Chapter 7
The Problem of Interpretation in Experimental Research

Davood Gozli

When we design an experiment, we set out to compare a number of conditions with respect to a set of dependent variable(s). The design serves its intended purpose when the conditions are similar except with respect to our independent (manipulated) variables. A well-known error arises when an independent variable is conflated with an unintended change (a confound). In the present chapter, we are interested in a particular type of confound, namely, the meaning participants assign to events in the experiment. By raising the question, "What else could the events mean for the research participants?," we are raising another closely connected question, "What else can the findings mean?" If the meaning assigned to events changes across experimental conditions, we can no longer assume that independent variables are manipulated within otherwise "controlled" conditions. Recognizing how the meaning of events might have changed for the research participants in turn changes the meaning of the experimental findings, which could deflate or undermine both the rationale and the theoretical significance of the research.

Let us illustrate the main point with a simple example. Imagine that we are interested in the effect of the loudness of task instructions on participants' performance. We divide participants into two groups, one receives instruction in normal voice and the other receives instruction in a loud voice. Some of the participants in the "loud" condition might interpret the experimenter's loudness as impatience, rudeness, or negative mood. If that is the case, describing the two conditions in terms of loudness alone would be inadequate. Consequently, any observed difference in performance cannot be attributed to the loudness of instructions alone. To explain a difference between the two conditions, we will have to consider that a loudly delivered instruction might have a different meaning, compared with the same instruction delivered in a normal voice. Such differences in interpretation can produce differences in the perceived context of research participation (Bergner, 2010, 2016). A condition in which participants follow the instructions of a rude researcher represents a qualitatively different context, compared to one where the researcher behaves politely (Johns, 2006).

D. Gozli (✉)
University of Macau, Taipa, Macau S.A.R., China

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_7

To formulate the problem more generally, researchers might want to test whether there is a relation between two variables (e.g., loudness of the verbal instructions and the speed of task performance) while assuming no change in the kind of processes under investigation. However, differences in the participants' interpretation can undermine that assumption (Toomela, 2008; Valsiner & Brinkmann, 2016). There is no guarantee for participants' perspective to uniformly conform to what is assumed in the experimental design and the categories applied in advance to what happens during the experiment. Across different experimental conditions, events might differ partly in ways that are intended by the experimenters (e.g., loudness of voice) and partly in ways that are unintended (e.g., perceived rudeness or impatience of the experimenter).

In the following sections, I first review several studies that have explicitly addressed the role of meaning and interpretation. The message from these studies is that changing the meaning assigned to a task (e.g., how the procedure is described to the participants), without changing anything else about the experimental situation, can change the results by influencing participants' sensitivity and responsiveness to the situation. Next, I turn to another set of studies that illustrate how people detect normative dimension in a given situation, including the experimental situation. These studies suggest that norms can be detected rapidly, automatically, and without explicit instructions. Participants can move from "X happened" to "X ought to happen" or from "X ought not to happen" to "X is acceptable." For our purpose, it is important that research participants might consider a type of action desirable or acceptable in a situation, without the researchers recognizing it. There might be, in other words, a mismatch between the normative situation perceived by the participants and those presupposed by the researchers. I will next turn to several examples in experimental research that involve neglecting possible changes in meaning. The possible mismatch between how participants experience events, on the one hand, and what researchers believe about the participants' experience, on the other hand, is important with regard to the meaning of the experimental findings. The present argument, therefore, intends to show the problem of interpretation in experimental psychology and the continuing relevance of theoretical psychology in experimental research.

