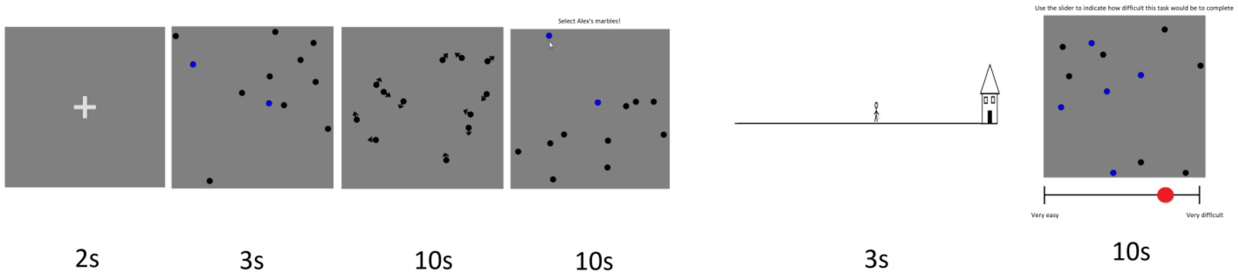

Figure 1. Schematic representation of a trial.

Following the completion of the tracking task, an image of a stick figure walking towards a house was presented for three seconds. After every tracking task, the figure appeared closer to the house, allowing participants to track the progress of the experiment. Finally, a new configuration of targets and distractors were presented in a still image. Participants used a slider (continuous scale 1-5) to rate how difficult they think it would be to accurately complete the presented task. Once the rating was complete, a new tracking task would begin. For tracking tasks, starting positions of targets and distractors were randomly chosen by the Psychopy3 script. For rating tasks, Psychopy generated an image of a random starting position with the targets colored blue and distractors colored black. The color of the targets did not change during rating. In both tracking and rating tasks, there were a total of 12 circles. The difficulty of the task was manipulated by varying the proportion of targets and distractors. The difficulties were coded as *very easy* (2 targets, 10 distractors), *easy* (3,9), *medium* (4,8), *difficul*t (5,7) and *very difficult* (6,6). There was a total of 50 tracking tasks and 50 rating tasks (10 per difficulty). Before the experiment, participants went through a practice session of 5 trials (1 per tracking and rating difficulty). In the practice session, the presentation order of tracking and rating trials was random, and participants were given feedback on their accuracy. In the testing session, the order was random with the exception that each tracking difficulty was followed by each of the rating difficulties the same number of times. At the conclusion of the experiment, participants were debriefed about its purpose. Completion time varied due to instructions and debriefing being self-paced (Mean completion time = 31 minutes).

3.3.2 Results

For both tasks, trials falling under 3 seconds were excluded as well as trials in which no targets were selected, or no rating was given. Following these criteria, 55 trials (3% of total) were excluded. In order to confirm that the difficulty manipulation was successful, an analysis of accuracy was conducted. Accuracy was calculated as a proportion of number of correctly identified targets divided by the number of targets chosen. The data were analyzed in JASP 0.14.1.0. (JASP Team, 2020). Mauchly's test confirmed that the assumption of sphericity was not violated ($p = 0.116$) and the data was analyzed using a one-way repeated-measures ANOVA with the five difficulty levels. The ANOVA confirmed that difficulty had a significant effect on accuracy ($F(4, 29) = 46.99$, $p < 0.001$, $\eta_p^2 = 0.62$) with lower accuracy scores in more difficult trials (see Table 1). Holm post-hoc tests confirmed that differences in accuracy were significant between all difficulty levels (all p's < 0.05).

	Very easy	**Easy**	**Medium**	**Difficult**	**Very difficult**
Mean	0.855	0.792	0.702	0.66	0.609
SD	0.155	0.18	0.194	0.173	0.156

Table 1. Accuracy ratings per tracking difficulty condition.

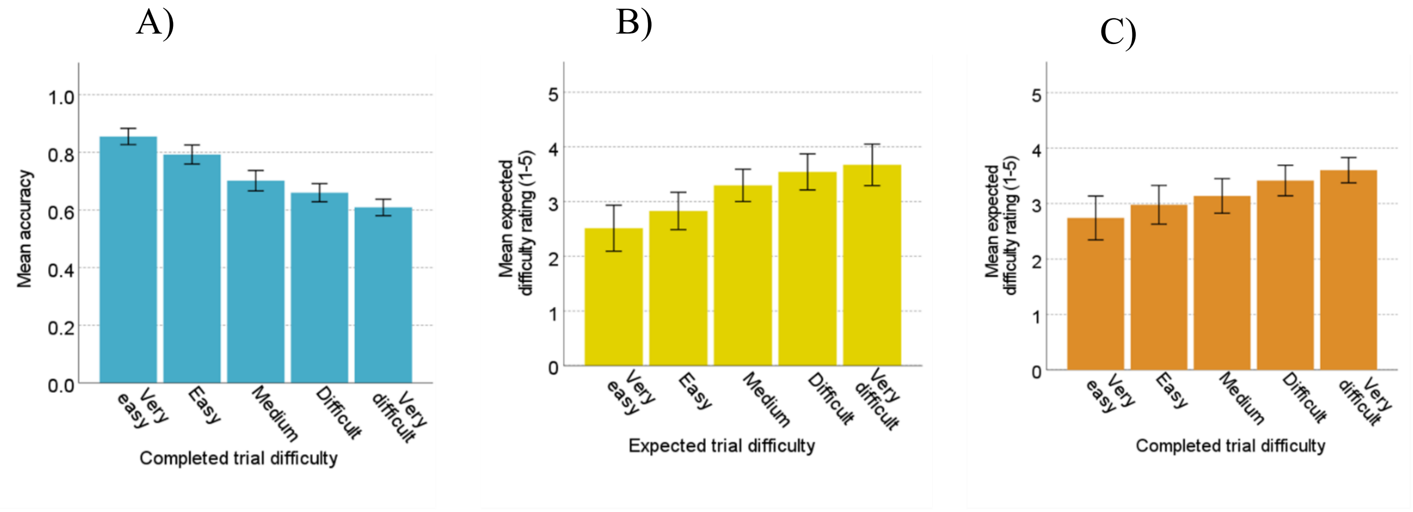

Figure 2. Results of Experiment 1. Mean accuracy per tracking condition (A). Mean response to the question "how difficult would it be to complete this task" (B). Mean response to the difficulty of a current trial is plotted on the Y-axis and the difficulty of the previous tracking trial is plotted on the X-axis (C). Error bars indicate SEM.

For participants' ratings of difficulty, Mauchly's test was significant, indicating a violation of the sphericity assumption $\chi^2(9) = 136.25$, $p < 0.001$. Greenhouse-Geisser correction was applied to the degrees of freedom. A one-way repeated-measures ANOVA showed a significant effect of difficulty $F(1.12, 27.69) = 18.15$, $p < 0.001$, $\eta_p^2 = 0.393$) with scenarios with more targets to keep track of judged as more difficult (see Table 2).

	Very easy	Easy	Medium	Difficult	Very difficult
Mean	2.513	2.829	3.296	3.543	3.671
SD	1.106	0.898	0.777	0.867	0.995

Table 3. Participants' ratings (1-5) of difficulty per condition.

Post-hoc tests (Holm) showed significant differences in ratings between all difficulties except Very easy and Easy (p = 0.16), Medium and Difficult (p = 0.26), Medium and Very difficult (p = 0.09) and Difficult and Very difficult (p = 0.43). Given that the tracking and rating tasks were presented one after another, an additional analysis was conducted to check for potential carryover effects. A one-way repeated-measures ANOVA was conducted on participants' ratings of difficulty as the dependent variable and the difficulty of the preceding tracking trial as the independent variable. The degrees of freedom were Greenhouse-Geisser corrected as Mauchly's test showed a violation of the sphericity assumption ($\chi^2(9) = 88.9$, $p < 0.001$). The results showed that the difficulty in the preceding tracking trial significantly affected the subsequent rating ($F(1.43, 22.68) = 16.758$, $p < 0.001$, $\eta_p^2 = 0.374$).

	Very easy	**Easy**	**Medium**	**Difficult**	**Very difficult**
Mean	2.74	2.987	3.137	3.414	3.6
SD	1.063	0.94	0.841	0.742	0.614

Table 2. Participants' ratings (1-5) of difficulty on the current trial based on the difficulty of the preceding tracking trial.

3.3.3 Discussion

The aim of the first experiment was to establish whether the difficulty manipulation leads to a decrease in performance accuracy and to confirm that participants are aware of the differences in difficulty. Consistent with previous findings in multiple object tracking literature (Intriligator & Cavanagh, 2001; Pylyshyn & Storm, 1988; Alvarez & Franconeri, 2007), the proportion of accurately selected targets was lower in more difficult trials. Specifically, accuracy rates significantly differed from each other on all five levels of difficulty. Participants' ratings reflected the difficulty manipulation to a large degree. However, the differences in the ratings between difficulties were not always significant. This could suggest that the subjective difficulty was not solely driven by performance accuracy. It is possible that the differences between some of the difficulties were too fine-grained for participants to detect. The results also showed a significant carryover effect with ratings of difficulty on a current rating task being influenced by the difficulty of the preceding tracking task. For example, easy tasks were rated as more difficult if the previous task featured difficult tracking. In this case, the difficult tracking task might have biased participants' estimates of their ability which led them to rate the current task as more difficult than they would have otherwise. Alternatively, a difficult tracking task might have signaled that multiple object tracking is a difficult activity, regardless of competence. Research in memory and learning shows that recently acquired information is retained better (Gupta, 2015). If over the course of the experiment participants were updating their beliefs about task difficulty, more recent evidence could have been given more weight. Finally, the tracking and rating tasks were presented immediately one after another, so it is possible that some participants mistakenly rated the difficulty of the tracking task they just completed.

3.4 Experiment 2

The aim of the second experiment was two-fold. Firstly, the design was modified so that it is clearer that rating tasks do not refer to previously completed tracking ones. The second goal was to establish that participants prefer to do easier rather than more difficult tasks. Previous research showed that people sometimes value exerting additional effort on a task but only if they manage to complete the task successfully (Eisenberger 1992; Norton et al., 2012). It is possible that some participants might prefer to do more difficult tasks because they would ascribe value to the additional effort required. For that reason, participants were informed that successful completion means an accuracy of 100% on a given trial.

3.4.1 Materials and methods

Participants

Thirty-two participants (ages 18-34, Mean age = 24, SD = 4.48, 16 female) participated in the experiment. The recruitment method and compensation were the same as in the previous experiment. The selection criteria were the same except that participants who completed the previous experiment were not eligible to take part in Experiment 2. Compensation varied by completion time (Mean payment = 9.76 GBP per hour). The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and was conducted in accordance with the Declaration of Helsinki (1991).

Design and procedure

The design and procedure were the same as in the previous experiment, with a few modifications. The instructions explicitly stated that the success on the task is defined as picking out the correct marbles without picking any of the incorrect ones. The number of tracking and rating difficulties was decreased from five to three. The selected difficulties were very easy (2 targets, 10 distractors), medium (4 targets, 10 distractors) and very difficult (6 targets, 6 distractors). The rationale was that fewer categories would make the differences in difficulty more salient. The order of trials and instructions were modified. The tracking trials were presented first (10 trials per tracking difficulty, random presentation order). This was followed by a rating block. The time allowed on a rating trial was increased from 10 to 15 seconds. Participants were told that the rating block will be followed by another tracking block but that only some tracking trials are predefined.

In the rating block they would rate the extent to which they would like to do the depicted tracking task in the future. They were led to believe that the trial list in the last tracking block would be determined by the ratings they give to presented tracking tasks. As part of the deception, the image of the stick figure did not reach its destination at the conclusion of the tracking block and participants were told that Alex will finish the walk back home during the next tracking block. Furthermore, after each rating trial was completed, a screen appeared with the text “Rating recorded, updating future tasks…”. Each of the difficulties were presented 10 times in a random order. Once participants completed the rating block, they were debriefed, and the experiment concluded.

3.4.2 Results

Due to a data logging error, judgments of difficulty were recorded for only 31 out of the 32 participants. For the multiple-object tracking task, 16 timed-out trials were removed (0.6% of the trials). The same was done for the 10 timed-out rating trials (0.1% of recorded trials). For accuracy ratings, Mauchly’s test confirmed that the sphericity assumption was not violated ($p = 0.49$). A one-way repeated-measures ANOVA showed a significant effect of task difficulty ($F(2,62) = 148.062$, $p < 0.001$, $\eta p2 = 0.827$), with accuracy being highest for Very easy trials ($M = 0.901$), followed by Medium ($M = 0.73$) and Very difficult ($M = 0.613$). Holm post-hoc tests confirmed that there were significant differences in accuracy between all levels of difficulty (all $ps < 0.001$). For participants’ preference ratings, Mauchly’s test revealed a sphericity violation ($\chi 2(2) = 9.55$, $p = 0.008$). A Greenhouse-Geisser corrected one-way rANOVA showed a significant effect of difficulty ($F(1.55, 44.99) = 97.177$, $p < 0.001$, $\eta p2 = 0.77$), with Very easy trials ($M = 4.54$) being preferred the most, followed by Medium trials ($M = 2.837$) and finally Very difficult trials (1.694). Post-hoc tests (Holm) revealed the difference in preference was significant between all levels of difficulty (all $ps < .001$).

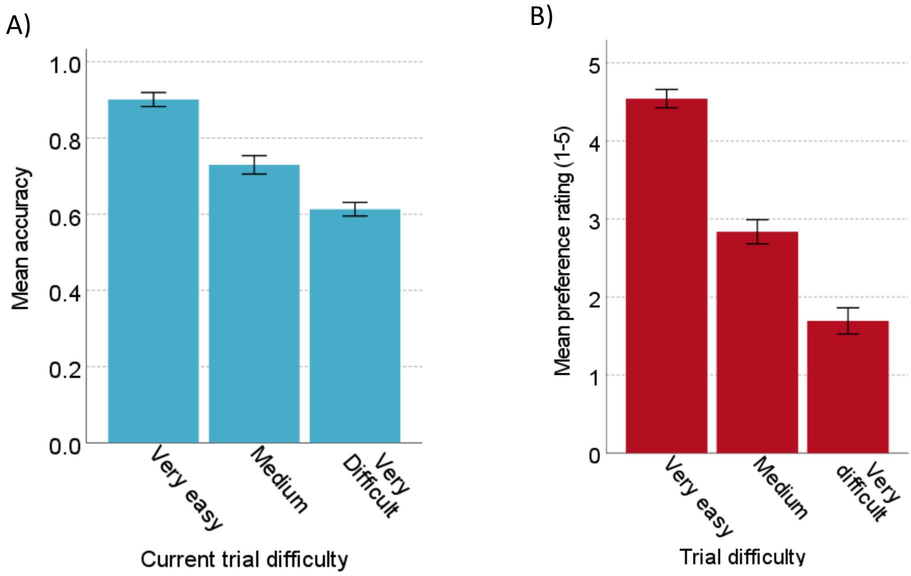

Figure 3. Results of Experiment 2. Mean accuracy per tracking condition (A). Mean response to the question "To what extent would you want to do this task in the future" (B).

	VE	M	VD
Mean	0.901	0.730	0.613
Std. Deviation	0.104	0.137	0.101

Table 3. Mean accuracy ratings and standard deviations per difficulty condition.

	VE	M	VD
Mean	4.542	2.837	1.694
Std. Deviation	0.645	0.842	0.917

Table 4. Mean preference ratings and standard deviations per difficulty condition.

Discussion

The aim of the second experiment was to establish that participants show a preference for doing easier over more difficult trials. A significant main effect of trial difficulty on participants' rating showed this to be the case. As in the previous experiment, higher difficulties also resulted in lower accuracy ratings. Furthermore, the results of the second experiment showed that using fewer difficulties makes it easier for participants to differentiate between them.

3.5 Experiment 3

The aim of the third experiment was to test how people distribute mental effort in a joint context. The first question was would this distribution be egoistic, altruistic or would it tend towards co-efficiency with participants placing equal weights on their own and their partner's effort. If participants tend towards co-efficient distribution, they should be willing to take upon themselves more effort if it means lowering the total effort (theirs + their partners) required to complete a task. A co-efficient strategy does not necessarily have to result in a completely even distribution of effort: individuals more competent at a task could complete it with less effort expenditure. However, objective competence on a task and one's assessment of how competent one is might not be completely in line with each other. It is possible that effort distribution is affected by both of these metrics or just one of them. Therefore, the following two questions were whether the distribution of effort would depend on participants' perceived and actual competence at the task.

3.5.1 Materials and methods

Participants

Sixty participants (ages 18-35, Mean age = 23.6, SD = 4.64, 28 female) took part in the experiment. The recruitment method and compensation were the same as in the previous experiment. The selection criteria were the same except that participants who completed the previous two experiments were not eligible to take part in Experiment 3. Compensation varied by completion time (Mean payment = 7.21 GBP per hour). The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and was conducted in accordance with the Declaration of Helsinki (1991).

Design

The third experiment used the same tracking task as the previous experiments and a modified rating task. The tracking task had three difficulties: *easy* (2 targets, 10 distractors), *medium* (3 targets and 9 distractors) and difficult (4 targets and 8 distractors). The rating task had two within-subject variables. *Self-difficulty* with levels of *easy* (2 targets), *medium* (3 targets) and *difficult* (4 targets). *Partner-difficulty* had the same three possible difficulty levels as self-difficulty. For the rating task, this resulted in a 3 x 3 within-subject design.

CONDITIONS (rating task)				
Participant difficulty	"Partner" difficulty	Participant targets	"Partner" targets	Total targets
Easy	Easy	2	2	4
Easy	Medium	2	3	5
Easy	Difficult	2	4	6
Medium	Easy	3	2	5
Medium	Medium	3	3	6
Medium	Difficult	3	4	7
Difficult	Easy	4	2	6
Difficult	Medium	4	3	7
Difficult	Difficult	4	4	8

Table 5. Difficulty conditions for rating tasks in Experiment 3.

Procedure

Before starting the experiment, participants were informed that they would be doing a tracking task with a partner. In reality, the partner was simulated by a computer-controlled script. The instructions presented the same story about Alex and the marbles as in the previous experiments with one modification. In the modified story, participants were told that their partner could see some of the marbles that Alex lost, and that the participant could also see only some of the marbles. Between them and their partner, they saw all the marbles dropped by Alex. Their objective was to gather all of Alex's marbles, together with their partner. They were also informed that their partner will see the same task as them but that they will be tracking different targets, which to the participant will appear as distractors. To enforce the cover story, participants were shown two images of the same tracking task – one from their perspective and the other from their partner's. After a practice session (2 trials per tracking difficulty), the tracking block started.

Participants went through 10 trials per tracking difficulty for a total of 30 trials. To further convince the participants that they are doing the task with a partner, finishing a tracking task quickly resulted in a screen telling the participants to wait until the partner has completed their part of the task. If participants finished a tracking task under 5 seconds, the waiting time for partner was a randomly selected amount of time between 0 and 5 seconds. After finishing the tracking trials, participants were informed that in the next block they would be shown images of potential future tracking tasks and that based on their ratings they would do some of those tasks in the third part of the experiment. For the rating block, they were shown their own targets (colored blue) as well as their partner's (colored green). The rating block consisted of five trials per Participant X Partner difficulty condition (see table X) for a total of 45 trials. To make it more plausible that ratings are used to generate a trial list for the next block, after every rating a screen saying "Standby, updating future tasks…" appeared. As was the case with tracking trials, finishing rating trials in under five seconds led to a "Waiting for partner" screen appearing for a random amount of time between 0 and 8 seconds. Once the rating block was completed, participants used a slider to rate how good they think they are at MOT tasks compared to other people. Following this, they were asked if they were convinced that their partner was a real person. Finally, participants were debriefed, and the experiment ended.

3.5.2 Results

Due to an error in communication between Pavlovia and Prolific, data was collected for 57 out of the target 60 participants. For accuracy analysis, timed-out trials were removed (0.3% of the data) and Mauchly's test confirmed that the sphericity assumption was not violated ($p = 0.624$). A one-way repeated-measures ANOVA was conducted for the three levels of difficulty as independent variables and accuracy as the dependent variable. $F(2,112) = 83.195$, $p < 0.001$, $\eta p2 = 0.598$). Post-hoc analyses (Holm) confirmed that there were significant accuracy differences between all levels of difficulty (all $ps < 0.001$) with the highest accuracy in the easy condition ($M = 0.921$) followed by medium ($M = 0.826$) and difficult ($M = 0.733$).

	Easy	**Medium**	**Difficult**
Mean	0.921	0.826	0.733
Std. Deviation	0.080	0.122	0.142

Table 5. Mean accuracy ratings and standard deviations per difficulty condition.

Twenty-six participants answered that they believed that they were doing the task with a partner and 36 reported that they forgot that they were doing the experiment with a partner or expressed doubt that the partner was a real person. We first ran a 3x3 repeated-measures ANOVA with factors of self and partner difficulty on the data of the 26 participants that reported being convinced that they were doing the task with a partner. Mauchly's test revealed sphericity violations for self-difficulty ($\chi2(2) = 20.68$, $p < .001$) as well as partner difficulty ($\chi2(2) = 22.88$, $p < .001$) and the interaction between the two ($\chi2(9) = 47.99$, $p < .001$). The results showed a significant main effect of self-difficulty ($F(1.26, 30.204) = 24.361$, $p < .001$, $\eta p2 = 0.504$) and partner difficulty ($F(1.227, 29.444) = 6.506$, $p < .05$, $\eta p2 = 0.213$) Furthermore, the interaction between self-difficulty and partner difficulty was significant ($F(1.85, 44.392) = 3.499$, $p < .05$, $\eta p2 = 0.127$). Holm post-hoc tests showed significant differences in self-difficulty in all conditions (see Appendix for a full post-hoc table). The effect of partner difficulty was significant only on the easy level of self-difficulty.

To check whether participants who believed that their partner was real distributed effort differently than participants who doubted the presence of a task partner, another repeated-measures ANOVA was conducted with the within-subject factors of *self-difficulty* (easy, medium, difficult) and *partner difficulty* (easy, medium, difficult) and the reported belief about the partner being a real person as a between-subject factor. The three-way interaction was not significant ($p = 0.648$), suggesting that effort was distributed in the same way by the two groups of participants. To test whether participants distributed the effort differently based on their assessed competence at the task, a further ANOVA was conducted. Based on participants' median estimate of their ability, each participant was grouped as *low-self assessed competence* or *high-self assessed competence*. A 3x3x2 repeated-measures ANOVA was conducted with the same within-subject factors and *self-assessed competence* as a between factor. There was no main effect of *self-assessed competence* and no interactions between this variable and *partner* and *self-difficulties* (all $ps > .05$), suggesting that participants' preferences for future tasks did not vary based on their perceived competence.

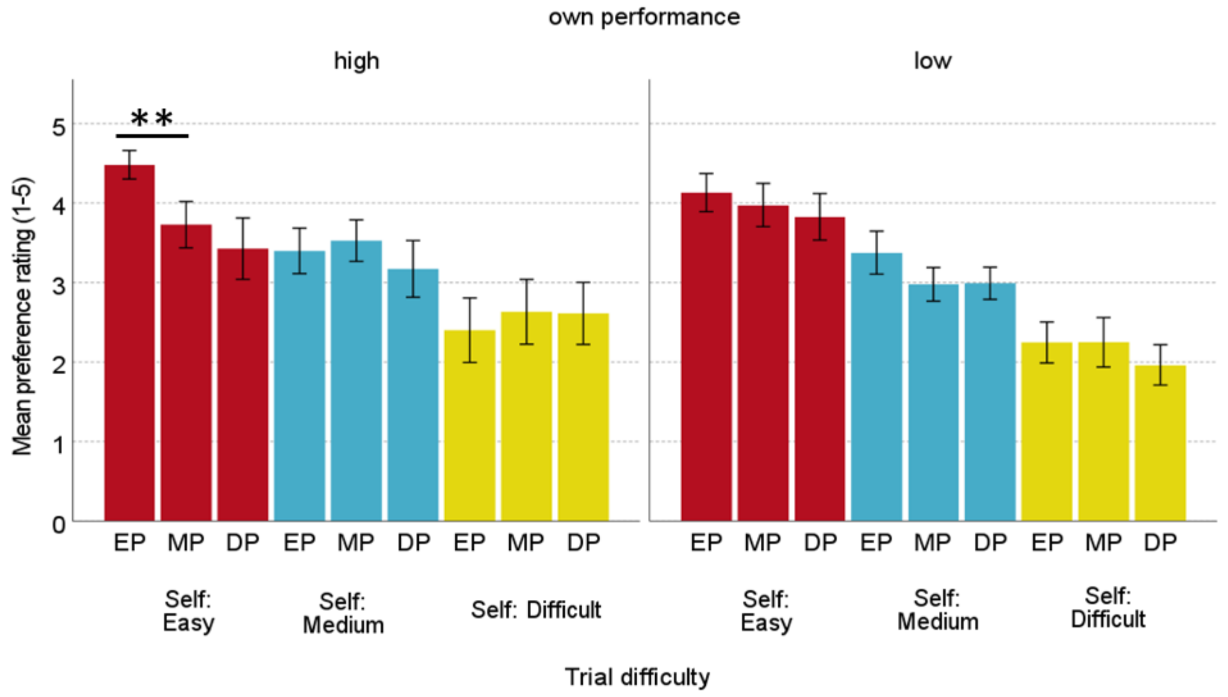

Figure 4. Mean preference ratings for reach participant (self) and partner difficulty in Experiment 3, split by participant performance on the tracking task. “EP”, “MP” and “DP” stand for “Easy partner difficulty”, “Medium partner difficulty” and “Difficult partner difficulty”. Error bars denote SEM.

To test whether participants distributed the effort differently based on their performance on the task, the data was split based on the median accuracy (see Figure 4). A 3x3x2 repeated-measures ANOVA was conducted with the same within-subject factors and *performance on task* (high versus low) as a between-subject factor. Mauchly’s test revealed sphericity violations for *self-difficulty* ($\chi^2(2) = 19.171$, $p < .001$) as well as *partner difficulty* ($\chi^2(2) = 22.52$, $p < .001$) and the interaction between the two ($\chi^2(9) = 39.081$, $p < .001$). A 3 x 3 x 2 repeated-measures ANOVA was conducted with the same within-subject factors and *performance* as a between factor. Greenhouse-Geisser corrections were used where needed. The results showed a significant main effect of *self-difficulty* ($F(1.265, 29.084) = 24.24$, $p < .001$, $\eta_p^2 = 0.513$) with participants preferring easy over medium and medium over difficult trials and a main effect of *partner-difficulty* ($F(1.219, 28.037) = 6.226$, $p = .004$, $\eta_p^2 = 0.213$). The three-way interaction between *self-difficulty*, *partner difficulty* and *performance on task* was also significant ($F(2.02, 46.452) = 4.191$, $p < .05$, $\eta_p^2 = 0.154$). Post-hoc testing revealed differences in preference ratings between all *self-difficulty* conditions. The effect of *partner difficulty* was significant only at the *easy* level of *self-difficulty*, with participants judging the *easy-partner* difficulty as preferable and making no differences between *medium-partner* and *difficult-partner* difficulties. This suggests that the effect of partner difficulty was driven by participants who perform better at the task in terms of accuracy. Interestingly, there was a strong positive correlation between mean accuracy of each participant and self-assessed competence ($r(25) = 0.554$, $p = 0.004$) but only accuracy had an influence on participants' preferences for future tasks. An additional rANOVA was run on the amount of time participants took to rate presented tracking tasks. The ANOVA showed no significant effects of *self-difficulty* or *partner difficulty*, and the interaction was also not significant (all ps > .05).

3.5.3 Discussion

The aim of the third experiment was to investigate if (i) people distribute mental effort co-efficiently, (ii) this distribution changes based on their competence at the task and (iii) this distribution changes based on self-assessed competence on the task. Participants' preferences for future trials seemed to have favored an egoistic effort distribution strategy, with a strong main effect of *self-difficulty*. Not only did participants prioritize their own effort over their partner's, in most conditions they did not prefer less effort for the partner even when the participant's own effort was the same. The difficulty of the partner was considered only when the participants' part of the task was easy. Further analyses showed that this preference for reducing the partner's effort was only shown by participants who performed better at the MOT task. In contrast, participants' estimates of how good they are at the tracking task were not related to preferences when it came to partners' effort.

These results run contrary to previous findings of cost-sharing in joint action (e.g., Ray & Welsh, 2011; Török et al., 2019). One possibility is that participants attempted to avoid potential free-riding from their partner. Punishing a partner by assigning more work or lesser rewards has long been established as an effective strategy to ensure cooperation in economic games (e.g., Fehr & Gächter, 2000). However, participants did not seem to have been after an equal distribution of effort: they were solely minimizing their own and had no systematic strategy when it came to deciding about their partner's. If concerned about free-riding, participants could have consistently assigned more effort to the partner. Another possibility is that participants were guided by certain assumptions about their partner's strategy. Participants believed that their partner was also giving their own ratings. They could have assumed that their partner would be largely egoistic, so they did the same. A further possibility is that when it is not immediately apparent what the partner's effort would be, participants do not pay attention to it. All conditions except *easy self + easy partner* fall at the limit of, or above the subitizing range. This might have not allowed for rapid estimation of total and relative efforts so participants might have adopted a strategy of estimating their own effort first and then their partner's under the condition that it is apparent at a glance. This might be supported by the relatively short deliberation times: participants were given 15 seconds to decide on a trial rating, but the mean deliberation duration was 3.8s.

Furthermore, it is worth pointing out that over half of the participants reported doubts about whether their partner was a real person. Even assuming that everyone who had doubts reported them when prompted, this number is unusually high. This is likely due to how the partner was simulated. When participants finished a trial on time, the Psychopy script checked their completion time against the approximate RTs from the previous two experiments. If the participant finished the trial faster, they would be told to wait for the partner to finish their part of the task. However, the tracking difficulties were easier in Experiment 3 and participants completed them faster so reaction times from Experiments 1 and 2 might not have been a good basis for approximation. This could have created the impression that the partner is taking an unusual amount of time to finish simple tasks. Combined with the fact that the study was conducted online without a confederate to stand in as the partner, it seems plausible that this created some suspicion among the participants.

3.6 Experiment 4

The aims of experiment 4 were to address two potential limitations of the previous experiment. To make sure that participants' strategies were not influenced by assumptions about partner's choices, the procedure was adjusted so that the participant believed that in each dyad, only one participant will be rating future trials. They were then informed that they were randomly chosen to be the one doing the ratings. Secondly, to make the "partner" more believable, their supposed reaction times were made to resemble the RT data collected in Experiment 3. Twenty RTs per difficulty were selected from the previous experiment. When the participant would finish their tracking trial, Psychopy would select one of the RTs and check its value against the participant's RT. If the participant was faster, they would see a "Waiting for partner" screen for the duration of the difference in the RTs. Otherwise, the next trial would begin.

3.6.1 Materials and methods

Participants

Sixty participants (ages 18-35, Mean age = 24.3, SD = 3.8, 31 female) took part in the experiment. The recruitment method and compensation were the same as in the previous experiment. The selection criteria were the same except that participants who completed the previous two experiments were not eligible to take part in Experiment 3. Compensation varied by completion time (Mean payment = 6.95 GBP per hour). The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and was conducted in accordance with the Declaration of Helsinki (1991).

3.6.2 Results

One participant was excluded from the analysis because they completed less than half of the trials. For the remaining data, before the analysis was conducted, timed out trials were removed from the data (0.4% of all trials). 42 out of 59 participants reported that they doubted whether their partner was a real person or that they forgot that they did the task with a partner.

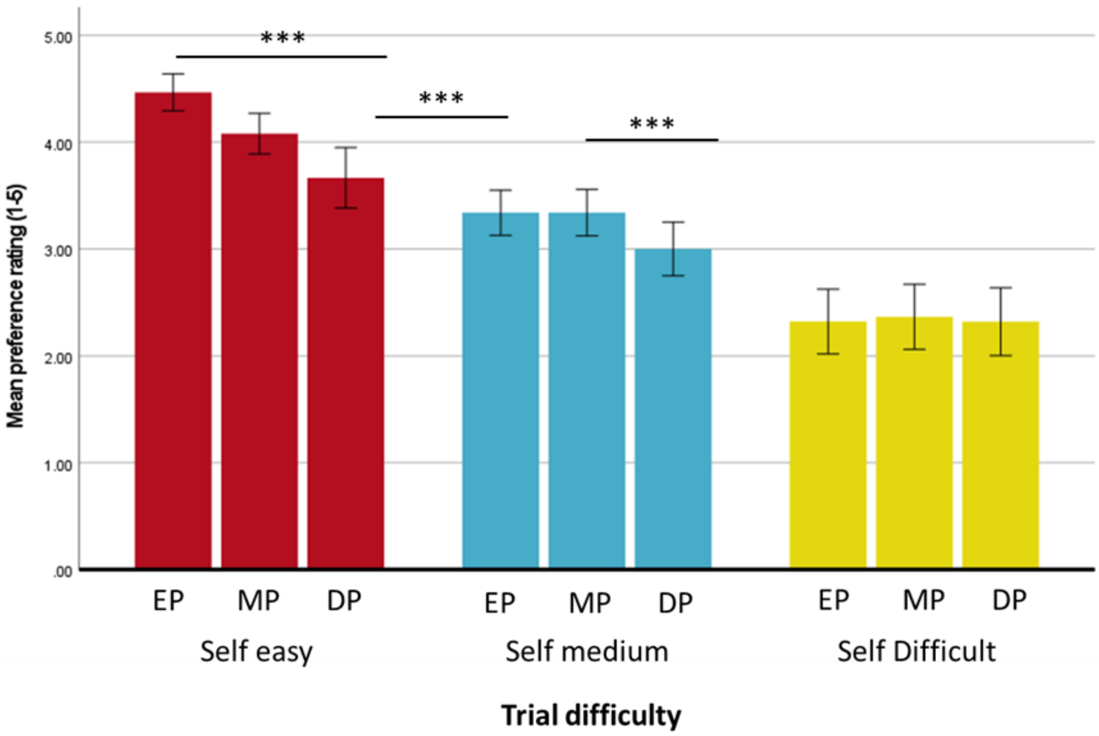

Figure 5: Mean preference for trial type for all participants in Experiment 4. Error bars denote ±2 SE.

A 3x3 repeated-measures ANOVA was conducted on the ratings of all 59 participants. Mauchly's tests showed sphericity violations for self-difficulty ($\chi^2(2) = 52.77$, $p < .001$) as well as the partner's difficulty ($\chi^2(2) = 28.06$, $p < .001$) but not for the interaction between the two (p

= .08). Greenhouse-Geisser corrections were applied where needed. The ANOVA revealed a significant main effect of self-difficulty ($F(1.25, 72.33) = 95.33$, $p < .05$, $\eta_p^2 = 0.62$), partner difficulty ($F(1.44, 83.53) = 10.88$, $p < .05$, $\eta_p^2 = 0.16$) and an interaction between the two ($F(4, 232) = 8.37$, $p < .05$, $\eta_p^2 = 0.126$). Bonferroni post-hocs were used to test for differences between conditions (see Figure 5). The results suggest that when participants had an easy trial, they tried to minimize the partner's difficulty, preferring the partner to have an easy trial over a medium one and a medium over a difficult one. In medium trials, participants seemed to prefer the partner's difficult trial the least while preferring an easy and medium trial for the partner equally. Finally, when the participant had a difficult trial, they did not take the partner's difficulty into account. Given that the majority of participants did not believe their partner was a real person, these results are puzzling: why would they want to minimize the effort of the partner when they did not believe that partner exists? One possibility is that all participants framed the task as if it were presenting them with a scenario in which they would act with a hypothetical partner. If this were the case, their belief of whether they are doing the task with a partner or not might not have had a substantial influence on the ratings. An alternative could be that the effect of minimizing partner's difficulty is driven by those participants who were convinced that the partner was a real person.

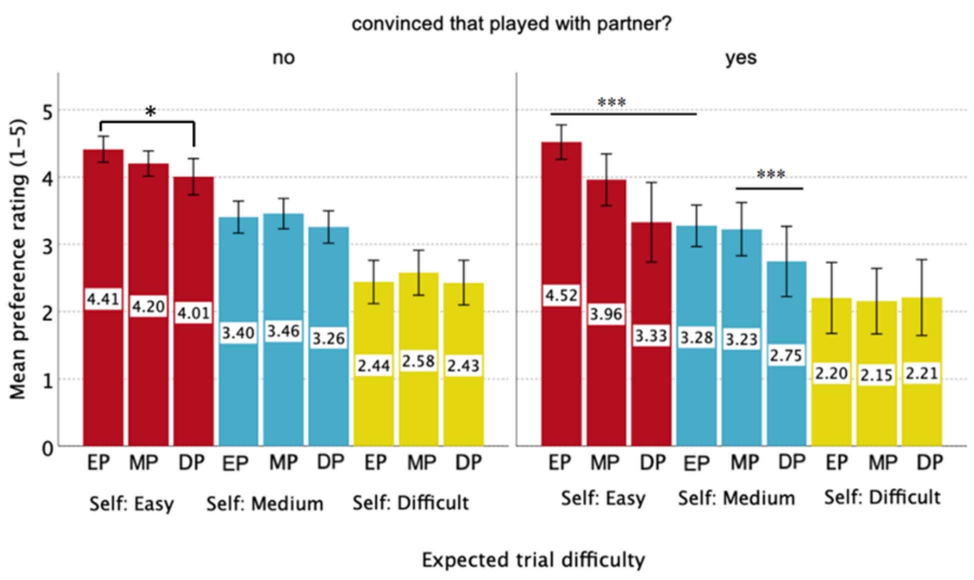

Figure 6: Mean preference for trial type for all participants in Experiment 4. Error bars denote ±2 SE.

To test this possibility, a repeated-measures ANOVA was conducted with within-subject factors of self-difficulty (easy, medium, difficult) and partner difficulty (easy, medium, difficult) and the between-subject factor of *reported belief about partner* (partner is real / partner is fake). Mauchly's test showed a sphericity violation for the main effect of self-difficulty ($\chi^2(2) = 51.85$, $p < .001$) and partner-difficulty ($\chi^2(2) = 25.28$, $p < .001$) but not the interaction ($p = .23$). Greenhouse-Geisser corrections were applied as needed. The main effect of self-difficulty was significant ($F(1.25, 71.08 = 77.32$, $p < .001$, $\eta_p^2 = 0.58$) as well as the main effect of partner-difficulty ($F(1.48, 83.61) = 14.48$, $p < .001$, $\eta_p^2 = 0.2$). The self-difficulty x partner-difficulty interaction was significant ($F(4, 228) = 11.25$, $p < .001$, $\eta_p^2 = 0.16$). Perhaps of most interest, the three-way interaction between self-difficulty, partner-difficulty and the reported belief about whether the partner was real was also significant ($F(3.63, 207.12) = 3.16$, $p < .05$, $\eta_p^2 = 0.53$). People who were not convinced that their partner was real minimized the partner's effort to a small extent: across self-difficulty conditions, the only difference in preferences about partner's effort was found in the self-easy condition and only between partner-easy and partner-difficult trials. On the other hand, those convinced that the partner was a real person consistently minimized their effort when their own trial was easy and to a small extent when their own trial was of medium difficulty.

3.6.3 Discussion

To ensure that participants were convinced that they were doing the task with a partner, the waiting times for the partner to complete their part of the task were closely modelled after real data from Experiment 3. Despite this, less than 30% of participants reported believing that the partner was a real person. Given that the partner's reaction times were constructed from real data, it does not seem probable that this skepticism arose from cues in the procedure of the experiment. A more likely explanation is that people are generally more suspicious when it comes to internetbased interactions in general.

The main aim of this experiment was to ensure that participants are not basing their ratings on the mistaken belief that their partner is also rating potential trials. To eliminate this kind of strategizing, participants were informed that future trials will be selected based only on their own ratings. The results showed that participants who believed that their partner was a real person minimized their effort more than those that did not. Even so, this was the case only when the participant had an easy or medium-difficulty trial. While these results do not show that participants' distribution of effort was co-efficient, they demonstrate that the partner is taken into account. This consideration decreases as the difficulty for the participant increases. The apparent disconnect between these results and previous studies demonstrating co-efficient behavior might be due to the MOT task being significantly more difficult than the simple motor tasks employed in (Török et al. 2019) and (Strachan & Török , 2020). This is particularly apparent in the self-easy condition where participants reliably tried to minimize the partner's effort in all three partner-difficulty conditions. A potential follow-up could be to make the entire task easier, bringing it closer to the differences in difficulty used in motor tasks.