Meaning of Events

In explaining human performance, it might appear that all we need is third-person knowledge of the structure of the task, including what features of the environment people are acting upon and what types of movements are available to them. Third-person knowledge of a task does, under some circumstances, give us some predictive ability. Before beginning to discuss variations in task performance, however, the participants' intention (e.g., agreeing to respond to stimuli according to the instructions), as well as their selective attention and interpretation of the situation, is assumed to (a) mirror the researchers' instructions and (b) be fixed throughout the experimental session and across conditions. Such assumed transparency enables researchers to bracket out the participants' interpretation and construct third-person

descriptions that appear complete and independent of the participants' interpretation (Mammen, 2017). In the present section, I review research from several fields in cognitive-experimental psychology, all sharing a common theme: participants' understanding of a task (first-person perspective) can change how the task is performed and, consequently, what factors that can further influence performance, even in the absence of any overt change (third-person description) in the situation and task features.

In simple stimulus-response (S-R) tasks, people are on average faster with congruent S-R arrangements (left and right keys paired, respectively, with left and right stimuli) than with incongruent arrangements (left and right keys paired, respectively, with right and left stimuli). A variant of this phenomenon is the effect of the relation between an irrelevant spatial feature of the stimulus and the response (Simon, 1990). For example, imagine a task in which people are instructed to use left/right keys to respond to high−/low-pitched tone. The tone could be emitted either from the left or the right side. The pitch of the tone is relevant to the task, and its location is irrelevant. Nevertheless, on average, responses are faster when the tone and response are congruent in their location, compared to when they are incongruent. This is known as the Simon effect (Hommel, 2011).

In a landmark study, Hommel (1993) reported that the Simon effect can be reversed just by changing the instructions delivered to the participants. In his experiment, participants responded to low−/high-pitched tones using left/right keys. Each key was connected to a light on the opposite side, such that pressing the right key would illuminate a light on the left-hand side and pressing the left key would illuminate a light on the right-hand side. What was special about this task was the ambiguity of the response location. How do you respond to a low-pitched tone? Do you press the *key* on the *left*-hand side or turn on the *light* on the *right*-hand side? Both descriptions were available to the participants. Hommel (1993) divided the participants into two groups. For the first group, the instructions emphasized the location of the keys ("press the left key if you hear low tone"). For the second group, the instructions emphasized the location of the lights that turned on with key-press responses ("turn on the light on the right side if you hear the low tone"). Despite identical physical properties of the two conditions, the Simon effect was reversed across the two groups. Thus, the intended response location, i.e., how participants conceived of the responses, made a qualitative difference in the interaction between stimulus and response features.

Another example of how interpretation can change a behavioral effect was found in the joint version of the Simon task. Imagine the Simon task performed by two participants, sitting side by side, such that each is responsible for one response (Sebanz et al., 2003). For instance, two participants respond to the onset of a single red/green visual stimulus that could appear on the left/right side of the screen. We instruct the participant sitting on the left side to respond to the green stimuli, ignoring stimulus location, while instructing the participant on the right-hand side to respond to the red stimuli, again ignoring stimulus location. In this setup, when a stimulus appears on the left side, regardless of its color, the participant sitting on the left tends to be (on average) faster than the participants sitting on the right side. Likewise, in response to a stimulus on the right side, the participant sitting on the

right would be, on average, relatively faster. This *joint* Simon effect suggests that two actors take into account, in addition to what they are instructed to do, the role of their co-actor.

Hommel et al. (2009) introduced an additional manipulation in a joint Simon task. Before the co-actors began performing the task together, the authors induced either a cooperative or a competitive relation between them. In the cooperative condition, the authors found a joint Simon effect, suggesting that participants took each other's role into account. By contrast, in the competitive condition, the authors did not find a joint Simon effect. These findings suggest that changing the way two people think about each other can affect the organization of a shared task and, subsequently, how the shared task is susceptible to further manipulations, even if the overt physical structure of the task remains the same.

With a task that is performed by one person, the presence of another person who observes task performance can change the meaning of the task, even when all other aspects are the task and instruction are kept the same. Sartori et al. (2009) asked participants to perform a series of manual actions (reaching and lifting) with some objects on a table. In the "individual" (control) condition, participants performed the actions alone according to the instructions. When they were being observed by a fellow participant, however, their movements differed in subtle ways. The researchers reasoned that the communicative intention—showing one's movement to the other person—changed performance characteristics. A follow-up experiment confirmed the role of communicative intention, as opposed to the mere presence of another person, by testing the effect of a blindfolded person. As predicted, the blindfolded person did not cause changes in movement characteristics, compared with the "individual" condition (Sartori et al., 2009; see also Quesque et al., 2013, for a related demonstration in experimental economics, and see Dana et al., 2007).

The next example is from a task-switching study. Experiments that investigate task switching require participants to learn and prepare for two distinct tasks, performing only one task per trial (e.g., Task 1, judging a number's parity; Task 2, judging if the number is smaller/larger than 5). A cue might determine which task is performed at any given time (e.g., the color of the number). The typical finding is a performance cost (slower response time and reduced accuracy), when the task switches from trial n to trial $n + 1$, compared with when the task repeats. But does the switch cost depend on participants' understanding that they are, indeed, performing two distinct tasks? To answer this question, Dreisbach et al. (2007) provided different instructions for the same task (involving eight stimuli and two responses). One group of participants was instructed to perform a single task involving eight alphanumerical items mapped, arbitrarily, onto two responses. For the second group, the task was described, less arbitrarily, in terms of two sub-tasks: a "number task" and a "letter task" (each sub-task associated with four stimuli and two responses). Dreisbach et al. found a robust switch cost (when the stimulus switched from letter to number and vice versa) in the latter group and found no switch cost in the former group. In short, the presence of a switch cost, which is itself indicative of how a task is organized, depended on participants' understanding that there were two distinct tasks (Gozli, 2019, Chap. 5).

Another example comes from a visual-search study. Huffman et al. (2017) used a visual search, in which participants looked for a non-salient target (a green circle among green squares). A salient distractor could also be present on the display (a red square among the green items), which was supposed to be ignored. In one condition, the location of the salient distractor changed randomly from trial to trial, while in another condition, it was predictable, moving in a clockwise pattern. First, Huffman et al. found that the salient distractor interfered with performance more when its location changed predictably, compared to when it changed randomly. More relevant for our purpose, the researchers compared participants who noticed the predictability of the distractor location with those who did not notice it. The former group was presumably more likely to keep track and *actively* ignore the distractor, but it was for this group that the distractor caused the largest interference. Therefore, for the same task, with the same search arrays and the same instructions, participants' understanding that a salient distractor is predictable and should be ignored increased the cost of the distractor on performance. One might argue that participants' awareness of the predictable distractor resulted from their noticing the higher-performance cost of the distractor and not the other way around. Regardless, what is important in the present context is the association between, on the one hand, different meanings assigned to the distractor (predictable vs. unpredictable) and, on the other hand, different performance costs of the distractor. Without inquiring about the participants' points of view and asking whether they were keeping track of the distractor, we could only attribute variations in performance to external factors.

As a final example for this section, we can turn to a study on the effect of perceived effort and commitment of a partner in shared task. Székely and Michael (2018) measured participants' commitment to a game, defined as the time taken before the participant quits a round of the game. The authors used a computer-mediated two-party game, which became increasingly boring over time. Participants were playing with a "partner" that was, in fact, a computer algorithm. At the beginning of each round, the "partner" had to unlock the round by solving a CAPTCHA problem that was either easy or difficult. The participant could either end the round, by pressing the space bar, or wait for their partner to solve the problem. Participants showed more commitment to the game as a result of the perceived effort of their partner. This was observed only when they believed the partner was another person (Experiment 1) and not when they believed the partner was a computer program (Experiment 2) or when they played the game alone (Experiment 3). Therefore, even when the superficial characteristics of a game remain the same, participants' decisions change based on the meaning they attribute to events.