On a related note, it remains an open question how participants interpreted the task. On the one hand, minimizing the partner's effort was more prevalent when they believed the partner is a real person. However, even those who were not convinced that the partner is real somewhat minimized their effort. This opens the possibility that, at least in part, the ratings were given about hypothetical scenarios rather than based on a belief that these trials will be done at a later time with a real partner. A possible follow-up could make the scenario such that it is more salient that it is a cooperative task with a real partner, rather than a possibly hypothetical one. Alternatively, beliefs about the partner could be manipulated more directly, comparing cost distributions between real and imagined partners.

3.7 Experiment 5

The previous experiments showed that people distribute mental effort mostly following an egoistic strategy, but still consider their partner if they are convinced that the partner is a real person and task difficulty is easy to moderate. Given that multiple object tracking is just one task requiring mental effort, the main aim of Experiment 5 was to establish whether the findings in the previous experiments generalize to other cognitive domains. Previous research shows that working memory elicits mental effort (Backs & Seljos, 1994; Choley et al., 2005). Furthermore, working memory is limited (e.g. Cowan, 2010; Oberauer et al., 2016) so in a joint task it would be rational to optimize the distribution of items to be remembered in order to achieve the best possible performance. Three additional modifications were made compared to the previous experiments. Firstly, to ensure that participants were convinced that their partner was real, the experiment was conducted in a live, rather than an online setting. Secondly, the information of own difficulty and partner difficulty was made more salient and, lastly, the overall difficulty of the task was lowered.

3.7.1 Materials and methods

Participants

20 participants (ages 19-28, Mean age = 25.3, SD = 3.4, 13 female) were recruited using the SONA System at the Central European University in Vienna. For their participation in the study, participants were compensated with approximately 10 euros. The study was approved by the Psychological Ethics Board (PREBO) Vienna.

Design

The experiment had the same structure as the ones previously reported: the first block consisted of a memory task of varying difficulty which was followed by a rating block. In the rating block, participants saw three possible difficulties for themselves and their partner and had to rate how much they would like to complete these trials in the future. The number of items to remember were three (easy), five (medium) and seven (difficult). This resulted in a 3 x 3 design for the rating task.

Procedure

To convince participants that they are doing the experiment together with a partner, testing was ostensibly done in pairs. The experimenter introduced the participants to each other, informed them that they will be doing a memory task together and then led them into separate rooms. The experiment was divided into three parts: a practice session, the memory task, and the rating task. First, participants completed a practice session to familiarize themselves with the memory task. A trial would begin with presentation of a series of blue and green squares occluding the numbers the participant would be asked to remember (see Figure 7). The blue cards covered the participant’s numbers while the green one covered their partner’s. Then the squares disappeared, revealing the participant’s numbers but not the partner's. Finally, they were asked to input the numbers by typing on a keyboard before proceeding to the next trial. The practice session consisted of six trials and participants received immediate feedback concerning their performance.

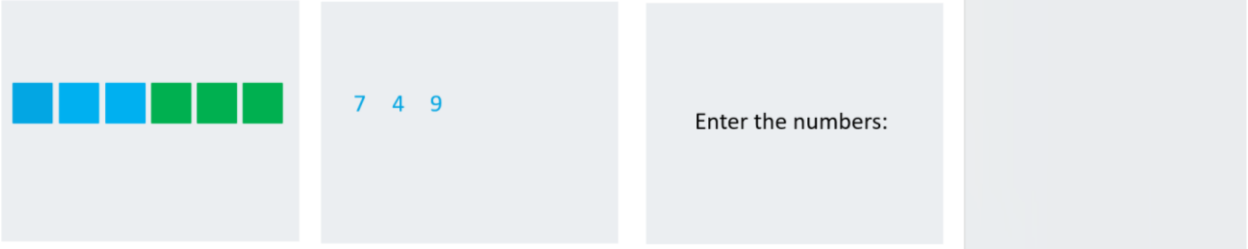

Figure 7. Schematic representation of a trial in Experiments 5 and 6.

After completing the practice session, a message on the screen informed participants that they will now be connected remotely with their partner and each of them will do their part of the memory task. It was emphasized that successfully completing a trial meant that both they and their partner remembered all the numbers correctly. The accuracy feedback was blocked such that participants received only joint accuracy feedback for each of the three blocks of trials. The first feedback was given after 9 trials and the second after 18 and the third after a further 18 trials for a total of 36 memory trials. The accuracy of the partner and the amount of time taken by the partner to complete the trials was based on pilot data.

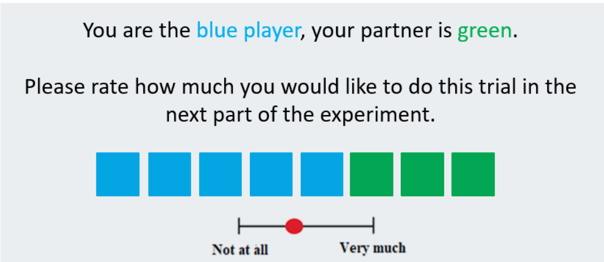

Figure 8. Example of a rating trial. In this trial, the participant would have a *medium* difficulty task while their "partner" would have an *easy* one.

This was followed by the rating block. The participants were told that they were randomly chosen to rate some possible tasks that they and their partner will do in the next part of the experiment. The rating trials presented participants with different difficulty configurations for themselves and their partner. The difficulties were easy/medium/difficult for the participant and the easy/medium/difficult for the partner. All the possible combination were presented four times, resulting in 36 rating trials. Once the participants completed the ratings, they were asked to remain seated until the experimenter came back into the room. This instruction was included as an additional precaution to maintain the illusion that the task was done together until the end of the experiment. Namely, as participants were positioned in adjacent rooms which were not soundproofed, one participant leaving could have made their partner suspicious as to whether they did the task together.

3.7.2 Results

Debriefing the participants revealed that all of them were convinced that they were doing the study with a real partner rather than a script-simulated one. An analysis was conducted to confirm that more difficult trials resulted in worse performance (see Figure 9). Mauchly's test indicated that sphericity was violated ($\chi^2(2) = 14.05$, $p = .001$) so a Greenhouse-Geisser correction was applied. A one-way repeated-measures ANOVA showed that there was a significant main effect of difficulty ($F(1.3, 24.65) = 26.34$, $p < .001$, $\eta_p^2 = 0.58$) with accuracy being highest in easy trials followed by medium and difficult ones (see Figure 9). Post-hoc (Bonferroni) tests confirmed significant differences between all three condition (ps < .001).

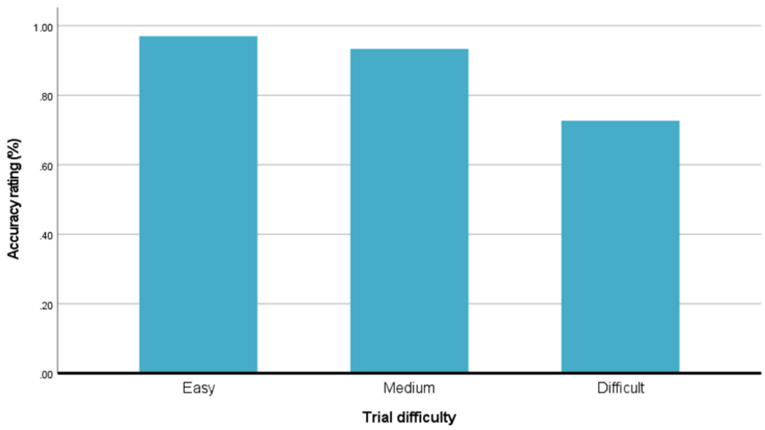

Figure 9. Accuracy percentages for easy (3 items), medium (5 items) and difficult (7 items) memory trials.

For participants' ratings, a 3 x 3 repeated-measures ANOVA was conducted with the three difficulties for the participant and the partner as the levels (see Figure 10). For the main effect of self-difficulty, Mauchly's test indicated a violation of the sphericity assumption ($\chi2(2) = 9.8$, p <

0.05). Greenhouse-Geisser correction was applied. The main effect of self-difficulty was significant ($F(1.36, 25.81) = 60.03$, $p < .001$, $\eta_p^2 = 0.774$) showing highest preference for own easy trials followed by medium ones and, lastly, for difficult ones (see Figure 10).

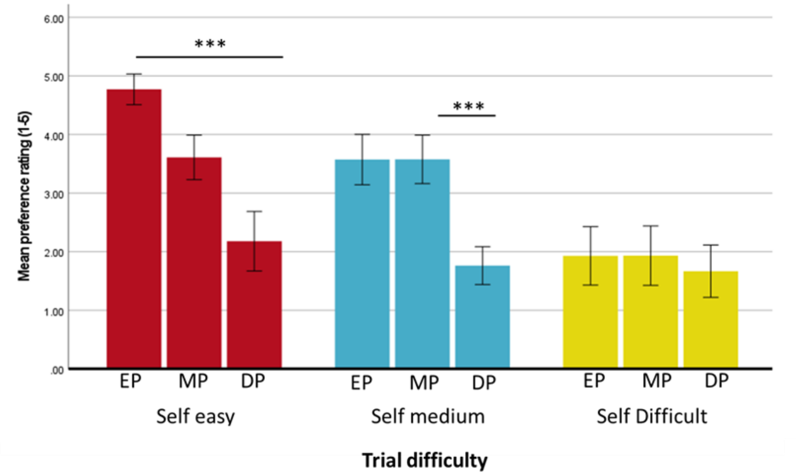

Figure 10. Mean preference ratings on a continuous scale (1-5) for own and partner's difficulty. Error bars denote ±2 SE.

Mauchly's test detected a sphericity violation for the main effect of partner difficulty ($\chi2(2) = 9.2$, $p < 0.05$). The main effect was significant ($F(1.41, 26.76) = 23.33$, $p < .001$ $\eta_p^2 = 0.55$), showing that participants preferred easy trials for their partner over medium or difficult ones. The interaction between *self-difficulty* and *partner-difficulty* was significant ($F(4, 76) = 24.99$, $p < .001$ $\eta_p^2 = 0.57$), demonstrating that participants differentiated between different levels of the partner's difficulty only in self-easy and, to a lesser extent, in the self-medium conditions.

3.7.3 Discussion

The aim of this experiment was to establish whether previous findings of distribution of mental effort generalize to other cognitive domains. Similarly to the previous experiments, the results showed that participants minimized the partner's effort the most when they had an easy trial, followed by a medium-difficulty trial and not at all when they had a difficult trial. However, unlike in the previous experiments, the same level of preference (or lack of thereof) was shown for difficult trials, regardless of whether the difficult part of the trial would have been solved by the participant or their partner. In contrast with previous experiments, this lack of preference did not change even when the participant had an easy trial. Simply put, equal weight was put on difficult trials, regardless of who would be the person assigned to them, suggesting that while the effort distribution strategy was not perfectly co-efficient, participants prioritized minimizing their partner's effort if it would have been on the extreme end of the task difficulty. Curiously, there was no difference between easy-self + medium-partner, medium-self + easy-partner and medium-self + medium-partner conditions. Given that in those three conditions there was no preference for an easy difficulty for the participant, it seems unlikely that this result is due to an egoistic strategy. A more plausible interpretation would be that the easy and medium memory tasks were quite close in their difficulty. While the differences in performance in the easy and medium memory task trials were significant, they were not very large. This leaves open the possibility that while statistically significant in terms of performance, the differences in mental effort might not have been very salient to participants.

3.8 Experiment 6

The aim of Experiment 6 was to test whether effort would be distributed differently if the participant thought they were better at the task than their partner. Unlike in Experiment 3, where participants reported how competent they thought the partner was, in Experiment 6 the apparent competence of the partner was directly manipulated.

3.8.1 Materials and methods

Participants

20 participants were recruited using the SONA System at the Central European University in Vienna. For their participation in the study, participants were compensated with approximately 10 euros. The study was approved by the Psychological Ethics Board (PREBO) Vienna.

Procedure

The procedure was identical to the one in the previous experiment with one key difference. The joint accuracy feedback was presented in a way that made the partner appear to be performing poorly at the task. For each difficulty, the partner's accuracy was lower by three standard deviations than participants' performance in the previous experiment. That resulted in participants being provided with much lower joint scores than in the previous experiment. The second change was adding an additional question to the debriefing. To ensure that perceived partner competence was successfully manipulated, at the end of the experiment participants were asked if they thought their partner had performed better, worse or about the same as them.

3.8.2 Results

Out of 20 participants, 8 reported that they were not convinced that they did the experiment with a real person. Given that excluding their data or splitting the data based on this belief did not lead to any qualitative change in the results, they were included in the analysis. Mauchly's test showed the sphericity assumption was violated ($\chi2(2) = 26.4$, $p < .01$) for the accuracy data (see Figure 11). The main effect of trial difficulty was significant ($F(1.13, 21.48) = 13.35$, $p < .001$ $\eta p2 = 0.41$). Bonferroni post hoc revealed significant differences between difficult conditions and easy and medium ones, respectively. The difference in performance between easy and medium trials was not significant.

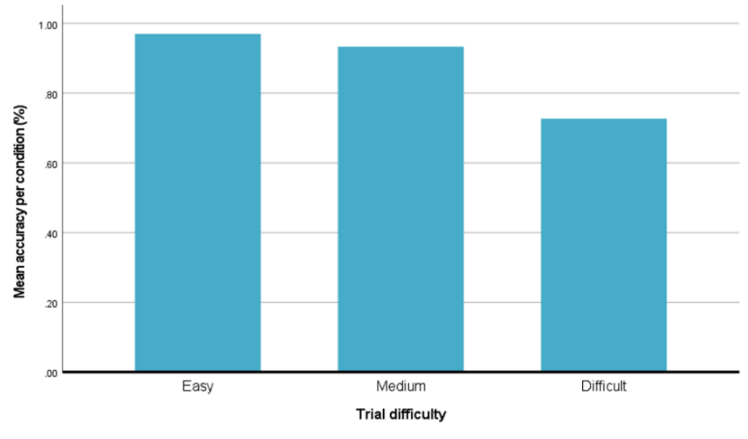

Figure 11. Mean accuracy for each of the difficulty conditions in Experiment 6.

For participants' ratings, Mauchly's test confirmed that the sphericity assumption was violated for the main effect of self-difficulty ($\chi2(2) = 25.1$, $p < .001$), partner difficulty ($\chi2(2) = 12.3$, $p < .05$), and the interaction between the two ($\chi2(2) = 17.42$, $p < .05$). Greenhouse-Geisser corrections were applied as needed.

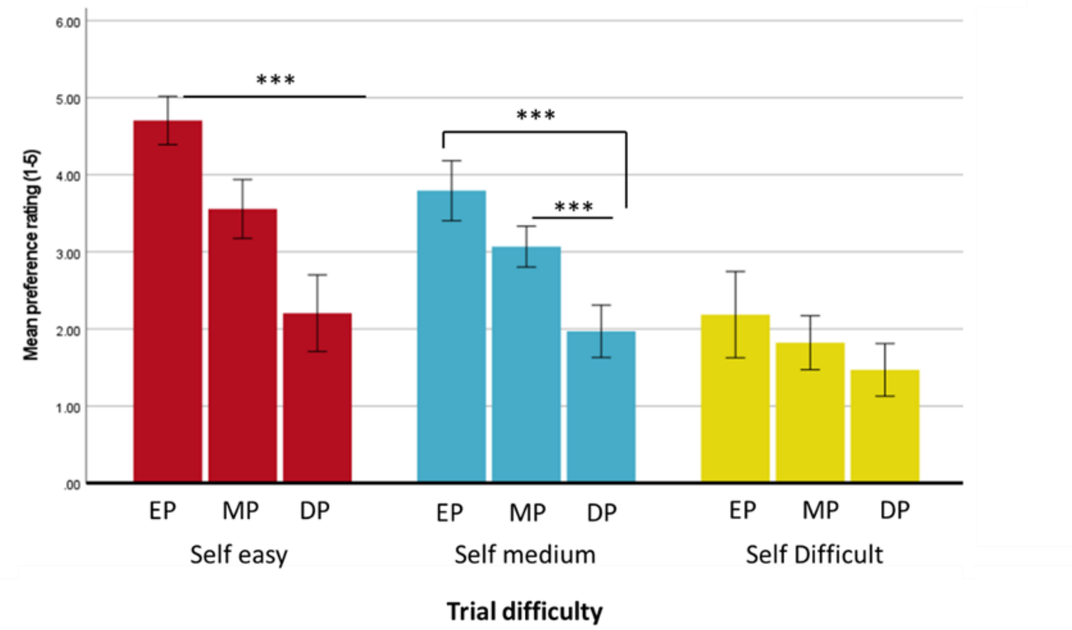

Figure 12. Ratings of preference for self and partner difficulty in Experiment 6. Error bars denote ±2 SE.

A 3x3 rANOVA (see Figure 12) showed a significant main effect of self-difficulty ($F(1.42, 27.7) = 38.54$, $\eta_p^2 = 0.67$), partner difficulty ($F(1.34, 25.42) = 32.21$, $\eta_p^2 = 0.63$) as well as a significant self x partner difficulty interaction ($F(2.67, 50.68) = 14.94$, $\eta_p^2 = 0.44$).

3.8.3 Discussion

The aim of Experiment 6 was to test whether explicitly establishing that a partner is less competent than the participant would make the participant more willing to act co-efficiently or even to choose more difficult trials for themselves. The results showed that participants were taking their partner's effort into account the most when their own task was easy, less so when it was medium and not at all when they had a difficult task. However, it should be pointed out that these results differ from those of Experiments 3-4 and are similar to the ones of Experiment 5. Namely, there seems to be a lack of preference for difficult trials in equal measure, regardless of whether they would be assigned to the deciding participant or their partner. For example, there was no difference in preference between having an easy trial for oneself and a difficult one for one's partner versus having a difficult trial for oneself and an easy one for the partner. While the results do not show co-efficiency across all conditions, on the extreme points of the difficulty range, participants' choices do seem to reflect co-efficiency. This could be due to a general strategy of avoiding costs that are on the extreme end of the spectrum. An alternative explanation is that the results did not show perfect co-efficiency due to how the task was implemented. For example, there was no significant difference in performance between easy and medium-difficulty memory tasks. That might explain why the preference differences for both self and partner's difficulty between some of the easy-medium combinations were not significant. If the different difficulties for the memory task did not result in differences in performance, one should not expect that they should be differentiated in participants' ratings of preference. On the other hand, conclusions based on the connection between performance and ratings should be taken with some reserve. The “partner’s” performance was 2 standard deviations lower than the average performance in the previous experiment yet eight out of twenty participants reported that their partner was equally good or better than they were at the task. This suggests that people might not be good at assessing how good they, or their partner, perform on the task.

3.9 Summary and general discussion

The present study investigated how people distribute mental effort associated with an attentional and a memory task. Specifically, the aim was to investigate four questions (i) is mental effort distributed co-efficiently, egoistically, or altruistically, (ii) does the distribution vary based on self-assessed competence at the task, (iii) does the distribution vary based on performance on the task and (iv) do distribution preferences stay the same across different tasks involving mental effort. Four experiments were conducted using a multiple object tracking paradigm. The goal of the first two experiments was to establish a suitable manipulation of mental effort. The first experiment confirmed that participants’ accuracy on MOT decreased as the number of targets increased and that participants’ assessments of difficulty largely matched the difficulty manipulation. Since expending mental effort is sometimes seen as valuable in itself (Eisenberger 1992; Norton et al., 2012) the second experiment was conducted to confirm that participants prefer easier over more difficult trials. The third experiment used a cover story to convince participants that they are performing the multiple object tracking task with a partner. Participants believed that the experiment consisted of three blocks: a tracking block followed by a rating block and finally another tracking block. In the rating block, they were presented with possible tracking tasks which they would rate based on how much they would like to do them in the future and informed that the final tracking block will consist of trials chosen by their ratings. The presented tracking tasks varied in difficulty of the task for the participant and their partner. Participants’ ratings suggested an egoistic strategy of effort distribution with preferences for easier difficulties for themselves and largely no difference in preferences when it came to partner's difficulties. The ratings of preference did not vary based on participants' self-assessed competence at multiple object tracking. Participants who were better at the task in terms of accuracy preferred that their partner would do an easy task over a difficult one but only when their own task was easy. A potential problem with the procedure in Experiment 3 is that participants might have made some assumptions that were not part of the experimental setup. For example, participants might have assumed that their partner is adopting an egoistic strategy and decided to do the same. This was addressed in Experiment 4, where participants were informed that their partner would not be doing the rating task. This resulted in minimizing the partner's effort to a certain extent, but more weight was still placed on one's own effort.