Detection and Adoption of Norms

Meaning of events and situations can vary across many different dimensions, and central among them are those groups under the "normative" category (Brinkmann, 2010). Norms feature in our experience not as isolated and detached individuals but

as members of communities and groups (Searle, 1995). A norm cannot remain a norm if everyone around you, or everyone in your group, violates it. Thus, the way we think about a norm is sensitive to how others—particularly others in our group—regard the norm. In reference to the meaning of actions, the normative meaning can change with a change in the social context (Bergner, 2016), particularly in response to how other people evaluate the given action.

If someone breaks a norm (or a rule) repeatedly, then their evaluation of the norm or their self-judgment might change in order to rationally adjust to their own behavior (Festinger, 1964). Imagine a person violating a norm for the first time (at time t_1), then for the second time (t_2), and so on until the tenth time (t_{10}). Among other considerations, we ought to consider whether and how the meaning of the norm changes in the perspective of this person. Particularly when breaking the norm is associated with a positive outcome, and no negative outcome, it is possible that the violation is regarded as more acceptable at t_{10} than at t_1.

Our general sensitivity to norms has been demonstrated in developmental studies with children. A study with 3-year-olds suggested that when children observe someone play with a toy for the first time, even once, they interpret the use in a normative sense ("one ought to play with the toy in this manner"; Schmidt et al., 2016). This claim was based on the observation that when the children later see someone else play with the toy in a different way, they object and try to enforce the way in which the toy "ought to" be used.

Efficient adoption of norms in children has been linked to the phenomenon of over-imitation. In imitating others, human children imitate both causally relevant and causally irrelevant steps of procedure, compared to non-human primates who tend to imitate only the causally relevant steps (Whiten et al., 2009). Kenward et al. (2010) asked whether over-imitation in human children should be attributed to norm learning or to a distorted causal learning. They designed an apparatus through which children were instructed to retrieve objects. Children were then asked to provide verbal explanations for their understanding of the causal mechanisms. The apparatus was in a transparent case, and the instructions to reach the objects included both causally relevant (moving the objects with a stick) and irrelevant actions (inserting the stick into an empty compartment). Although the children's action included the irrelevant action, their explanations suggested that they performed the unnecessary action out of norm learning and not for misunderstanding the causal mechanism.

The label "unnecessary action" (i.e., instrumentally superfluous) should be used with caution. Picking up and conforming to seemingly unnecessary actions play a crucial role in communication and in sustaining our social-cultural reality (Toomela, 2016). A polite gesture could appear unnecessary while being *not* unnecessary (such double-negations are discussed by Engelsted, 2017), if it signifies something about the relationship, its cultural embeddedness, its history, and its anticipated future. Thus, our capacity to detect norms or rules rapidly and flexibly should be viewed in light of the complex and dynamic nature of our social reality (Mammen, 2008). Indeed, the presence of another person can change the likelihood of rule-breaking, depending on their stance on (the meaning they assigned to) rule-breaking. Simons-Morton et al. (2014) studied rule violations in young males' driving with a driving

simulator with or without a passenger (confederate). They compared traffic rule violations when participants drove (a) alone, (b) with passenger who was accepting of risk, and (c) in the company of a passenger who was aversive to risk. They found that the presence of a risk-accepting passenger can increase the likelihood of committing traffic rule violation. These findings demonstrate the social nature of rules and norms and the role of others' perspective in determining what one ought to do.

The influence of others can be exerted indirectly, based on the observable effects of their actions. It might be possible, for instance, that observing the violation of one norm could promote the violation of other norms. Keizer et al. (2008) described such phenomena as "cross-norm inhibition" (p. 1683). In their first study, Keizer et al. staged different conditions in public bicycle-parking areas. In their so-called "disorder" condition, they covered the wall with graffiti right next to a clearly visible sign prohibiting graffiti. In the "order" condition, they kept the wall clean. Moreover, flyers were attached to the handlebars of the parked bicycles, and the owners had to remove it before riding the bicycle. The question was whether people litter when they remove the flyers attached to their bicycles, particularly in the presence of an already existing violation. The authors found that, when seeing rule violation in one domain (graffiti against the rules), people become more likely to violate a rule in another domain (littering).

Neglecting Meaning

In this section, I turn to several studies that seem to have neglected the role of interpretation and meaning. Explanations are offered with reference to situational factors, but not with reference to the possible differences in subjective meaning assigned to those factors. Researchers might only be interested in how participants respond overtly to stimuli and how those responses change with changes in situational factors, without concern for the meaning assigned to those events. Careful experimental designs and instructions, which negotiate and clarify the meaning of events prior to collection of data, are attempts to side-step to the issue of meaning (Wachtel, 1973). Nevertheless, it is possible that the intended manipulations result in unintended changes in participants' interpretations, with some degree of individual differences. The following studies are selected from a diverse set of topics related to decision-making, rule violation, second-language effects, meditation, and cooperative/punitive behavior. The studies share in common an insistence on describing the experimental manipulation and their effects from a third-person viewpoint, neglecting the possibility that the (first-person) meaning assigned by the participants might be a confounding factor and perhaps the primary explanation of the findings. Taking changes in meaning into account, as we shall see, results in a view of the research findings that fundamentally diverges from the one provided by the researchers.

The experimental-cognitive studies of rule violation have insisted on adopting a third-person perspective on participants' behavior, which requires fixing the

meaning of "rule" and "rule-breaking" across different conditions, ignoring possible variations of meaning in the participants' perspective. For instance, research initiated by Pfister and colleagues examines the potential costs of rule violation on performance (Pfister et al., 2016a, b). More generally, this research aims to offer description of rule-following and rule-breaking action in terms of relatively low-level, sensorimotor characteristics of actions, although this requires specifying the task requirement at the higher level of abstraction (i.e., determining what counts as rule violation). Researchers have found, using target-directed movement tasks, that hand movements that violate a rule tend to be on average slower, both in their initiation and in their completion, and their trajectory tends to deviate from a straight path, compared to movements that conform to the rule (Pfister et al., 2016b; Wirth et al., 2016).

Wirth et al. (2018a) found that the costs associated with rule violation are eliminated if the rule violation (a) is performed frequently and (b) has been committed recently. The authors assumed that increasing the frequency of rule-violation trials and their recency does not change the meaning of the behavior with respect to the rules. This assumption would be inconsistent with the research on norms, which suggests we adopt norms rapidly and flexibly in response to the changing contexts and in response to changes in our own behaviors (Ting, 2018). If we accept that the meaning of rule (violation) changed across the conditions, then recency and frequency are confounded with meaning. Rather than stating that participants are now committing the same action ("rule-breaking") with more efficiency, we might have to state that participants are performing a different type of action ("breaking a relatively strict rule" vs. "breaking a nominal and flexible rule"). The meaning assigned to an action might change during the experiment, as an unintended outcome of experimental manipulations.

Another example highlights how tasks might be differently interpreted by different groups of participants. Kozasa et al. (2012) aimed to test the effect of long-term meditation practice on attention. They compared the performance of regular meditators and non-meditators in a Stroop task (Stroop, 1935). In a Stroop task, participants are instructed to report the color (ink) of words, one word at a time, while ignoring the meaning of the words. For each color-word stimulus, the color and the meaning of the word could be congruent (RED typed in red), incongruent (GREEN typed in red), or unrelated (SHELF typed in red). Compared to neutral trials, performance is usually better on congruent trials and worse on incongruent trials. Kozasa et al. (2012, p. 745) reasoned that the Stroop task requires "attention and impulse control" and asked whether these abilities are superior in regular meditators. Using brain imaging, the authors found increased activity on incongruent trials relative to congruent trials but only for non-meditators. The increased brain activity, associated with incongruent color-word stimuli, was not observed in regular meditators.