Experiments 5 and 6 extended the previous findings using a working memory task. The results showed a similar pattern, with one's own difficulty given priority overall, but participants were more willing to consider the partner's difficulty. This was particularly apparent in the most difficult trials where participants placed equal weight on their own and their partner's difficulty. Curiously, in Experiment 6, the fact that almost half of the participants did not believe that the partner was real did not produce any qualitative change in the results. This suggests that at least to some extent the presented future trials were judged as if they were hypothetical, rather than trials which would be allocated to the participant and their partner in a later stage in the experiment.

A potential limitation of the reported experiments that should be noted is that they do not establish a firm or quantifiable relationship between effort and task difficulty. While effort tracks difficulty, it does not necessarily always match it (Brehm & Self, 1988; Inzlicht et al., 2018). For example, while remembering seven items is clearly more difficult than remembering three, the magnitude of the difficulty increase is not completely clear. This is especially salient in

Experiments 5 and 6, where participants did not have a strong preference for easy over medium trials, even for themselves. Going forward, it might be useful to delineate difficulties based on a common criterion. This would allow more informed comparisons between different difficulty levels as well as potentially comparing difficulties in multiple object tracking and memory experiments with each other. An example in this direction would be the study by Potts and colleagues (2018) where they established that using subjective time needed to complete a task allows comparisons between difficulties of a wide range of physical and cognitive tasks.

Two key differences between the presented experiments and previous findings of co-efficiency should be pointed out. In the experiments reported here, participants were judging whether they want to do the task while in previous studies (Török et al., 2019; 2020; Strachan & Török, 2020) the task was presented, and they chose *how* to complete it. It is possible that given the choice, people will choose egoistically but when performing a task, they will try to optimize group performance. A follow-up experiment could present participants with a tracking or memory task in which they could divide effort co-efficiently or egoistically without the option to avoid the task altogether. For example, they could be presented with two possible tasks, one of which would be easier for them, but the overall difficulty would be higher while the other one would be easier for them but with a higher overall difficulty. After making the choice of task, the participant, and their partner, would immediately solve the chosen task. A further benefit of this modification would be that it would eliminate ambiguity about which actions are co-efficient. In the current setup, there is no single trial which is co-efficient. Rather, co-efficiency might arise through a pattern of decisions across trials. The second difference was the setting in which the experiments were conducted. Previous findings result from live experiments in the laboratory whereas most of the ones reported here were online based (due to the Covid pandemic at the time). While Experiments 5 and 6 were done live, participants performed the tasks in separate rooms. It is possible that doing the experiment face-to-face with salient contributions of each partner increases willingness to act co-efficiently.

4 Chapter 4: Joint efficiency informs action prediction

4.1 Introduction

If I were to see my friend holding a cup near the top using only two fingers, I might assume that the cup has hot coffee in it. Likewise, most people observing the same event would probably not conclude that my friend thinks that this is the best way to pick up a cup but, rather, they would try to find alternative explanations for their behavior. For the most part, we interpret others' behavior as rational. We assume that they have mental states – goals that they desire to reach and beliefs on how to best reach those goals (Dennett, 1987). This attribution of rational behavior driven by goals and constraints emerges early in childhood (Csibra & Gergely, 2007). Six-month old infants observing an agent trying to reach an object are surprised if the agent jumps when there is no obstacle in their way (Gergely et al., 1995; Csibra et al., 1999) or if the agent jumps higher than would be necessary to clear the obstacle (Liu & Spelke, 2017). Older children are capable of attributing more complex goals to agents. In particular, children between five and seven years old can take into account an agent's preferences and competences. In a study by Jara-Ettinger and colleagues (2015), children saw an agent reaching for various objects placed on boxes of different heights. If one of the objects was on a taller box and the agent reached for it, children inferred that the agent preferred that object over the one on a lower box. Such expectations of efficient actions have also been found in other species. For example, macaques are surprised if they observe a human taking an item using an inefficient trajectory (Rochat et al., 2008).

Expectations of efficiency have good grounds and reflect the tendency of humans to minimize effort. When making free choices between reaching actions, people tend to choose the one that is the biomechanically easiest in terms of least energy expenditure and the shortest required trajectory (Anderson & Pandy, 2001; Cos et al., 2011; 2014; Todorov & Jordan, 2002; Todorov, 2004). When moving objects, people try to achieve a comfortable posture at the final position of movement (Rosenbaum & Jorgensen, 1992) and avoid extreme joint angles when possible (Rosenbaum et al., 1990).

In social situations, when people interact, they tend to forgo individual efficiency if it would reduce the effort of others. Santamaria and Rosenbaum (2011) demonstrated that whether and how long people hold doors open for others is influenced by how close others are to the door and how many people are approaching. The closer people were to a door, the more likely that another person would open it for them, and they held it open longer if more people were following behind them. In turn, people approaching the door would speed up to reduce the effort of the person holding the door open. Santamaria and Rosenbaum (2011) interpret these results in terms of a *shared-effort model*; the door will be held open as long as holding it decreases the effort of the entire group to a larger extent than it increases the effort of the individual holding it. In support of the shared-effort model, other studies found that when passing objects to others, people do so in a way that will result in a comfortable end-state posture of their partner (Gonzalez et al., 2011; Meyer et al., 2013). In a study by Dötsch and Schubö (2015), participants had to cooperate in moving a cup from one end of a table to the other. The first participant would place the cup in the middle of the table, halfway between themselves and their partner after which the partner would pick up the cup and place it on their end of the table. When placing the cup in the middle, participants tended to rotate it so that the handle was facing the other and its orientation would make it easier for their partner to pick it up.

In an effort to directly test whether people act co-efficiently, Török and colleagues (2019) conducted a study where the participants performed a sequential object transport task on a touchscreen. Participants were presented with different layouts of barriers and were tasked to cooperatively move a ball from one location to another. The first participant would bring the ball from the starting point on the bottom of the screen to one of two possible openings in the middle, after which the second participant would pick up the ball and move it to the end position. The layouts of the barriers were varied so that the individually shortest paths and overall shortest path could either match or mismatch. In the congruent condition, the individually shortest path was also shortest for the dyad while in the incongruent condition, taking the individually shortest path meant that the overall path for the dyad was longer. Participants tended to choose individually shorter paths in the congruent condition and individually longer paths in the incongruent one. These findings showed that people take upon themselves some additional effort if that would mean minimizing the total effort of the group. A subsequent study (Török et al., 2021) confirmed that this result was due to participants placing equal weight on their effort as they did for their partner's.

While these studies on joint action suggest that people tend to act co-efficiently, a much less investigated question is whether we expect others to act that way in joint actions and if this expectation can inform action predictions. In a recent study, Mascaro and Csibra (2022) presented 14-month olds with a scenario similar to the ones in (Török et al., 2019). In the experiment, two animated agents (shapes) were cooperatively moving an object such that the left agent would hand the object to the right agent who would then carry it to the final location. The agents were divided by a wall which had two openings through which the left agent could hand objects to the right agent. One of the openings was closer to the left agent while the other was farther away. The left agent always chose the individually efficient opening. The wall configurations were varied so that in one condition (collective coherent), the left agent's individually efficient choice was also efficient for the entire dyad while in the other condition (collective incoherent) it was not. Infants were surprised by the incoherent condition, which was reflected in longer looking times compared to the coherent condition. Surprise at individually efficient but jointly inefficient choices of action suggests that the 14-month-olds were expecting co-efficiency.

When it comes to predicting actions, expectations of individual efficiency inform prediction of trajectories of ongoing movements. Hudson and colleagues (2018) showed participants a hand moving to pick up an object. The hand would disappear mid-transit and participants were asked to report the last location of the hand's index finger. The time of the hand's disappearance was semi-randomized to introduce ambiguity about its final location. In some conditions there was an obstacle between the hand and the object and in others, the path was clear. When there was an obstacle between the hand and the object, the last location of the hand was reported as higher than when the path to the object was clear. In reality, the last locations in both conditions were the same. This effect was reduced if the hand was replaced with a non-intentional agent such as a ball (McDonough et al., 2019). A subsequent study (McDonough et al., 2020) showed that hand shape also modulated predictions. Participants saw a hand approaching two objects, one affording a precision grip (a strawberry) and the other a power grip (an apple). The hand could either be in a precision grip or power grip configuration and would disappear so that its distance from the strawberry and apple was the same. When the hand was in a precision grip, it was reported to disappear closer to the strawberry and when it was in a power grip it was reported closer to the apple. These results have been interpreted within the predictive coding framework (Kilner et al., 2007, Friston 2010, Clark, 2013). According to this account, actions are represented hierarchically, with high-level expectations about overall goals informing expectations of sub-goals which in turn inform predictions of kinematics. Within this framework, action prediction can be described along the lines of Bayesian inference. When predicting an action, prior expectations are taken into account (in this case, efficiency), as well as the location of the stimulus. If the location of the stimulus (e.g. a moving hand) is not ambiguous and does not match prior expectations, a prediction error is generated, fed back up the chain and higher-level predictions are then modified. On the other hand, if the prior is strong enough and the final location of the stimulus is sufficiently uncertain (in this case, disappearing at an unpredictable time and location), the final location of the stimulus will be perceived as closer to the prior than it was.

These studies suggest that expectations of individual efficiency inform on-line prediction of ongoing actions. The aim of the following experiments was to test whether people have expectations of joint efficiency when observing an interaction and whether these expectations inform action prediction. To test this question, the experiments utilized a phenomenon called representational momentum, previously used to test if individual efficiency informs action perception (Hudson et al., 2019; McDonough et al., 2018; 2019). Representational momentum is a tendency to remember the final location of a moving object further along its trajectory than it actually is. In the original experiment, participants were presented with a series of images of a rectangle. The first image depicted the rectangle in an upright position with each subsequent image showing it at an increasingly rotated angle. After four images were presented, the task was to compare the final image to a probe rectangle. Participants tended to report the final rectangle as more tilted to the direction of rotation than it was (Freyd & Finke, 1984). One proposed explanation of representational momentum is that it is due to eye-movements overshooting the location of the object as it suddenly disappears, leading to the object being remembered farther along the trajectory. Supporting this interpretation are studies by Kerzel (2000) and Kerzel and colleagues (2001) where they introduced two tracking conditions: in one, participants tracked the object as it moved while in another, they were fixating on a stationary cross. When participants were not tracking the object, representational momentum was reduced or eliminated. Other studies have cast doubt on whether eye movements can provide a full explanation of the effect. For example, the magnitude of representational momentum is influenced by the size of the object and its velocity (Hubbard & Bharucha, 1988; Kozhevnikov & Hegarty, 2001) which cannot be accounted for by eye movements alone. Perhaps more strikingly, the magnitude of representational momentum is influenced by conceptual knowledge; movement of an ambiguous object labeled as a “rocket” will produce stronger representational momentum than if the same object is labeled as a “steeple” (Reed & Vinson 1996; Vinson & Reed, 2002). Hubbard and Favretto (2003) showed that the inferred source of movement of the target also modulated the magnitude of representational momentum, reflecting expectations rooted in naïve physics.

Taken together, these studies suggest that representational momentum is affected by high-level expectations and is sensitive to conceptual knowledge, making it a good measure of expectations of action efficiency. The following experiments utilized representational momentum to test whether people have expectations of co-efficiency and if those expectations inform online predictions of ongoing actions.

4.2 Experiments

The experiments consisted of observing a cooperative game where two observed agents had to drive a ball from one end of the screen to the other. All the experiments used the same layouts as Török and colleagues (2019) and the movement of the ball was modelled from the movement data provided by Török and colleagues. In the first two experiments, the bottom agent could choose a path that is individually efficient or inefficient. Crucially, some of the individually inefficient paths were jointly efficient. In some of the trials, the interaction was cut short, with the ball disappearing before reaching its destination, at which point participants had to report the ball's last location. The main prediction was that the bias in perceptual estimates would be influenced both by expectations concerning the individual efficiency of the agent initiating the interaction as well as the joint efficiency of the entire interaction.

4.3 Experiment 1

4.3.1 Materials and Methods

Participants

An a-priori power analysis revealed that a sample size of 88 is required to achieve 90% power to detect a small-medium effect size (Cohen's d = .35) with an alpha level of .05 (two-tailed). 111 participants were tested online using the Prolific platform (https://www.prolific.co/), 22 of which were excluded based on predefined criteria specified in the Data Analysis section. The inclusion criteria were fluency in English, an approval rating at or above 95% on Prolific, an age range between 18 and 40 years, and having normal or corrected-to-normal-vision. The exclusion criterion was participation in the pilot version of this experiment.

Design

The experiment used a 2 x 2 x 2 within-subject design with the factors of *trajectory* (whether the bottom player moved a ball leftwards or rightwards), *bottom barrier* (whether the bottom player had a barrier on one of their potential paths or not) and *top barrier* (whether the top player had a barrier on one of their potential paths). This created a setup where, based on the bottom agent's choice, the bottom barrier would influence their own efficiency while, again based on the bottom agent's choice, the top barrier could influence the efficiency of the top player. The joint efficiency was influenced by different combinations of the top and bottom barriers. In particular, no bottom barrier combined with a top barrier would make the rightward path individually inefficient for the bottom agent but jointly efficient. The dependent variable was the perceptual error on the X axis in each trial, calculated as the difference between the actual point of the ball's disappearance and the point reported by the participant. Given that the gaps through which the ball could be passed were parallel, the data from the Y-axis did not allow for straightforward predictions between conditions and is therefore not reported. If participants expect the bottom agent to act efficiently for themselves, in conditions where there is no bottom barrier (Figure 1, b) and d)), the rightward bias should be stronger than when there is a barrier on the bottom (Figure 1, a) and c)). Expectations of the bottom player considering the efficiency of the top player would result in a main effect of the top barrier. Namely, when the rightward bias should be stronger in conditions where there is a top barrier (Figure 1, b) and c)) than where the way on the top is clear (Figure 1, a) and d)). A significant interaction between top and bottom barrier would suggest that participants formed predictions based on the combined action costs of the two observed players.

Stimuli

The stimuli consisted of four layouts taken from Török et al. (2019), an image of a ball with a cursor superimposed at its center and a series of animations of the ball's movement derived from the data provided by the authors of Török et al. (2019). The static images were of a 1920 x 1080 resolution, and the size of the ball was 192 x 108 pixels. The layouts had four different configurations of barriers: *barrier-top, barrier-bottom, barrier-both* and *barrier-none* (see Figure 1). While Török et al. (2019) varied the length of the barriers, in this experiment the barrier could either have the maximum length presented in Török et al. (2019) or would be absent (corresponding to 1 and 0 barrier lengths in Török et al. (2019)).

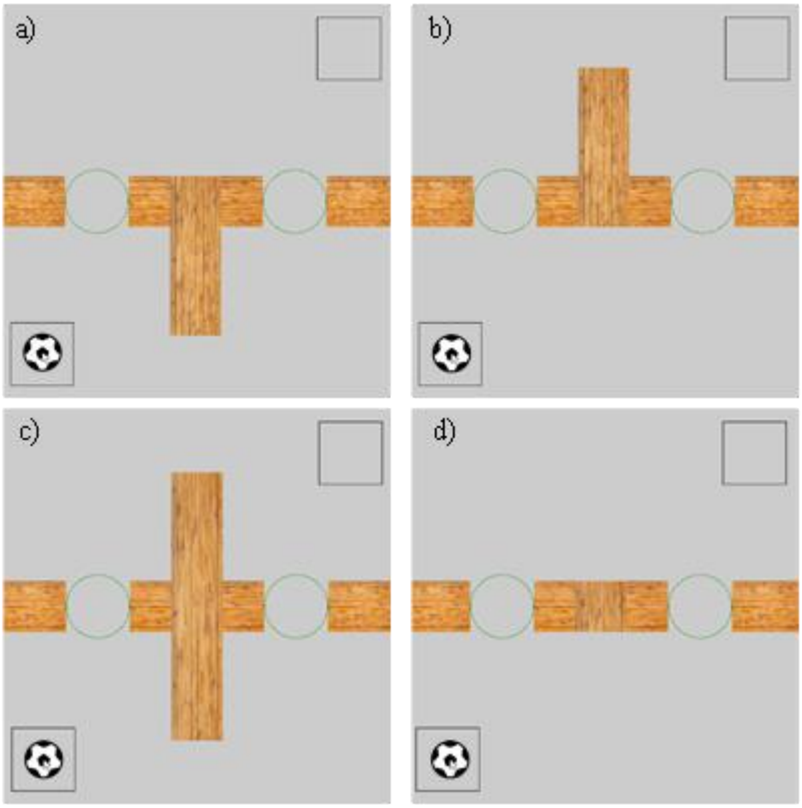

Figure 1: Possible barrier configurations shown with the ball and cursor. Barrier bottom (a), barrier top (b), barrier both (c) and barrier none (d).

The animations of ball movement were derived from participant data from Török et al. (2019). The trajectories of the ball were split into 80ms bins and the average of the X and Y axes positions in each bin was taken to create each frame of the stimuli. *Full trajectories* consisted of the bottom agent driving the ball to one of the central apertures, followed by the top agent's mouse cursor picking up the ball and leading it to the top right of the screen. There were 8 full trajectories. One for each of the barriers (top/bottom/both/none) x trajectory (left/right) combinations. *Test trajectories* showed the bottom agent leading the ball for a minimum of 320ms and a maximum of 560ms. There was a total of 12 different test trajectories. Three were taken from trials where a participant was driving the ball to the right while avoiding the barrier (curved right trajectories), an additional three when they were driving the ball rightwards without an obstacle (straight right) and six were taken from trials in which a participant drove the ball towards the left aperture (straight left).

Procedure

The experiment was coded in Inquisit and hosted on the Millisecond platform (https://www.millisecond.com/). Before the experiment started, participants went through instructions telling them that they would be viewing and making judgments about recordings from a game that two people had played previously. The game consisted of the bottom player bringing a ball to one of the central gaps and the top player picking it up and driving it to the top right of the screen. They were told that the object of the game was to cooperatively bring a ball from the bottom of the screen to the top as fast as possible but without the ball bumping into any obstacles. Finally, for each of the possible layouts, the individual difficulty for the top and bottom player was explained (e.g.,, that for barrier-top the left path is shorter for the bottom player but longer for the top player). After the instructions, participants completed two blocks of practice trials. In the first block, they were presented with trials showing a full interaction, with the bottom agent's cursor seen driving the ball to one of the openings, followed by the top agent’s cursor moving from the top right corner to pick up the ball and take it back to the top right of the screen. Once the interaction was complete, participants were asked if the bottom agent chose the opening/path which they expected, to which they responded by clicking on a “YES” or a “NO” box on the screen. The first practice block consisted of 8 full trials (one for each of the barrier top/bottom/both/none) x trajectory (left/right) combinations.

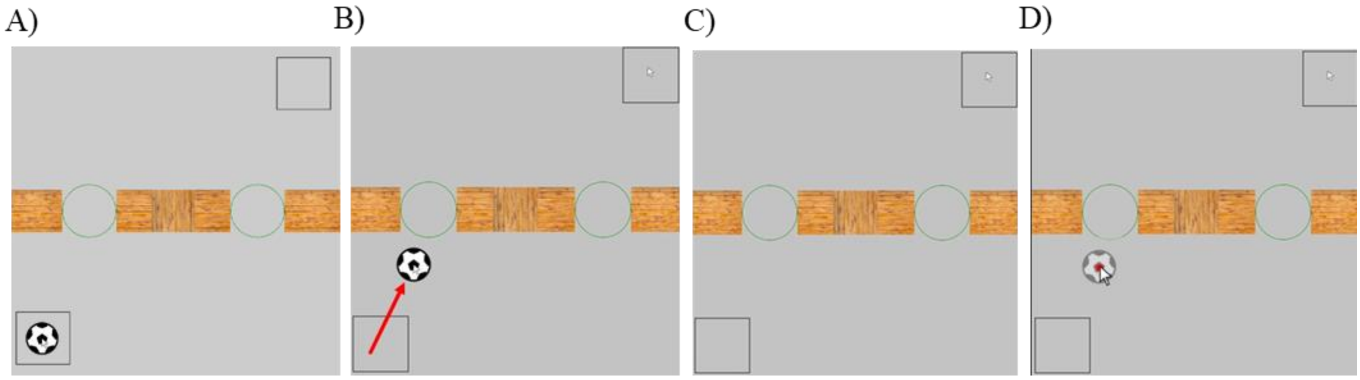

Figure 2. Schematic representation of a trial. The ball appeared and stayed stationary for one second (A), after which it started moving (B) and disappeared (C). The participants' task was to click on the tip of the mouse cursor that was leading the ball (C). The red arrow is illustrative of a possible direction of movement and was not shown in the experiment. The translucent image of the ball (D) is a reference to the ball's last location and was not shown in the experiment.