How do we know that the meditators and non-meditators had the same understanding of the Stroop task? Recall that Huffman et al. (2017) showed how an active approach to ignoring a distractor might increase the cost of the distractor. Could a similar and costly strategy have been adopted by the non-meditators in the Kozasa et al.'s (2012) study? If we assume that there is only one way to understand and

engage with this task, or at least both groups of participants understood the task in the same manner (deploying the same set of cognitive capacities), then we could conclude that meditators were relatively more efficient than the non-meditators, in exercising a set of common capacities. If, however, the meditators took a different approach to the Stroop task and, in effect, performed a different task, then we cannot conclude that the meditators had a quantifiable increase with regard to the same capacities. For example, the regular meditators might have taken a more passive approach to the irrelevant word meaning—letting go of the irrelevant stimulus feature rather than actively suppressing it—making it unnecessary to exert cognitive control on monitoring color-word conflict. In other words, rather than enhancing the ability to cognitive control (in response to interference), meditation might promote an understanding of task performance that reduces the necessity for cognitive control (by *letting go* of the source of interference).

In social and cooperative situations, different interpretations of actions can similarly result in different outcomes. Six-year-olds have been found to intervene in an unfair interaction, even when that is costly for themselves (McAuliffe et al., 2015). When observing another child unfairly and selfishly dividing candies with a partner, children demonstrate a tendency to intervene; they refuse the allocation, thereby punishing the "selfish" allocator. We might describe the perspective of the child who can intervene as a detached third party, and this characterization might be reasonable in certain conditions. However, some manipulations might change the child's perspective. Consider, for instance, that the intervention itself can become costly (i.e., intervention could cost the child a candy). Here, the intended manipulations are described as the *fairness* of the allocation decision and the *cost* of intervention, though the two factors might interact and change the meaning assigned to the act of intervention. Choosing a costly intervention might involve the child's standpoint having shifted from a neutral third party to an "ally of the underdog," given that both are placed in positions of disadvantage. If so, what happens in the experiment might not fit the neat and clear categories imposed by the experimental design. While experimenters claim to establish a causal link between variables, the cost of intervention, and the probability of intervention, they end up instead constructing different dramaturgical scenarios (narratives) in the different conditions (Harré, 1993). If we acknowledge the differences between conditions as differences between the dramaturgical scenarios, we will not assume the same variables are at play across the conditions.

Turning to a line of research quite different from what I have discussed above, though it helps further illustrate my main point, let us now consider the so-called second-language effect on cheating behavior. The background for this line of research are studies on cheating, in which participants roll a die and report whether the outcome matched a pre-specified target number (in which case they would receive monetary reward). The die is rolled inside a cup, and no one other than the participant can see the outcome. There is, therefore, an incentive to cheat, and participants can cheat without being detected, although the rate of cheating can be estimated at the group level (Hilbig & Thielmann, 2017). Interestingly, Bereby-Meyer et al. (2018) reported that the rate of cheating is lower when people use their

second language (L2) during the study session, compared to when they use their native language (L1). The authors interpreted this effect in terms of the different efficiency of information processing across different languages. Presuming that we are less efficient with L2, communicating in L2 would prolong decision-making, increasing the probability that the relatively slow and rational modes of thought dominate our decision. Alternatively, presuming that using L2 and cheating are both associated with extra cognitive effort, participants might decide to behave honestly merely to avoid the additional effort (Pfister et al. 2016a, b). These interpretations both treat cheating as the same type of action (qualitatively equivalent) across the two language conditions.

It is possible, however, that different languages evoke different norms, leading participants to assign different meanings to a dishonest expression depending on the medium of expression. Communicating in L2, particularly for novices, might be associated with a range of relationships and sociocultural positions (teachers, fellow language students, being in the out-group, etc.), distinct from those that are dominantly associated with L1. It is possible that communicating in L2 evokes situations in which dishonesty is either less accessible or evaluated more negatively. Wirth et al. (2018b) have argued that rule-breaking sensitizes participants to the concept of authority. This phenomenon might itself depend on the medium of communication. In particular, the concept of authority might be more accessible, and more easily evoked, when using L2. If so, then a dishonest expression in L1 and L2 ought to be considered as qualitatively different actions, by virtue of evoking different meanings (Bergner, 2016; Gozli, 2017). This interpretation stands in contrast to the approach that compares the probability of cheating, *as the same type of action*, across different conditions (for a review, see Gozli, 2019).

The second-language effect has been investigated on other forms of decision-making bias. The studies on the so-called framing effect show that, for the same decision, people can become more or less tolerant of risk, depending on whether their choice is presented with respect to its potential loss or its potential gains (Tversky & Kahneman, 1981). Recent studies have shown the framing effects can decrease if the choices are presented in L2, compared to when they are presented in L1 (Costa et al., 2014; Keysar et al., 2012). Keysar et al. (2012) found no evidence of a framing effect in L2. They interpreted the results in terms of a cognitive and emotional distance that comes with using L2, which in turn could result in more rational responses. Similarly, Costa et al. (2014) presented participants with moral dilemmas and found that a higher number of participants responded in a utilitarian ("rational") manner, e.g., saving five people by killing one person, when the dilemma is presented in L2.

Again, it is possible that certain heuristics—or decision "shortcuts"—are less accessible in L2. Embodied cognition perspectives are relevant for these lines of research (Barsalou, 1999; Gallese & Lakoff, 2005). The meaning of words is grounded in perceptual-motor experience and bodily affective sensations. Concepts expressed in L1 could be associated with stronger embodied correlates compared to L2. An embodied feeling of loss/gain, and its impact on decision-making, might be more easily evoked in L1; similarly, a utilitarian response to a moral dilemma might

evoke more aversive connotations, or embodied meaning, when considered in L1, compared to when it is considered in L2.

A recent study by Korn et al. (2018) fits with the idea that the effect of L2 on the framing effect might be due to differences in meaning, rather than differences in cognitive effort. Korn et al. attempted to generalize the foreign-language effects in terms of the more abstract construct, such as cognitive fluency or cognitive effort. To do so, instead of presenting problems in L1 and L2, they presented them either in an easy-to-read font (fluent, low effort) or in a hard-to-read font (dysfluent, high effort). In their first two experiments, recruiting 158 and 271 participants, the authors found no effect of font type. In an online version of the experiment, recruiting a larger sample of 732, they found a small effect of font type in the expected direction: the framing effect was smaller with the hard-to-read font. The difficulty in obtaining the effect, as reported by Korn et al. (2018), lends support to my alternative account of the second-language effect. Beyond variations in cognitive fluency or efficiency, the differences between L1 and L2 might be related to variations in meaning and interpretation.

A final example comes from research comparing human-human and human-computer interaction. Tenbrink et al. (2010) compared how people gave route instructions to (1) another person and (2) to a computer program. Instructions varied more widely in the former condition, compared to the latter. In other words, when humans addressed a computer, they remained within a narrower range of expressions. While the intended manipulation in the study was the addressee type (two categories: human vs. computer program), the unintended manipulation (meaning) might have been the perceived linguistic competence of the addressee. Assuming that a higher level of linguistic competence in an addressee might lead us to use language more freely and flexibly, consequently introducing more variety in our expressions. If this explanation holds, we would expect to see a similar pattern of findings if we compare how people communicate to computer programs of different degrees of sophistication. Moreover, in human-human interactions, we would also expect differences in how people communicate to native speakers of a language and novice speakers. It would, therefore, be worthwhile to explore variations in how people interpret the context and their addressees.