The second practice block (see Figure 2) had the same conditions and number of trials as the first. In this block, participants would see a layout and were asked to press the space bar once they decided which path the bottom agent should take. Once they pressed space, the ball with a mouse cursor appeared on the screen, and stayed stationary for one second after which it started moving. During this time, the participants' mouse cursor was invisible to prevent them from using it to track the ball. After a variable amount of time (320, 400, 480, 560ms), the ball would disappear, and participants needed to click on the last location of the mouse cursor that had been leading the ball. Presentation durations were varied to increase the ambiguity of the last location of the stimulus and therefore increase the influence of prior expectations. After a response was made, the scene would be replaced with an attention cross which needed to be clicked so that the participants' mouse was re-centered. Once the cross was clicked, the next trial would begin.

Once practice was concluded, the experiment began. Trials followed the same procedure as in the second practice block. There was a total of 96 trials, with 12 trials for each combination of barriers (top, bottom, both, none) combined with the two trajectories (left/right). In random order, the eight full interactions were interspersed, and participants were asked if the bottom agent had chosen the path they were expecting. Halfway through the experiment, there was a 30 second break. The median time to complete the experiment was 24 minutes.

Data Analysis

Data exclusion criteria were adopted from previous studies using representational momentum to study action prediction (e.g., Hudson et al., 2016; McDonough & Bach, 2023). For each participant, individual trials were excluded if the first response (pressing spacebar) time was less than 200ms or their second response (mouse click) was less than 200ms or more than 2000ms or if the Pythagorean distance between the real final coordinates and participants' mouse response was more than 300 pixels (in a standardized coordinate system). Participants who had too few trials remaining (<50%) were excluded. Individual participants were excluded if their average Pythagorean distance between the real final coordinates and their responses exceeded 150 pixels or if the correlation between the real final coordinates and their responses was less than 0.7.

For the trials where participants had to report the ball's last location, perceptual error on the X-axis was calculated as the participant's reported last location minus the ball's real disappearance point, in pixels. Positive values indicate perceptual errors to the right while negative ones indicate errors to the left (a value of 0 would mean an accurate report of the ball's last seen location).

4.3.2 Results and Discussion

The first thing to note is that there was a general rightward bias across all conditions (see Figure 3). This is consistent with previous experiments showing that representational momentum is generally stronger when objects are moving from left to right than vice versa (Halpern & Kelly, 1993; Kerzel, 2003). The rightward effect in those studies was relatively small (0.77° and 0.18° respectively) and given that the current experiment was conducted online, it is not possible to re-calculate the bias in terms of visual angle. Moreover, other studies found no differences in the magnitude of representational momentum between left-to-right and right-to-left motion (Hubbard 1990; 1995). A more probable reason for the overall rightward bias is that the starting position of the bottom player was at the bottom left of the screen so they would have to move the ball somewhat rightwards even if they were driving the ball to the left gap (see Figure 4). This was reflected in a paired-samples t-test showing that the magnitude of rightward bias was stronger for rightward than leftward trajectories ($t(88) = 32.91$, $p < 0.001$, $d = 3.49$).

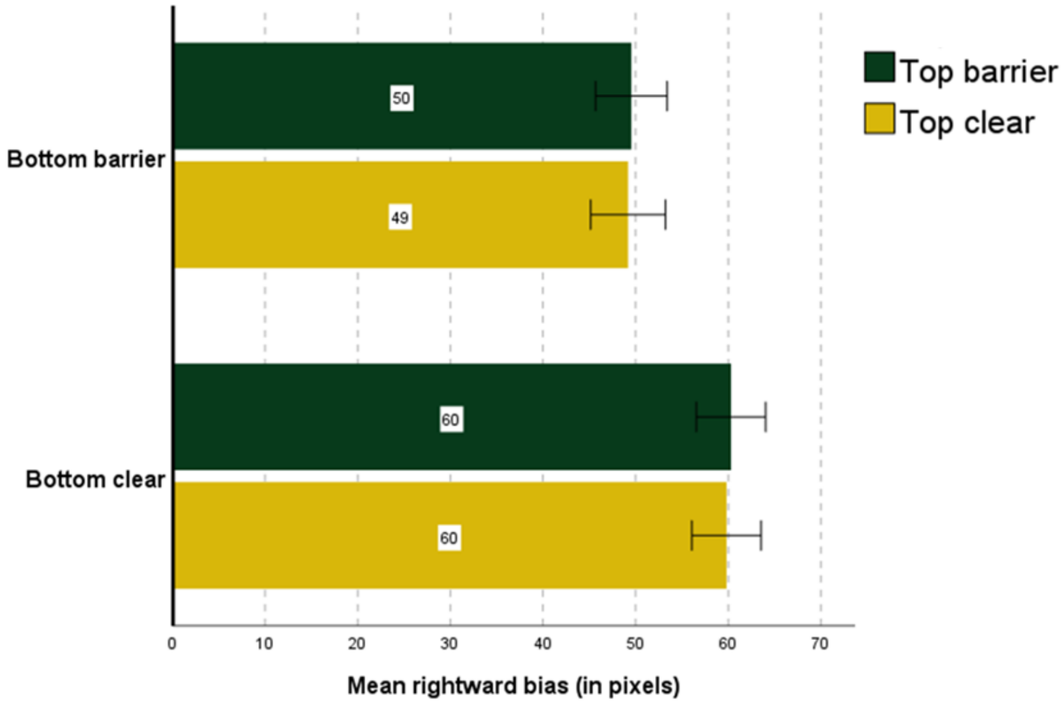

Figure 3. Mean rightward bias in pixels for each condition represent 2 SEM.

A 2x2 rANOVA was conducted with the factors of *top barrier* (barrier/clear) and *bottom barrier* (barrier/clear) with perceptual errors (in pixels) as the dependent variable. Results showed a significant main effect of bottom barrier ($F(1, 88) = 81.797$, $p < 0.001$, $\eta_p^2 = 0.48$) while the main effect for top barrier ($p = 0.63$) and the interaction were not significant ($p = 0.93$).

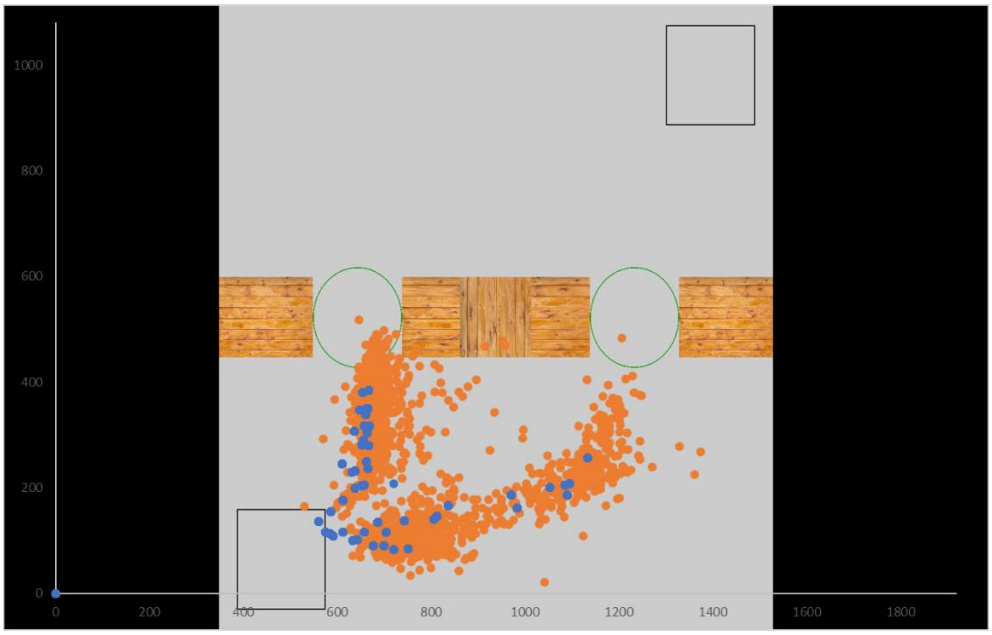

Figure 4. Participants' responses (before exclusions) in conditions with no barriers. Blue dots represent disappearance points of the ball while the orange dots represent participants' responses.

The significance of the main effect of the bottom barrier means that the participants' bias was modulated based on whether the bottom agent had a barrier or not, with the bias towards the right being stronger when there was no barrier on the path to the right, replicating previous findings (Hudson et al., 2018, McDonough et al., 2019). The bias showed that participants predicted the bottom player to choose the left or right path depending on whether the bottom player had an obstacle. This could be taken to suggest that participants expected Player 1 to consider Player 2's action costs in addition to their own; if they expected the bottom player to act purely egoistically, their expectations would be that the bottom player would always choose the left path given that the left path was shortest for the bottom player in all scenarios. Put differently, if the bottom player would not take into account the top player's obstacles, the individually rational path should be the left one, which should not result in a significant main effect of bottom-barrier. It has to be acknowledged, however, that during the practice and the experiment participants saw both rightward and leftward trajectories, which would have lowered their expectations of seeing only leftward ones. The fact that the main effect of the top barrier was not significant also speaks against the interpretation that participants expected Player 1 to consider Player 2's action costs . If people expected the bottom agent to take the effort of the top one into account, the main effect of the top barrier would have been significant. In particular they should have expected the bottom agent to go rightwards if the top agent had a barrier. Finally, the non-significant interaction suggests that participants' predictions were not influenced by co-efficiency.

Two more likely interpretations are that participants either did not represent the scenario as a joint task between the players or that they did not pay attention to the entire layout of the barriers. These possibilities are not mutually exclusive, and both might have resulted from the structure of the task. Apart from the practice trials with full trajectories at the beginning of the experiment, participants only encountered trials where they had to report the end position of the ball while it was still in the bottom part of the screen. In order to complete the task successfully, they did not need to attend to the part of the screen where the barriers (or lack thereof) of the top player were located. It is possible that this setup led to participants focusing only on the bottom part of the screen. To account for this possibility, participants' answers to full trajectories which were periodically shown during the experiment were analyzed (see Figure 5).

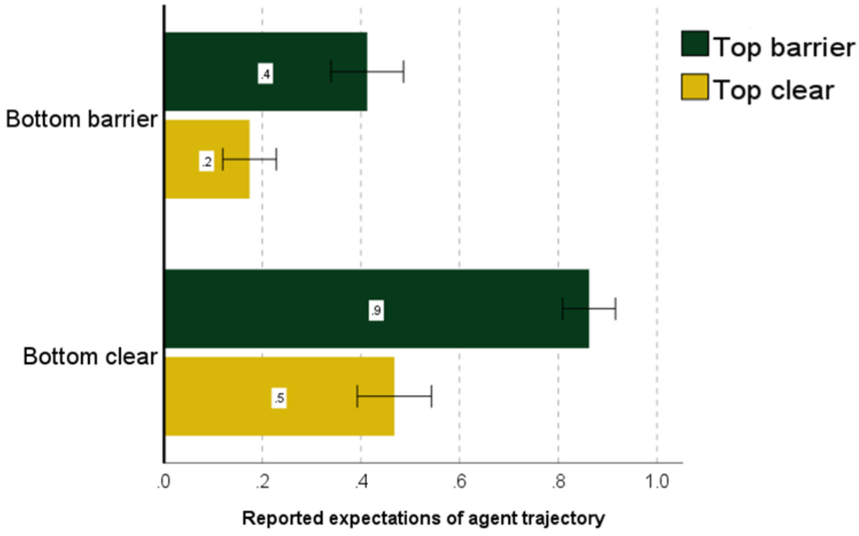

Figure 5. Mean response to whether participants expected a leftward or rightward trajectory per condition. The coding is 0 for expected leftward trajectory and 1 for a rightward one.

For each condition for each participant, reported expected trajectory was coded as 0 (reported expecting leftward trajectory) or 1 (reported expected rightward trajectory). For example, if the trajectory was a leftward one and the participant reported they did not expect this trajectory, it was coded as 1. A 2x2 repeated-measures ANOVA was ran with barrier top (barrier/clear) and barrier bottom (barrier/clear) as independent variables and the mean expected trajectory as the dependent variable. Results showed a main effect of barrier bottom ($F(1, 108) = 145.08$, $p < 0.001$, $\eta_p^2 = 0.571$), barrier top ($F(1, 108) = 104.09$, $p < 0.001$, $\eta_p^2 = 0.491$) as well as a significant interaction between the two ($F(1, 108) = 6.76$, $p = 0.011$, $\eta_p^2 = 0.059$)). Participants expected rightward trajectories more often when the bottom agent had a clear path, and the top agent would have to go around a barrier to pick up the ball in the left gap. In other words, they expected the bottom agent to act co-efficiently and take somewhat longer rightward paths if that would mean that their partner would not have to take a significantly longer path around a barrier. These results support the interpretation that the perceptual bias was not modulated based on the top agent's barriers. This could be because participants attended the entire screen for explicit predictions but not when reporting the location of the disappearing ball. An alternative explanation would be that there is a dissociation between reported expectations and the perceptual bias based on those expectations. Previous research suggests that this is unlikely. For example, using a similar paradigm, Hudson et al. (2016) directly manipulated participants' expectations by showing a hand either approaching an object or moving away from it. The hand's movement was preceded by the grasping agent stating that they will take the object or leave it where it is. The magnitude of representational momentum increased when the direction of movement matched the declared intention and decreased when it did not. This shows that high-level expectations can inform prediction in a way that is measurable by using representational momentum.

4.4 Experiment 2

The aim of the second experiment was to adapt the procedure so that the participants would be more likely to pay attention to the entire scene, rather than just the bottom half of the screen. The structure of the experiment was largely the same as the previous one with a few modifications. Firstly, the full interactions were only presented during the instruction phase of the experiment. Secondly, catch trials were added. On these trials, the participants were shown the layout and when they pressed the space bar to indicate that they were ready, the layout disappeared, and they were asked "Which path should the bottom player take while taking into account the top player?". This was done to encourage participants to look at the entire layout before indicating that they are ready; including the catch trials meant that when indicating that they are ready participants did not know if on a particular trial they would be asked about the paths, to answer which they would need to look at both the top and the bottom barrier configuration. There was a total of 24 catch trials presented at random times. To increase believability that the game was played by real people, avatar images of the two agents were positioned next to their starting locations. The last modification was adding another break and showing the instructions and full interaction videos during both breaks. This was done to remind participants that the scenario they were observing had two agents and that the agents' goal was to cooperatively move the ball.

4.4.1 Materials and methods

Participants

The inclusion and exclusion criteria were the same as in the previous experiment. Due to an error in scheduling participants on Prolific, instead of the target sample of 88 participants, 91 were recruited. Since excluding the final three participants did not produce any qualitative change in the results, they were kept in the analysis.

4.4.2 Results and discussion

As in the previous experiment, representational momentum was calculated by subtracting the position of the ball on the X-axis from the position of the participants' response. The data was then aggregated and a 2x2 within-subjects repeated-measures ANOVA was conducted with factors of *bottom barrier* (barrier/clear) and *top barrier* (barrier/clear) (see Figure 6). The main effect of the bottom barrier was significant ($F(1, 90) = 27.83$, $\eta_p^2 = 0.236$, $p < .001$) with a higher rightward bias when there was no barrier compared to when the barrier was present. This suggests that participants considered it more likely that the bottom agent would go rightwards when there was no barrier as opposed to when there was. This finding reflects predictions of individual efficiency – it is more efficient for the bottom agent to move towards the left gap when there is a barrier to their right. Furthermore, this finding is a conceptual replication of previous studies showing expectations of individual efficiency influencing representational momentum (Hudson et al., 2019, McDonough et al., 2019).

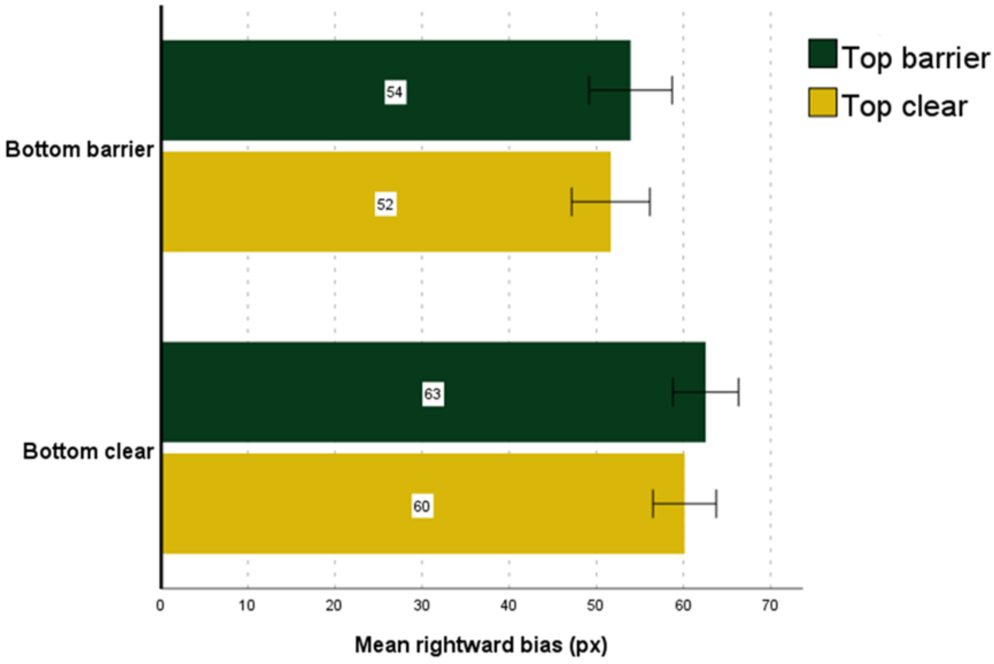

Figure 6. Mean rightward bias (in pixels) per condition.

The main effect of the top barrier was also significant ($F(1, 90) = 5.35$, $\eta_p^2 = 0.056$, $p < .05$), with the rightward bias being stronger when the top barrier was present compared to absent. If the top agent had to go around a barrier to reach the ball at the left opening, participants predicted that the bottom agent would be more likely to pick the right path over the left. It is worth noting here that, for the bottom agent, the left path is always individually efficient. In other words, the prediction of the action of the bottom agent was informed by the constraints and potential effort of the top player. The interaction between the top and bottom barrier was not significant ($p > .05$), suggesting that the effects of barriers for the top and bottom agent were additive. While this shows that participants predict that the bottom agent will consider their own and their partner's barriers, this is different from expecting co-efficiency. The two conditions in which co-efficient choices can be made are the bottom barrier + top clear and bottom clear + top barrier. In the first case, the co-efficient choice would be the bottom agent moving to the left, avoiding the barrier. The top agent would have to traverse a longer distance but in the aggregate the path would be shorter than if the bottom agent chose the right path. The second condition is when there is no barrier on the bottom but a barrier on the top. Here, the co-efficient choice would be for the bottom agent to go to the right. If participants expected these co-efficient actions (and they influenced representational momentum), it would have been shown in an interaction.

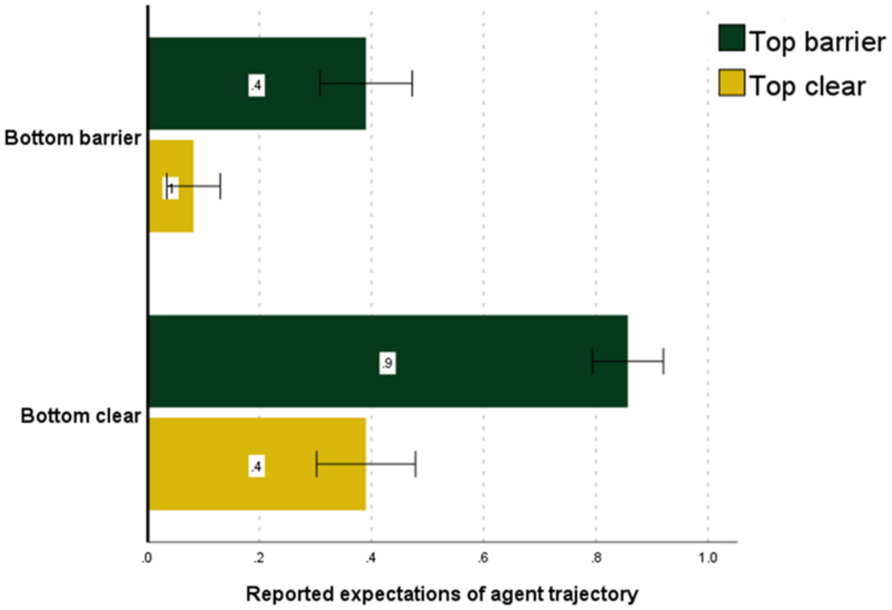

Figure 7. Mean response to whether participants expected a leftward or rightward trajectory per condition. The coding is 0 for expected leftward trajectory and 1 for a rightward one.

For trials where the entire interaction was shown and participants were asked if the agent took the path they expected, a leftward expectation was coded as 0 and a rightward one as 1 (see Figure 7). A 2x2 rANOVA was conducted with the within-subject factors of barrier bottom (barrier/clear) and barrier top (barrier/clear). The main effect of the bottom barrier was significant ($F(1, 90) = 183.94$, $\eta_p^2 = 0.671$, $p < .001$), showing that participants expected the bottom agent to take the left path more often if there was a barrier blocking the right path. The main effect of the top barrier was significant as well ($F(1, 90) = 134.37$, $\eta_p^2 = 0.599$, $p < .001$), showing that when the top agent would have to go around a barrier to reach the left gap, participants expected the bottom agent to take the right path instead. Finally, the interaction between the top and bottom barrier was significant ($F(1, 90) = 4.69$, $\eta_p^2 = 0.05$, $p < .05$). The interaction showed that when the bottom agent's right path was clear and the top agent had a barrier to their left, participants expected the bottom agent to take the right (longer) path (add reference to figure). In other words, when the bottom agent could take a longer path for themselves but significantly shorten the top agent's path, participants expected them to do so. This finding suggests that participants predicted a co-efficient interaction; it mirrors the pattern of results that Török and colleagues (2019) obtained when people played the ball-moving game rather than observed it (Experiments 1, S1 and S2).

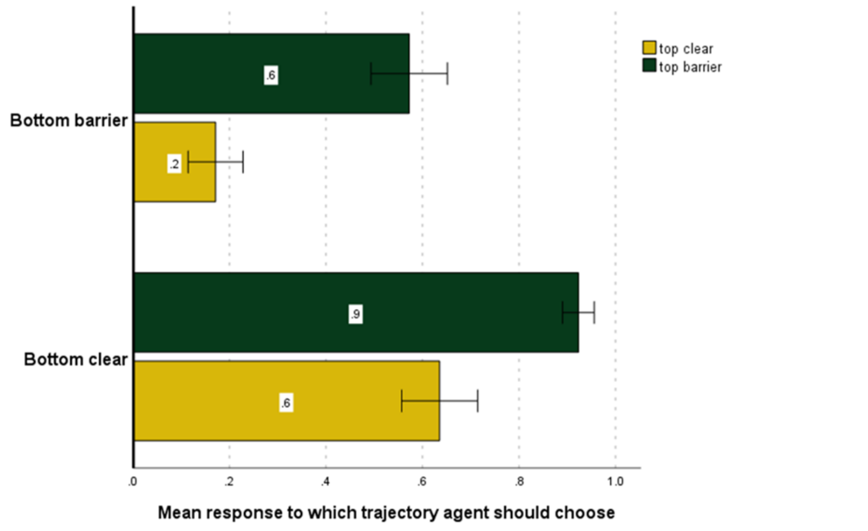

Figure 8. Mean response to whether participants thought the bottom agent should choose the left path (coded as 0) or the right (coded as 1)

For catch trials, participants' answers were again coded as 0 if they indicated that the bottom player should go to the left and 1 if they responded that the agent should take the right path. A 2x2 rANOVA was conducted showing significant main effects of bottom barrier ($F(1, 90) = 254.05$, $\eta_p^2 = 0.738$, $p < .001$), top barrier ($F(1, 90) = 170.84$, $\eta_p^2 = 0.665$, $p < .001$) and a marginally significant interaction between the two ($F(1, 90) = 3.567$, $\eta_p^2 = 0.04$, $p = .058$). The main thing to note here is that the catch trial analysis showed the same pattern of results as participants' explicit expectations, with participants reporting that the bottom agent should take longer paths if it reduces the overall length of the path for the dyad.

In summary, Experiment 2 showed that participants had the same explicit expectations about which path the bottom agent should choose as in the previous experiment. These predictions accurately reflected real path lengths for the dyad. The results thus show that people have explicit expectations of co-efficiency when observing third-party interactions. It remains unclear to what extent these expectations inform predictions of ongoing actions. The measure of representational momentum showed that expectations of individual efficiency of the bottom agent biased participants' judgments of where the ball disappeared. Furthermore, these predictions were also informed by the barriers of the top agent, showing that predicting actions of one person is informed by the constraints of their partner in an interaction. However, the lack of a significant interaction suggests that the overall path length might not have affected the perceptual bias, as would be predicted based on explicit expectations of co-efficiency.

4.5 Experiment 3

The limited extent to which considerations of co-efficiency influenced predictions of ongoing actions in Experiments 1 and 2 could be due to prediction mostly being informed by individual efficiency. If this is true, a replication of Experiment 2 with a single agent should result in predictions being biased by the entire potential path of their movement, resulting in significant main effects of top and bottom barriers and an interaction. Alternatively, it is possible that prediction is informed only or mainly by the immediate (bottom) obstacles. In this case, the bottom half of the layout should dominate predictions, even if a single agent is leading the ball through the entire layout. In Experiment 3, the scenario was modified so that it involved only the bottom agent leading the ball from the bottom left, through one of the central openings and to the end position at the top right of the screen. Participants were told that the bottom agent had to lead the ball as quickly as possible without hitting any of the barriers. Before the experiment, examples of path length were given for each of the layouts. The rest of the experiment followed the same structure as Experiment 2.

Participants

The inclusion and exclusion criteria were the same as in the previous experiment with 88 participants remaining after exclusions.

4.5.1 Results and discussion

The same exclusion criteria for individual trials and participants were applied as in the previous two experiments. Representational momentum was calculated by subtracting the position of the ball on the X-axis from the position of the participants' response. A 2x2 rANOVA was conducted with the within-subject factors of *bottom barrier* (barrier/clear) and *top barrier* (barrier/clear) (see Figure 9). The main effect of *bottom barrier* was significant ($F(1, 87) = 43.78$, $\eta_p^2 = 0.34$ $p < .001$) but neither the main effect of *top barrier* nor the interaction between the two ($ps > .005$).

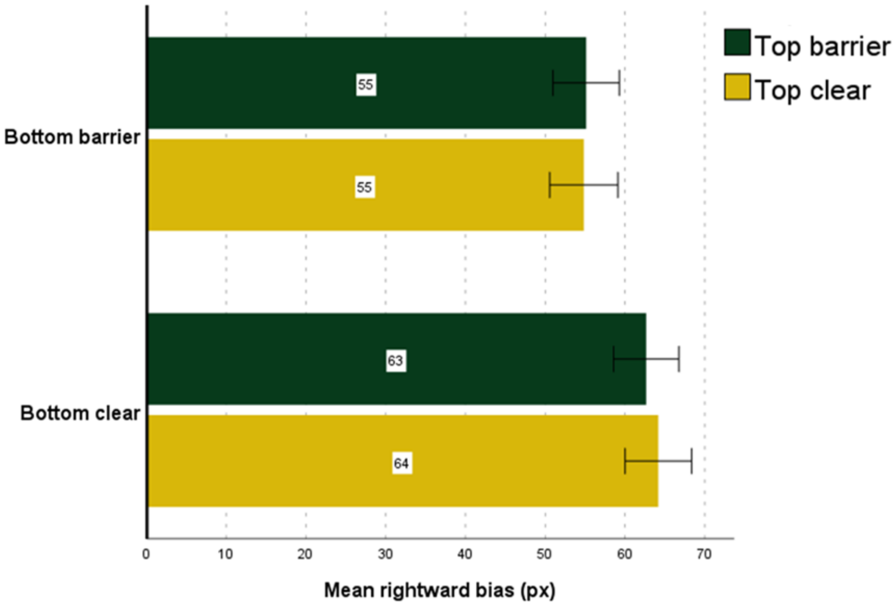

Figure 9. Mean rightward bias per condition (Experiment 3).

This shows that when there was no barrier on the bottom, rightward bias was stronger than when the path to the right was clear, suggesting that participants expected the agent to avoid going around the barrier. From an efficiency standpoint, this is to be expected – in all trials where there is a bottom barrier, taking the left path is either equally or more efficient. The lack of a significant main effect of top barrier or interaction is puzzling, however. Given that the agent has to go through the entire layout, they should choose their path based on both top and bottom barriers. However, this was not reflected in participants' bias, suggesting that they might have focused on the bottom part of the screen.

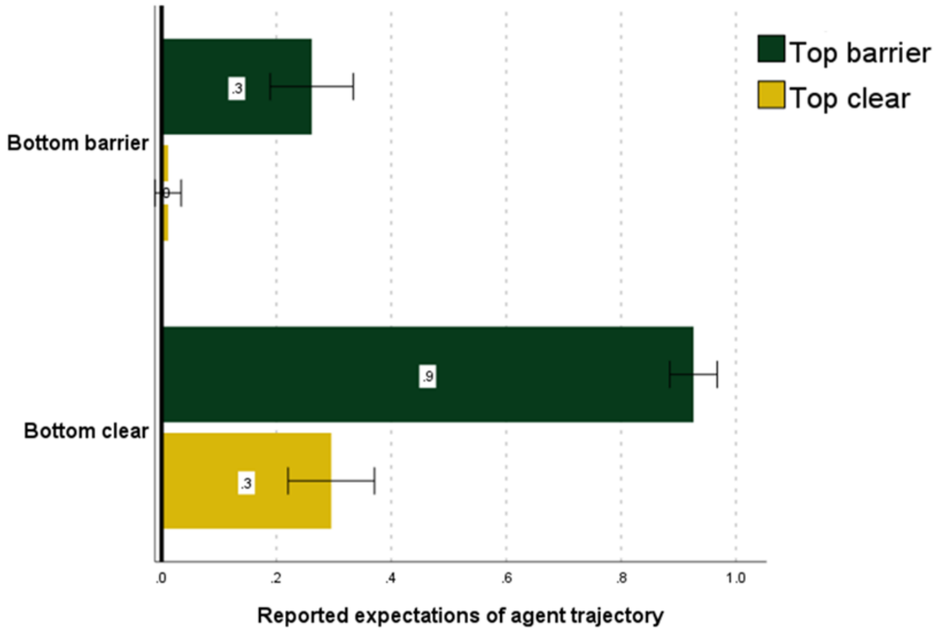

Figure 10. Mean response to whether participants expected the agent to choose the left (coded as 0) or the right path (coded as 1).

The analysis of participants' reported expectations when viewing full interactions revealed significant main effects of bottom barrier ($F(1, 87) = 364.27$, $\eta_p^2 = 0.8$, $p < .001$), top barrier ($F(1, 87) = 359.96$, $\eta_p^2 = 0.8$, $p < .001$) and a significant interaction between the two ($F(1, 87) = 36.32$, $\eta_p^2 = 0.29$, $p < .001$). This confirmed the findings from the previous two experiments: when asked about explicit expectations, participants took into account the configuration of both the top and bottom barrier. For example, in the bottom clear-top barrier condition, the shortest path is the one rightwards whereas in the bottom barrier-top clear condition, the overall shortest path is a rightward one (add reference to figure). Participants' expectations accurately reflected this, and they reported that they expected the agent to take the overall shortest path.

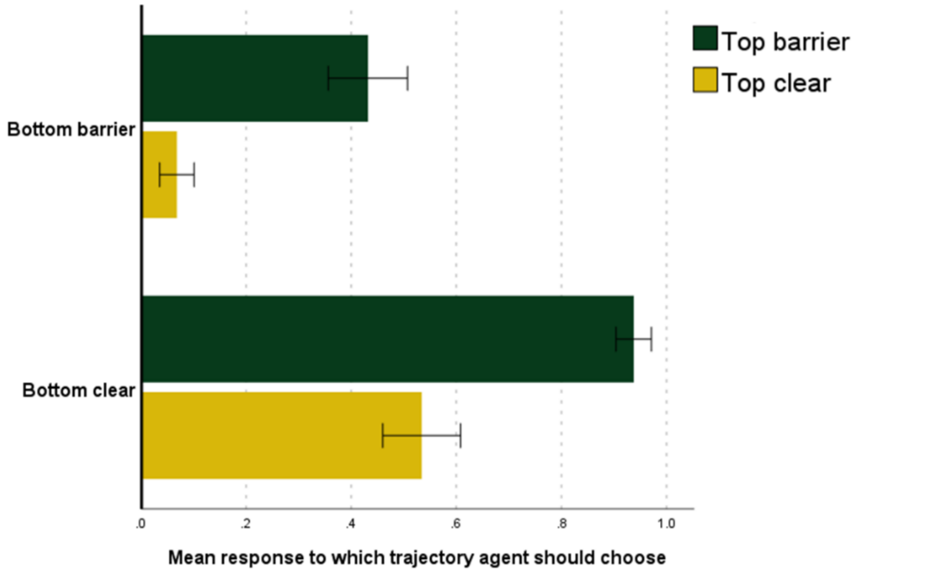

Figure 11. Mean response to whether participants thought the agent should choose the left path (coded as 0) or the right (coded as 1)

For trials where the layout instantly disappeared once participants indicated that they are ready, the analysis showed a main effect of the bottom ($F(1, 87) = 331.67$, $\eta_p^2 = 0.79$, $p < .001$) and top barriers ($F(1,87) = 307.66$, $\eta_p^2 = 0.78$, $p < .001$) but the interaction was not significant ($p > .05$). Similarly as with the full interaction trials, participants reported that the agent should choose the overall shortest paths when taking into account both the top and bottom barriers.

The aim of Experiment 3 was to extend the findings of Experiment 2 by probing expectations of individual rather than joint efficiency. Curiously, the results showed a disconnect between the different measures of expectation. Explicit questions about expectations and prompts to indicate which path the agent should take revealed that participants are both able to ascertain which path is the shortest and that they expected the agent to take it. However, only the presence or absence of the bottom barrier modulated representational momentum, showing that these expectations were not fully translated into modulations of representational momentum. A possible explanation could be the complexity of the layout. When observing an ongoing action on the bottom of the screen the bottom barrier is the one that is relevant for predicting where the ball will be in the following moment. Taken together, the results of Experiment 3 showed that participants expected individual efficiency from the agents but that these expectations did not consistently shape predictions of ongoing actions.

4.6 Summary and general discussion

The present study investigated if people have expectations of individual and joint efficiency and whether those expectations inform predictions of ongoing actions. In three experiments, efficiency was operationalized as the length of a path that an agent (or two agents) would have to follow to move a ball to a set location. In the first experiment, the goal was to investigate if people are sensitive to co-efficient actions when observing a third-party interaction. The results showed that when asked explicitly about expectations, people were sensitive to the overall path length and to the obstacles both agents were facing. On the other hand, action prediction as measured by representational momentum, was only informed by the presence of the bottom barrier. A likely reason for the disconnect between explicitly expressed expectations and action prediction was that the task required participants to attend only the bottom of the screen.

The second experiment introduced catch trials to ensure that participants paid attention to the entire layout as well as additional instructions and avatar images of agents. The results showed the same pattern for explicit judgments as in Experiment 1, but representational momentum was modulated by both the barrier of the top and bottom players, suggesting that action prediction was informed by efficiency of both the top and bottom players. However, if action prediction was guided by jointly efficient paths, it would have been biased towards actions that would incur a small cost for the bottom agent while benefiting the top one.

Experiment 3 tested whether results of the previous two experiments were due to expectations of joint efficiency not informing action prediction, or, rather, whether the results reflect a more general process where predictions are dominated by barriers closest to ongoing movement. To test this, the procedure in Experiment 3 was modified so that only the bottom agent performed the whole task. While explicit judgments and reports of expectations were similar as in the previous two experiments, judgments about the location of the disappearing object were influenced only by the presence or absence of the bottom barrier. This suggests that the results of the previous two experiments were not due to joint efficiency being uninformative for action prediction. If this was the case, in the experiment with a single agent, representational momentum should have reflected individual efficiency across the entire path rather than just for the bottom barrier. A more likely explanation for the results of Experiment 3 is that action prediction was informed by the closest barriers to the current position of the ball, which in this case would be the presence or absence of the bottom barrier.

However, in Experiment 2, there were effects of both top and bottom barriers, even though the top barriers were not immediately in the bottom agent's path. While action prediction might not have been informed by the total length of the path, it was influenced in equal measure by immediate obstacles of the bottom agent and potential (but from their perspective immediate) future obstacles of the top agent. In other words, action prediction did not reflect joint efficiency operationalized as total path length across two agents. Using the same operational definition, prediction did not reflect individual efficiency either – a puzzling finding given a substantial amount of research suggesting that it should (Hudson et al., 2017; McDonough et al., 2018; 2019; 2020). Across Experiments 2 and 3, it seems that only the nearest barrier for the agent was informative for action prediction, except that in Experiment 2, the nearest barrier for both agents informed prediction of actions of the bottom player. The difference between the findings of Experiments 1 and 2 is probably due to instructions and catch trials in Experiment 2 made the efficiency for the top player more salient. However, the same catch trials and instructions were used in Experiment 3, with the only difference being that in Experiment 2 there were two agents. This shows that action prediction was equally informed by the efficiency of both players, if not at a sufficiently fine-grained level to be classified as joint efficiency in Experiment 2.

A possibility that should not be discounted is that the starting positions of the agents influenced how the task was represented. The scenario was seen from a bird's-eye view, with the bottom agent being closer to the participant. This might have caused participants to adopt the perspective of the bottom agent, especially as the task was for them to predict the bottom agent's actions. If this were the case, results of Experiment 2 might have been due to participants representing the scene as if they were the bottom player. In this scenario, they would have represented their "own", bottom, barriers, and the barriers of the top agent as individual constraints of each of the agents. Previous research suggests that when we engage in a task with another person, we represent their visuo-spatial perspective as well as physical constraints in their path even when they are irrelevant for our own part of the task (Freundlieb et al. 2016; 2017; Schmitz et al., 2017). Perspective-based interference or facilitation is not limited to real partners. In a study by Ward and colleagues (2018), participants were tasked to report whether letters in different orientations were canonical or mirrored. In some conditions, an observer was added to the scene. When people judged individually, the further the letters were orientated away from them the longer their reaction times. In the presence of another person, however, this effect was reduced. Thus, letters were recognized more quickly the closer to the upright angle they were either to the participant or the observer, suggesting that the other's visuo-spatial perspective was (quasi-) perceptually represented.

There are a few ways in which follow-up experiments could be modified to account for this. Perhaps the simplest option would be to rotate the layout by 90 degrees so that neither the top nor the bottom agent appear closer to the participant. Alternatively, agents could take turns so that in some trials the bottom agent would initiate the interaction while in other trials the top agent would do so. This particular modification would also be beneficial for controlling for the overall rightward bias across the experiments: it should be the case that in trials initiated by the bottom agent the overall bias would be towards the right while for the ones initiated by the top agent it should be to the left. Furthermore, such a scenario might make it more salient that the whole path needs to be traversed, regardless of the initiating agent. A more significant modification could tackle the abstractness of the whole set-up. Action prediction might not extend to abstract scenarios where interactions are observed virtually (and in a birds-eye view) in the same way that it operates when watching everyday actions. Using the same measure of representational momentum, more realistic displays of joint interactions could be designed. Such designs would allow for less instructions than were used in Experiments 1-3 which would create options for more fine-grained manipulations of expectations. For example, a simpler layout would allow for manipulating beliefs about agents' characteristics, allowing for comparisons of the extent to which action prediction is informed by perceived effort, path length, altruism, or ability. Another modification would be to conduct the experiment live which would allow to control for the size of the screen so that the pixel-size of the bias would relate to the same physical size across participants. A minor addition could be to use a touch screen instead of a mouse response. An online replication of a study using representational momentum found smaller effect sizes compared to conducting the study live on a touch screen (McDonough et al., 2019 cf. McDonough & Bach, 2022).

In summary, the present study has shown that explicit expectations are in line with joint efficiency. At the same time, predictions of ongoing actions were informed by the closest barrier to the acting agent (Experiments 1, 2 and 3) but not the overall path length, either joint or individual. This is likely due to the salience of the bottom barrier and complexity of the scene. Most importantly, Experiment 2 showed that predicting the actions of the bottom agent was informed by both the bottom and top barriers. Therefore, these findings show that predicting ongoing actions in an interaction can be informed by considerations of efficiency of both agents, if not by fully optimal paths for one or both of the agents in the interaction.

5 Chapter 5: General discussion

The aim of this thesis was to investigate how people represent their own and their partner's effort, how they distribute mental effort and whether the effort involved in an interaction can guide action prediction. The underlying assumption behind the studies was that organisms act as rational agents when trying to accomplish their goals (Dennett, 1988), which is extensively supported by research with humans (Anderson & Pandy, 2001; Cos et al., 2011; 2014; Todorov & Jordan, 2002; Todorov, 2004; Kool et al., 2010) as well as non-human animals (Walton et al., 2006; Bergman et al., 2001; Hughes et al., 1992; Rochat et al., 2008). In comparison to other species, humans are uniquely cooperative. Our cognitive and motivational processes are geared towards sharing intentions and collaborating with others which enables a wide range of activities ranging from everyday interactions to creation of complex cultural artefacts (Tomasello, 2005; 2009). This makes us adept at considering others' mental states, tasks, perspectives, physical abilities and predicting what they will do (Csibra et al., 1999; Gergely et al., 1995; Freundlieb et al., 2016; Rochat, 1995; Stoffregen et al., 1999). In collaborative contexts, we minimize the effort of our partners, be it in motor tasks (Constable et al., 2016; Gonzalez et al., 2011; Meyer et al., 2013) or during communication (Duran et al., 2011; Galati et al. 2013; Mainwaring et al., 2009; Schober 2009).

The first question that this thesis aimed to address was whether agents represent a task partner's effort and if this representation can influence their perception of the environment. The second aim was to test whether people, when sharing mental effort, do so in a way that conforms to principles of joint rationality. The third question investigated in this thesis was whether observers expect others to act coefficiently and if this informs people's action prediction.

5.1 SUMMARY OF KEY RESULTS

5.1.1 Chapter 2

Chapter two investigated whether effort modulates perceptual estimates of the environment when task demands are controlled for with appropriate cover stories. Our initial aim was to test if people engaged in a collaborative task represent their partner's effort and if this influences their estimates of height and distance. To establish that this effect can be reliably elicited with the intended manipulation, four experiments were conducted with individual participants. The first experiment tasked participants to lift items of varying weights off shelves of varying heights. The prediction was that estimates of the height of the shelf would be influenced by the weight of the item, with heavier items inducing overestimation of the height of the shelf. Furthermore, given that it is more difficult to lift items from higher shelves (regardless of the item's weight), we hypothesized that such overestimation would increase with shelf height, as was the case in previous studies with slant and distance (Bhalla & Proffitt, 1999; Proffitt et al., 1995). As findings of effort influencing perception were previously criticized as arising from task demands (Durgin et al., 2012; Firestone, 2013; Firestone and Scholl, 2014), cover stories were employed. The cover stories were designed to be as unobtrusive as possible, given that previous studies were criticized for employing conspicuous cover stories which might have eliminated the effect effort might have on perceptual judgments (Proffit, 2013; Witt, 2011; Witt et al., 2016). Debriefing confirmed that the deception was successful and none of the participants correctly guessed the purpose of the experiment. Going against predictions based on the literature reporting effects of effort on perception, our results showed that the weight of the object did not influence estimates of its height. A follow-up study tasked participants to judge distances which they would have to traverse either walking (low-effort condition) or hopping on one leg (high-effort condition). As in the previous experiment, a cover story was successfully employed, and estimates of distances were not modulated by the effort necessary to traverse them. The third experiment introduced two changes: effort was manipulated by carrying a heavy or empty backpack and participants did not traverse the distance after giving their estimates. This setup was almost identical to the original study by Proffitt and colleagues (2003) with the only difference being that it employed a cover story. As in Experiments 1 and 2, the effort manipulation did not result in different estimates of distance between the high and low effort conditions. The fourth experiment was identical to the previous one with the only difference being that no cover story was employed. The reasoning for dropping the cover story was that effects of effort on perception were previously criticized as arising from participants guessing the purpose of the manipulation (Durgin et al., 2009; 2012; Firestone, 2013; Firestone and Scholl, 2014). If the effects are due to construal rather than effort modulating perception, repeating the experiment without a cover story should bring it about. Results showed that the effort manipulation had an impact on distance estimates, with distances being judged longer in the high-effort than the low-effort condition. The debrief confirmed that all the participants correctly guessed the purpose of the effort manipulation. Since the only difference between Experiments 3 and 4 was participants' belief about the manipulation, findings of Experiment 4 are likely due to task demands rather than effort influencing perception. These results did not open the way to testing if corepresentation of effort in a joint task can influence both agents' perceptual estimates. However, taken together, Experiments 1-4 extend our knowledge on the mechanisms through which effort influences perceptual estimates (Proffit, 2013; Witt, 2011; Witt et al., 2016), In particular, our findings underline the importance of task framing and the need for employing successful cover stories to rule out task demands. One potential follow-up experiment would be to replicate the cover story manipulation using height estimates as was the case in Experiment 1. This could be broadened to size and slant perception with employing or not employing a cover story being a between-subject factor. These modifications would allow for testing the generality of task construal influencing perceptual estimates. In sum, results from experiments in Chapter 2 provide insights about methodology, urging caution when framing instructions and using cover stories. An important implication of Chapter 2 is that effects of effort on perception should be checked against cover stories successfully masking the intended manipulation from subjects. It remains an open question whether these findings would generalize to other areas such as height and slant perception.

5.1.2 Chapter 3

The aim of Chapter 3 was to investigate how people engaged in a cooperative task distribute associated cognitive costs. In particular, the main question was whether they employ a strategy that minimizes their own, their partner's, or the total amount of mental effort. A further question was whether this strategy was informed by their performance, assessed self-competence, and assessed competence of their partner. Six experiments eliciting mental effort were conducted online using a multiple-object tracking paradigm (Experiments 1-4) and two experiments were done live using a working memory task (Experiments 5 and 6). The first experiment aimed to confirm that adding additional targets increases the difficulty of the task (reflected in decreased accuracy) and that participants were aware that this is the case (reflected in participants' ratings of how difficult a presented task would be). Experiment 1 confirmed that the difficulty manipulation worked as intended and that participants could tell the different difficulties apart. Experiment 2 aimed to confirm that people prefer easy over difficult tracking trials. In the first part of the experiment, participants completed multiple-object tracking trials. In the second part, they were presented with hypothetical tracking trials and had to rate the extent to which they would like to do them in the (supposed) third part of the experiment. Participants' ratings confirmed that they preferred to do easier trials over more difficult ones. Taken together, Experiments 1 and 2 showed that the paradigm worked as intended both in terms of difficulty manipulation and participants' preferences. The third experiment was a "joint" one – participants believed that they were going to be doing the same task as in the previous experiments but this time with a partner. The task was the same except that they were informed that success on a trial depends on both their own and their partner's accuracy. They believed that the partner was seeing the same MOT task, except that they were tracking their own set of targets. After the first tracking block, participants were told that they will be rating a series of tracking tasks with different difficulty distributions for themselves and the partner. Based on their ratings, they and their partner will do some of these tasks in the third part of the experiment. Results showed a predominately egoistic strategy, with minimization of own effort and no attempt to minimize the partner's. This was true regardless of self-assessed competence on the task. Better performers showed a preference for partners' easy over medium and difficult trials. However, this was only the case when the participants' part of the trial was easy as well. The subsequent experiment had a slight modification; to ensure that egoistic strategies were not dominant because people were assuming their partner was also giving ratings, it was made clear during the instructions that only the participant's ratings will influence the selection of subsequent trials. In this experiment, a similar pattern of self-effort minimization emerged but the partner's difficulty was taken more into account. When the participant had an easy part of the trial, they preferred that the partner had an easy over a medium over a difficult trial. In participants' trials of medium difficulty, they only differentiated between partner's easy and medium trials on the one hand and difficult ones on the other. Finally, on trials where the participant had the difficult task, there was no preference for any of the three difficulties of the partner. This might suggest that distributing the task co-efficiently, or even just taking the partner into account comes with an additional cognitive cost. As their own difficulty increases, participants' resolution of the partner's difficulty (and consequentially of the whole task) becomes more difficult to track. Therefore, in self-easy trials, they might be willing to disambiguate different difficulties for the partner but as their part of the task becomes more difficult, computing the partner's cost introduces additional effort. The final two experiments used the same structure but with a working memory task where difficulty was manipulated by the number of items participants and their supposed partners would have to recall. The aim was to replicate the previous findings using a different task. The final difference between these experiments and Experiments

1-4 was that they were done live. Experiment 5 largely replicated the findings of the previous experiments with one key difference. Ratings reflected a strategy where all difficult trials, regardless of whether they were the participants' or the partners were rated equally low. In other words, scenarios where the participant would have an easy part of the trial and their partner the difficult one were judged as equally desirable if the difficulties were reversed. This is a marked departure from results of Experiments 1-4 and might be due to the fact that the experiment was conducted live, and participants met before the experiment began. Experiment 6 aimed to test if effort distribution could be influenced by showing to participants that their partner was less competent at the task than they were. The procedure was identical to Experiment 5 with one key modification: periodically, feedback was given in a joint format (“You and your partner solved X trials correctly”). The feedback for the supposed partner's accuracy was calculated based on performance in Experiment 5. More precisely, the supposed partner's accuracy was 2 standard deviations lower than the average performance of participants in Experiment 5, creating an impression of an incompetent partner. While over half of the participants reported that they thought the partner was worse at the task than they were, the distribution of effort followed the same pattern as in Experiment 5.

5.1.3 Chapter 4

Chapter 4 (Study 3) tested whether observers of an interaction expect agents to minimize joint effort and if these expectations inform action prediction. These experiments built on previous findings that expectations of individual efficiency can be tested using representational momentum – remembering moving object's locations closer to expected trajectories than they were in reality (Hudson et al., 2017; McDonough et al., 2018; 2019; 2020). Three experiments were conducted in which participants viewed an interactive game where the agents had to lead a ball from the bottom to the top of the screen while avoiding barriers. Effort or costs were operationalized as path length. In Experiments 1 and 2, the bottom agent would lead the ball to one of two central locations whereupon the top agent would take over the ball and lead it to the final location. The barriers were varied so that some paths were individually (but not jointly) efficient while others were individually inefficient but jointly efficient. The first experiment showed that explicit judgments of expectations were in line with joint rationality: participants expected the bottom agent to choose paths that reduced the overall length of the path. However, this was not reflected in the measure of representational momentum, which only showed an effect of the barrier closest to the bottom agent but not their partner. The second experiment followed a similar procedure except that it introduced catch trials which prompted studying the whole layout rather than just the bottom part of the screen. As was the case in Experiment 2, explicit expectations were in line with co-efficiency. With regard to action prediction, both the top and bottom barriers informed predicted trajectories. However, these predictions, while taking into account both agents, were not biased to the path that would be overall the shortest. More precisely, predictions were biased away from barriers of the top and bottom agents separately, while the most efficient path was the result of different barrier combinations for both agents. The third experiment aimed to test if the results from Experiments 1 and 2 were due to expectations of co-efficiency not being informative for action prediction or were a reflection of a more general bias against immediate obstacles for the bottom agent. The procedure was modified so that the entire path was traversed by only one agent. Only the closest barriers informed action prediction, suggesting that results of Experiment 2 reflected expectations that the bottom agent would take both their own efficiency and their partner's into account, if not the overall length of the path.

5.2 Conclusions and open questions

5.3 Effort and perceptual judgments

The first study (Chapter 2) showed that effort does not influence perceptual judgments of the environment if task construal is controlled for. The initial goal of experiments in Chapter 2 was to establish a working paradigm for testing of corepresentation of effort and whether it would influence perception. While findings of task construal being the driver behind effort influencing perception precluded testing the second part of the question, corepresentation of effort could be investigated using different measures. Potts and colleagues (2017) tested people's preferences for various cognitive and physical tasks and found that choices on which ones to do were formed by the subjective time needed to do the task. Subjective time could be instead used as a measure of corepresentation in a cooperative setting: if the task difficulty of one's partner is represented, it should have an effect on estimates on how long it would take to complete a task. One possibility would be a joint memory task where the difficulty of the participants' part of the task and the partner's time to complete their part of task are kept constant but the difficulty of the partner's part of the task was varied. Participants' estimates of the amount of time needed to complete tasks could indicate corepresentation of their partner's effort: the more difficult the partner's task, the longer it would be judged to take to complete.

5.4 Distribution of mental effort

The second study (Chapter 3) showed that when it comes to cognitive costs, effort distribution does not seem to follow the same pattern as it does with motor or physical costs. Specifically, when sharing physical costs, people place equal weight on their own and their partner's effort (Török et al., 2022) whereas in experiments in Chapter 3, own mental effort was prioritized while the partner's mental effort was minimized inconsistently. This could be due to mental effort being generally aversive and laborious for people. Supporting this, Vogel and colleagues (2020) had people perform an N-back task in which they were sequentially presented with a series of items and they had to decide if the current item matches an item that was presented a number (N) of trials before. In some trials, participants had a choice between doing the task or receiving a painful thermal stimulus. As the difficulty of the N-task increased, people were increasingly more likely to opt for the thermal stimulus instead. This speaks to mental effort being unpleasant and, possibly, one would be less inclined to act coefficiently while engaging in a cognitive task compared to a motor or physical one. While there is evidence of people sharing mental effort in studies on communication (Duran et al., 2011; Galati et al. 2013; Mainwaring et al., 2009; Schober 2009), it is less clear whether this behavior reflects coefficiency in the sense that it leads to increasing own's effort to the extent that this increase would decrease the effort of their partner.

A connected question is that of the overall difficulty of the task. Previous studies showing coefficiency or approximate coefficiency used easier tasks such as moving items on a touch screen (Török et al., 2019; 2021) or holding a door open (Santamaria & Rosenbaum, 2011). Experiments in Chapter 3 showed that, in low effort conditions, people are good at detecting and assisting with another's effort. As the difficulty increases, either the resolution for the partner's effort or the wiliness to reduce it decreases. A related issue is that we based the difficulty manipulation on the assumption that each additional item to track or remember increases the difficulty by the same amount. This might not have been the case. For example, in multiple-object tracking tasks, performance is affected by a number of dimensions such as the number of targets, distractors, their similarity, speed and trial duration (Alvarez et al., 2007; Bettencourt & Sommers, 2009; Holcombe & Chen, 2012). A series of experiments could keep the decider's effort comparably low while varying the partner's level of effort to find a low-effort scenario of the decider where they are sensitive to the partner's costs. Once this baseline is established, subsequent work could increase the decider's costs until their resolution for the partner's effort starts decreasing. This would allow for testing of boundary conditions of coefficiency for sharing of cognitive costs.

Another possibility is that the rating task, rather than the memory or tracking task, was too difficult. In the current setup, to reach a coefficient distribution of effort, participants would have to keep in mind, across trials, the three levels of their own difficulty combined with the three levels of their partner's difficulty. Furthermore, the slider measure used to indicate preference would need to be sufficiently sensitive to reflect these computations. It is possible that taking all these considerations into account introduced a level of difficulty during rating trials that people were either unable or unwilling to engage in. This could also explain the lack of minimization of the partner's effort; participants focused on making their task easier but weighing the different difficulties of their partner's task (and combinations of their own difficulties) made judgments of preferred future trials too complex. In line with this, in a study on social loafing involving cognitive costs, Weldon and Gargano (1988) showed that when people are not held accountable or do not feel responsible for the outcomes of a task, they put in less effort. In the reported experiments on mental effort, both accountability and responsibility are, at least to an extent, ambiguous. Experiments 3 and 4 were done online while Experiments 5 and 6 were done live but participants sat alone in the room. Both scenarios minimize accountability as participants did not know each other beforehand. As for responsibility, participants were informed that their judgments will affect the choice of trials in the supposed third part of the experiment but, apart from higher ratings increasing the probability of the trial occurring later, no clear relationship between ratings and trial probabilities was established. This opens a wide range of possibilities for participants to interpret how (and whether) their ratings influence future trials, leaving the interpretation of responsibility open to each participant. A way forward would be to both make it easier to disambiguate individually and jointly efficient choices as well as making consequences of choices less hypothetical. A variation of the experiment could be to have participants draw cards from three decks; a deck that minimizes the participants' effort, one that minimizes the partner's effort or a third deck that minimizes the overall effort but there is a 50-50 chance that it would increase or decrease the choosing participant's effort compared to the individually efficient deck. Once a card is drawn, the task would be immediately presented. This kind of setup would solve the problem of accountability (as the choosing participant is obviously accountable) and would minimize the effort associated with computing individual and joint costs across many different self-other difficulty combinations.

5.5 Expectations of co-efficiency and action prediction

The third study (Chapter 4) demonstrated that joint efficiency guides action prediction. While prediction was not fully sensitive to the total path length of both agents in the interaction, this was also the case for individual agents. Potential follow-ups should concentrate on making the scene simpler and to have constraints of both agents more readily available to perceivers. Additionally, research on individual efficiency demonstrated that high-level information about agents and their intentions can influence predictions (Hudson et al., 2016). As previous research on joint rationality mostly concentrated on emergent behavior rather than actively manipulating high-level expectations, an interesting avenue would be to see if people's knowledge about the agents involved in the interaction (e.g., "The bottom agent is interested in doing their part as easily as possible") affects their judgments. Chapter 4 tested coefficiency expectations when observing an interaction. A further step could be to test if people involved in the interaction have the same expectations and whether they inform action prediction. If so, expectations of coefficiency could prove to facilitate coordination and monitoring in a wide range of everyday interactions.

5.6 Conclusions

This thesis shows that perception and effort are linked (Chapter 4) but that this connection should be taken with careful consideration (Chapter 2) and that sharing of cognitive costs does not follow typically jointly rational strategies as sharing other kinds of effort (Chapter 3). In other areas of research far removed from perception, such as moral reasoning, varying irrelevant dimensions of a task can produce significant changes in the results (Petrinovich and O'Neill, 1996; Cao et al., 2017). The second chapter showed that perceptual estimates can be influenced by how the situation is construed, suggesting that contextual factors should be taken into account even when investigating potentially impenetrable cognitive mechanisms. The experiments on action prediction demonstrated that observers expect that agents will minimize each other's effort and that this expectation guides action prediction. Previous research on joint efficiency mostly concentrated on studying whether people make coefficient decisions. The results of the studies reported in Chapter 4 raise the possibility that ongoing interactions are informed by expectations of coefficiency. If this is the case, coefficiency could be an important aspect of successful coordination in a range of interactions, ranging from simple motor tasks to more complex planning and decision-making. Finally, the results show that mental effort distribution does not follow the same pattern as is the case with physical or motor effort. This might be because it is more aversive or that coefficient distribution of mental effort requires more forethought than is the case in the motor domain. This opens avenues for future research on the calculus of joint cognitive costs.

6 References

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., and Kievit, R. A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res. 4:63. doi: 10.12688/wellcomeopenres.15191.2

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of vision*, *7*(13), 14-14.

Anderson, F. C., & Pandy, M. G. (2001). Dynamic optimization of human walking. *J. Biomech. Eng.*, *123*(5), 381-390.

Anderson, S. J., Yamagishi, N., & Karavia, V. (2002). Attentional processes link perception and action. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1497), 1225-1232.

Anstis, S. (1995). Aftereffects from jogging. Experimental Brain Research, 103(3), 476- 478.

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Experimental brain research*, 211(3-4), 371-385.

Balcetis, E., and Dunning, D. (2010). Wishful seeing: more desired objects are seen as closer. Psychol. Sci. 21, 147–152. doi: 10.1177/0956797609356283

Banerjee, P., Chatterjee, P., and Sinha, J. (2012). Is it light or dark? Recalling moral behavior changes perception of brightness. Psychol. Sci. 23, 407–409. doi: 10.1177/0956797611432497

Bergman, C. M., Fryxell, J. M., Gates, C. C., & Fortin, D. (2001). Ungulate foraging strategies: energy maximizing or time minimizing?. *Journal of Animal Ecology*, *70*(2), 289-300.

Bettencourt, K. C., & Somers, D. C. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of vision*, *9*(7), 9-9.

Bhalla, M., & Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *Journal of experimental psychology: Human perception and performance*, 25(4), 1076.

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Giving a helping hand: effects of joint attention on mental rotation of body parts. *Experimental brain research*, *211*, 531-545.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European journal of neuroscience*, 13(2), 400-404.

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009). Broken affordances, broken objects: a TMS study. *Neuropsychologia*, 47(14), 3074-3078.

Cao, F., Zhang, J., Song, L., Wang, S., Miao, D., & Peng, J. (2017). Framing effect in the trolley problem and footbridge dilemma: Number of saved lives matters. Psychological reports, 120(1), 88-101.

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2011). The space of affordances: A TMS study. *Neuropsychologia*, 49(5), 1369-1372.

Chang, C. H., Wade, M. G., & Stoffregen, T. A. (2009). Perceiving affordances for aperture passage in an environment–person–person system. *Journal of Motor Behavior*, 41(6), 495-500.

Cho, D. T., & Proctor, R. W. (2010). The object-based Simon effect: Grasping affordance or relative location of the graspable part?. *Journal of Experimental Psychology: Human Perception and Performance*, 36(4), 853.

Cho, D. T., & Proctor, R. W. (2011). Correspondence effects for objects with opposing left and right protrusions. *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 737.

Clore, G. L., and Proffitt, D. R. (2016). The myth of pure perception. Behav. Brain

Cos, I., Bélanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of neurophysiology*, *105*(6), 3022-3033.

Cos, I., Duque, J., & Cisek, P. (2014). Rapid prediction of biomechanical costs during action decisions. *Journal of Neurophysiology*, *112*(6), 1256-1266.

Craighero, L., Fadiga, L., Umiltà, C. A., & Rizzolatti, G. (1996). Evidence for visuomotor priming effect. *Neuroreport*, 8(1), 347-349. 21

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica*, *124*(1), 60-78.

Doerrfeld, A., Sebanz, N., & Shiffrar, M. (2012). Expecting to lift a box together makes the load look lighter. *Psychological Research*, 76(4), 467-475.

Dötsch, D., & Schubö, A. (2015). Social categorization and cooperation in motor joint action: evidence for a joint end-state comfort. *Experimental brain research*, *233*(8), 2323-2334.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., and Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. Psychon. Bull. Rev. 16, 964–969. doi: 10.3758/pbr.16.5.964

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J., & Williams, M. (2012). The social psychology of perception experiments: Hills, backpacks, glucose, and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1582.

Durgin, F. H., Pelah, A., Fox, L. F., Lewis, J., Kane, R., & Walley, K. A. (2005). Selfmotion perception during locomotor recalibration: more than meets the eye. *Journal of Experimental Psychology: Human Perception and Performance*, 31(3), 398.

Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V., & Pavesi, G. (1998). Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. *Neuropsychologia*, 37(2), 147-158.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, 73(6), 2608-2611.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, 17(8), 1703-1714.

Firestone, C. (2013). How “paternalistic” is spatial perception? Why wearing a heavy backpack doesn’t—and couldn’t—make hills look steeper. Perspect. Psychol. Sci. 8, 455–473. doi:10.1177/1745691613489835

Firestone, C., & Scholl, B. J. (2014). “Top-down” effects where none should be found: The El Greco fallacy in perception research. *Psychological science*, 25(1), 38-46.

Firestone, C., and Scholl, B. J. (2016). Cognition does not affect perception: evaluating the evidence for “top-down” effects. Behav. Brain Sci. 39:e229.

Frak, V., Paulignan, Y., & Jeannerod, M. (2001). Orientation of the opposition axis in mentally simulated grasping. *Experimental Brain Research*, 136(1), 120-127.

Franchak, J. M., Celano, E. C., & Adolph, K. E. (2012). Perception of passage through openings depends on the size of the body in motion. *Experimental brain research*, *223*, 301-310.

Freundlieb, M., Kovács, Á. M., & Sebanz, N. (2016). When do humans spontaneously adopt another’s visuospatial perspective?. *Journal of experimental psychology: human perception and performance*, *42*(3), 401.

Freundlieb, M., Sebanz, N., & Kovács, Á. M. (2017). Out of your sight, out of my mind: Knowledge about another person’s visual access modulates spontaneous visuospatial perspective-taking. *Journal of experimental psychology: human perception and performance*, *43*(6), 1065.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126.

Galetzka, C. (2017). The story so far: how embodied cognition advances our understanding of meaning-making. Front. Psychol. 8:1315. doi: 10.3389/fpsyg. 2017.01315

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in cognitive sciences*, *7*(7), 287-292.

Gibson, J. J. (1979). The Ecological Approach to Visual Perception. Boston, MA: Houghton Mifflin.

Glover, S., & Dixon, P. (2017). The role of predictability in cooperative and competitive joint action. *Journal of Experimental Psychology: Human Perception and Performance*, 43(4), 644

Gonzalez, D. A., Studenka, B. E., Glazebrook, C. M., & Lyons, J. L. (2011). Extending end-state comfort effect: Do we consider the beginning state comfort of another?. *Acta psychologica*, *136*(3), 347-353.

Halpern, A. R., & Kelly, M. H. (1993). Memory biases in left versus right implied motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 471.

Hoffman, D. D., Singh, M., and Prakash, C. (2015). The interface theory of perception. Psychon. Bull. Rev. 22, 1480–1506. doi: 10.3758/s13423-015-0890-8

Holcombe, A. O., & Chen, W. Y. (2012). Exhausting attentional tracking resources with a single fast-moving object. *Cognition*, *123*(2), 218-228.

Hubbard, T. L. (1990). Cognitive representation of linear motion: Possible direction and gravity effects in judged displacement. *Memory & Cognition*, *18*, 299-309.

Hubbard, T. L. (1995). Cognitive representation of motion: Evidence for representational friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 1-14.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic bulletin & review*, *12*, 822-851.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*, 211-221.

Hubbard, T. L., & Favretto, A. (2003). Naïve impetus and Michotte's " tool effect": Evidence from representational momentum. *Psychological Research*, *67*(2), 134-152.

Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018). Perceptual teleology: expectations of action efficiency bias social perception. *Proceedings of the Royal Society B*, *285*(1884), 20180638

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition*, *146*, 245-250.

Hunt, D. F., Hunt, H. B., and Park, J. H. (2017). Bioenergetic costs and state influence distance perception. Physiol. Behav. 180, 103–106. doi: 10.1016/j. physbeh.2017.08.003

Isenhower, R. W., Richardson, M. J., Carello, C., Baron, R. M., & Marsh, K. L. (2010). Affording cooperation: Embodied constraints, dynamics, and action-scaled invariance in joint lifting. *Psychonomic Bulletin & Review*, *17*(3), 342-347.

Janyan, A., & Slavcheva, G. V. (2012). When left feels right: asymmetry in the affordance effect. *Cognitive processing*, 13(1), 199-202.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14-23.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, 14(1), S103-S109.

Josa, R. V., Camus, T. P., Murday, V., Morgado, N., Palluel Germain, R., Brunel, L., et al. (2019). The action constraints of an object increase distance estimation in extrapersonal space. Front. Psychol. 10:472. doi: 10.3389/fpsyg.2019.00472

Kerzel, D. (2000). Eye movements and visible persistence explain the mislocalization of the final position of a moving target. *Vision Research*, *40*(27), 3703-3715.

Kerzel, D. (2003). Attention maintains mental extrapolation of target position: Irrelevant distractors eliminate forward displacement after implied motion. *Cognition*, *88*(1), 109-131.

Kerzel, D., Jordan, J. S., & Müsseler, J. (2001). The role of perception in the mislocalization of the final position of a moving target. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 829.

Kozhevnikov, M., & Hegarty, M. (2001). Impetus beliefs as default heuristics: Dissociation between explicit and implicit knowledge about motion. *Psychonomic Bulletin & Review*, *8*, 439 453.

Krogh-Jespersen, S., Henderson, A. M., & Woodward, A. L. (2020). Let's get it together: Infants generate visual predictions based on collaborative goals. *Infant Behavior and Development*, *59*, 101446.

LeBoeuf, R. A., and Shafir, E. (2003). Deep thoughts and shallow frames: on the susceptibility to framing effects. J. Behav. Decis. Mak. 16, 77–92. doi: 10.1002/ bdm.433

Lee, Y., Lee, S., Carello, C., and Turvey, M. T. (2012). An archer's perceived form scales the "hitableness" of archery targets. J. Exp. Psychol. Hum. Percept. Perform. 38, 1125–1131. doi:10.1037/a0029036

Lessard, D. A., Linkenauger, S. A., & Proffitt, D. R. (2009). Look before you leap: Jumping ability affects distance perception. *Perception*, 38(12), 1863-1866.

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, *160*, 35-42.

Mark, L. S. (2007). Perceiving the actions of other people. *Ecological Psychology*, 19(2), 107-136.

Mascaro, O., & Csibra, G. (2022). Infants expect agents to minimize the collective cost of collaborative actions. *Scientific Reports*, *12*(1), 17088.

McDonough, K. L., & Bach, P. (2023). Expectations of efficient actions bias social perception: a pre-registered online replication. *Royal Society Open Science*, *10*(2), 220889.

McDonough, K. L., Hudson, M., & Bach, P. (2019). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports*, *9*(1), 1-10

Meagher, B. R., & Marsh, K. L. (2014). The costs of cooperation: Action-specific perception in the context of joint action. *Journal of experimental psychology: human perception and performance*, 40(1), 429.

Meyer, M., Bekkering, H., Paulus, M., & Hunnius, S. (2010). Joint action coordination in 2½-and 3-year-old children. *Frontiers in Human Neuroscience*, 4, 220.

Meyer, M., Van Der Wel, R. P., & Hunnius, S. (2013). Higher-order action planning for individual and joint object manipulations. *Experimental brain research*, *225*, 579-588.

Molto, L., Morgado, N., Guinet, E., Fazioli, L., Heurley, L. P., and PalluelGermain, R. (2020). Motor simulation in tool-use effect on distance estimation: a replication of Witt and Proffitt (2008). Psychon. Bull. Rev. 27, 301–306. doi: 10.3758/s13423-019-01686-7

O'Shea, H., & Moran, A. (2017). Does motor simulation theory explain the cognitive mechanisms underlying motor imagery? A critical review. *Frontiers in human neuroscience*, 11, 72.

Petrinovich, L., and O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. Ethol. Sociobiol. 17, 145–171. doi: 10.1016/0162-3095(96) 00041-6

Pezzulo, G., Iodice, P., Donnarumma, F., Dindo, H., & Knoblich, G. (2017). Avoiding accidents at the champagne reception: A study of joint lifting and balancing. *Psychological science*, *28*(3), 338-345.

Philbeck, J. W., Woods, A. J., Arthur, J., & Todd, J. (2008). Progressive locomotor recalibration during blind walking. *Perception & psychophysics*, 70(8), 1459-1470.

Proffit, D. R. (2013). An embodied approach to perception: by what units are visual perceptions scaled? Perspect. Psychol. Sci. 8, 474–483. doi: 10.1177/1745691613489837

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on psychological science,* 1(2), 110-122. doi: 10.1111/j.1745-6916.2006.00008.x

Proffitt, D. R., and Linkenauger, S. A. (2013). "Perception viewed as a phenotypic expression," in Action Science: Foundations of an Emerging Discipline, eds W. Prinz, M. Beisert, and A.

Herwig (Cambridge, MA: The MIT Press), 171. doi: 10.7551/mitpress/9780262018555.003.0007

Proffitt, D. R., Bhalla, M., Gossweiler, R., and Midgett, J. (1995). Perceiving geographical slant. Psychon. Bull. Rev. 2, 409–428. doi: 10.3758/BF03210980

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science*, 14(2), 106-112.

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 839.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human movement science*, 26(6), 867-891. 23

Rochat, M. J., Caruana, F., Jezzini, A., Intskirveli, I., Grammont, F., Gallese, V., ... & Umiltà, M. A. (2010). Responses of mirror neurons in area F5 to hand and tool grasping observation. *Experimental brain research*, 204(4), 605-616.

Rochat, P. (1995). Perceived reachability for self and for others by 3-to 5-year-old children and adults. *Journal of Experimental Child Psychology*, 59(2), 317-333.

Roepstorff, A., and Frith, C. (2004). What's at the top in the top-down control of action? Script-sharing and 'top-top'control of action in cognitive experiments. Psychol. Res. 68, 189–198. doi: 10.1007/s00426-003-0155-4

Rosenbaum, D. A., Vaughan, J., Barnes, H. J., & Jorgensen, M. J. (1992). Time course of movement planning: selection of handgrips for object manipulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1058.

Rosenbaum, D. A., Marchk, F., Barnes, H. J., Vaughan, J., Slotta, J. D., & Jorgensen, M. J. (2018). Constraints for action selection: Overhand versus underhand grips. In *Attention and performance XIII* (pp. 321-342). Psychology Press.

Satta, E., Ferrari-Toniolo, S., Visco-Comandini, F., Caminiti, R., & Battaglia-Mayer, A. (2017). Development of motor coordination during joint action in midchildhood. *Neuropsychologia*, 105, 111-122.

Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. Ecological Psychology, 9(3), 189-206.

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. Journal of Experimental Psychology: *Human Perception and Performance*, 43(8), 1480.

Schnall, S., Zadra, J. R., and Proffitt, D. R. (2010). Direct evidence for the economy of action: glucose and the perception of geographical slant. *Perception* 39, 464–482. doi: 10.1068/p6445

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, 10(2), 70-76.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own?. *Cognition*, 88(3)

Shaffer, D. M., McManama, E., Swank, C., and Durgin, F. H. (2013). Sugar and space? Not the case: effects of low blood glucose on slant estimation are mediated by beliefs. Iperception 4, 147–155. doi: 10.1068/i0592

Shapiro, L. (2011). Embodied Cognition. New York, NY: Routledge.

Stenzel, A., Chinellato, E., Bou, M. A. T., del Pobil, Á. P., Lappe, M., & Liepelt, R. (2012). When humanoid robots become human-like interaction partners: corepresentation of robotic actions. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1073.

Stoffregen, T. A., Gorday, K. M., Sheng, Y. Y., & Flynn, S. B. (1999). Perceiving affordances for another person's actions. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 120.

Symes, E., Ellis, R., & Tucker, M. (2005). Dissociating object-based and space-based affordances. Visual Cognition, 12(7), 1337-1361.

Taylor, J. E. T., Witt, J. K., and Sugovic, M. (2011). When walls are no longer barriers: perception of wall height in parkour. Perception 40, 757–760. doi: 10.1068/p6855

Thompson, M. E. (1944). Learning as a function of the absolute and relative amounts of work. *Journal of Experimental Psychology*, *34*(6), 506.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience*, *7*(9), 907-915.

Tomasello, M. (2009). *Why we cooperate*. MIT press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675-691.

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in joint action: MaximiFzing coefficiency in coordination. *Psychological science*, *30*(6), 930-941.

Török, G., Stanciu, O., Sebanz, N., & Csibra, G. (2021). Computing joint action costs: Co-actors minimize the aggregate individual costs in an action sequence. *Open Mind*, *5*, 100-112.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human perception and performance*, 24(3), 830. doi: 10.1037/0096-1523.24.3.830

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual cognition*, 8(6), 769-800. 24, doi: 10.1080/13506280042000144

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta psychologica*, 116(2), 185-203.

van der Wel, R. P., & Fu, E. (2015). Entrainment and task co-representation effects for discrete and continuous action sequences. *Psychonomic bulletin & review*, 22(6), 1685- 1691.

Vesper, C., van der Wel, R. P., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental brain research*, 211(3-4), 517-530.

Viger, C. (2006) „Is the aim of perception to provide accurate representations? A case for the 'no' side“, in: Stainton, R. J. (ed.) Contemporary Debates in Cognitive Science, Oxford: Blackwell Publishing, 275-288

Vinson, N. G., & Reed, C. L. (2002). Sources of object-specific effects in representational momentum. *Visual Cognition*, *9*(1-2), 41-65.

Vogel, T. A., Savelson, Z. M., Otto, A. R., & Roy, M. (2020). Forced choices reveal a trade-off between cognitive effort and physical pain. *elife*, *9*, e59410.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural networks*, *19*(8), 1302-1314.

Wang, X. T., and Johnston, V. S. (1995). Perceived social context and risk preference: a re-examination of framing effects in a life-death decision problem. J. Behav. Decis. Mak. 8, 279–293. doi: 10.1002/bdm.3960080405

Warren Jr, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3), 371.

Warren, W. H. (1984). Perceiving affordances: visual guidance of stair climbing. *Journal of experimental psychology: Human perception and performance*, 10(5), 683.

Weldon, E., & Gargano, G. M. (1988). Cognitive loafing: The effects of accountability and shared responsibility on cognitive effort. *Personality and Social Psychology Bulletin*, *14*(1), 159-171.

Wilson, M. (2002). Six views of embodied cognition. Psychon. Bull. Rev. 9, 625–636. doi: 10.3758/bf03196322

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., and Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. Proc. Natl. Acad. Sci. U.S.A. 104, 7780–7785. doi: 10.1073/pnas.0701644104

Witt, J. K. (2011). Action's effect on perception. Curr. Dir. Psychol. Sci. 20, 201–206. doi: 10.1177/0963721411408770

Witt, J. K., & Proffitt, D. R. (2008). Action-specific influences on distance perception: a role for motor simulation. *Journal of experimental psychology: Human perception and performance*, 34(6), 1479.

Witt, J. K., Linkenauger, S. A., & Wickens, C. (2016). Action-specific effects in perception and their potential applications. *Journal of applied research in memory and cognition*, 5(1), 69-76. doi:10.1016/j.jarmac.2015.07.008

Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: a role of effort and intent. Perception 33, 577–590. doi: 10.1068/p5090

Witt, J. K., Proffitt, D. R., and Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. J. Exp. Psychol. Hum. Percept. Perform. 31, 880–888. doi:10.1037/0096-1523.31.5.880

Witzel, C., Valkova, H., Hansen, T., and Gegenfurtner, K. R. (2011). Object knowledge modulates colour appearance. Iperception 2, 13–49. doi: 10.1068/i0396

Wolpert, D. M., & Landy, M. S. (2012). Motor control is decision-making. *Current opinion in neurobiology*, *22*(6), 996-1003.

Woods, A. J., Philbeck, J. W., and Danoff, J. V. (2009). The various perceptions of distance: an alternative view of how effort affects distance judgments. *J. Exp. Psychol. Hum. Percept*. Perform. 35, 1104–1117. doi: 10.1037/a0013622

Xiang, Y., Vélez, N., & Gershman, S. J. (2023). Collaborative decision making is grounded in representations of other people's competence and effort. *Journal of Experimental Psychology: General*.

Zadra, J. R., and Proffitt, D. R. (2016). Optic flow is calibrated to walking effort. *Psychon. Bull. Rev*. 23, 1491–1496. doi: 10.3758/s13423-016-1017-6

Zadra, J. R., Weltman, A. L., & Proffitt, D. R. (2016). Walkable distances are bioenergetically scaled. *Journal of experimental psychology: human perception and performance*, 42(1), 39

ACKNOWLEDGEMENT

TO EXTERNAL FUNDING AGENCIES CONTRIBUTING TO PHD DISSERTATIONS

Name of doctoral candidate: Vjeran Kerić

Title of dissertation: PERCEPTION AND DISTRIBUTION OF EFFORT IN JOINT ACTION

Name of supervisor(s): Natalie Sebanz, Gunther Knoblich

External funding agency: **European Research Council**

Acknowledgement:

This research has received funding from the European Research Council (ERC) under the European Union's Seventh Framwork Programma(FP7/2007-2013) under **grant agreement No. 616072- JAXPERTISE**