Conclusion

The aim of the present chapter was to highlight the role of interpretation (on the part of research participants) in experimental research (Gozli, 2019; Toomela, 2008; Valsiner & Brinkmann, 2016). Aiming to discover causal relations among variables of interest, experimental researchers attempt to isolate variables in relatively simple laboratory settings (Gozli & Deng, 2018), in which events and behaviors are described with respect to a set of already determined categories (Gozli, 2017; Mammen, 2017). When the effect of rule-breaking is concerned, for example, researchers aim to keep all else equal across conditions, except for whether a given

action reflects "conforming to" or "violating" a rule; similarly, when the effect of long-term meditation is concerned, we aim to keep all else equal, except for the prior meditation experience of the participants. Accordingly, we assume that the meaning of events does not change systematically—from the perspective of the research participants—as we introduce other manipulations. We might implicitly assume that rule-breaking preserves its meaning within the context of the task, regardless of the frequency and recency of rule-breaking. Similarly, we might assume that the meaning assigned to a distracting item (i.e., how one ought to think, and what one ought to do, about the distractor) is the same in the meditator and non-meditator groups.

My argument consisted of three stages. First, I reviewed several examples from experimental psychology, where task interpretation was explicitly manipulated—by changing the task instructions—and was found to impact performance in the task. These changes do not require changing the physical features of the task. Second, I discussed a particular dimension of meaning, namely, the normative dimension, within which actions could be identified as "good," "bad," "unacceptable," "desirable," and so forth. Research has demonstrated our ability to detect norms rapidly and flexibly, our sensitivity to the presence of others, and our sensitivity to the perspective of others in evaluating our own actions. Third, I reviewed several different lines of research (on rule-breaking, meditation, intervention by children, and second-language effects), where there are clear possibilities for the intended manipulations to be confounded with changes in participants' interpretation of the events. If my criticisms are valid, it would mean that the experiments did not isolate the variables of interest (e.g., meditation ➔ distractor suppression) within controlled processes (i.e., one and the same task). Rather, different conditions correspond to qualitatively different tasks, which could not be neatly compared only with respect to the intended manipulations and measures.

Disentangling meaning of an event from experimental manipulations is not straightforward (Bergner, 2016). Meaning is not an isolable part of the experimental setting, but has to do with how the setting is framed. When the meaning of a given task/event changes from the point of view of the research participants, it could mean that (a) what the participants view as potentially available is different, including potential social incentives and prohibitions; (b) what the participants regard as acceptable, good, bad, and so forth, in the experimental context might be different; and (c) the presumed purpose of the experimental session might change. All these might happen despite the researchers' belief in the validity of their prior categories of description and evaluation.

Once we recognize that our intended manipulations can change the meaning of events from the perspective of research participants, we will allow ourselves to describe differences between conditions in qualitative terms, in terms of tasks that differ in kind, and with respect to different normative standards. What else could change, for instance, when a problem is presented in L2, rather than L1? Might long-term meditators think differently about the task, which subsequently renders them differently prepared for the events in the experiment? Might a child interpret intervention differently, once we associate the act of intervention with personal

cost? Pursuing such questions requires including the participants' perspective, rather than an insistence on a uniform, third-person description. Social situations, including experimental sessions that include researchers and participants, are usually open to multiple descriptions and transformation, which is to say they have depth (Mammen, 2017, 2019). What actually happens, therefore, as a result of our experimental manipulations might not fit within the pre-specified categories of description and evaluation.

Acknowledgments I would like to thank Carol Ting, Heath Matheson, and Kieran O'Doherty for reading earlier drafts of this chapter and offering helpful comments.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–660.

Bereby-Meyer, Y., Hayakawa, S., Shalvi, S., Corey, J., Costa, A., & Keysar, B. (2018). *Honesty speaks a second language*. Topics in Cognitive Science. Online first.

Bergner, R. M. (2010). What is descriptive psychology? An introduction. In K. Davis, F. Lubuguin, & W. Schwartz (Eds.), *Advances in descriptive psychology* (Vol. 9, pp. 325–360). Descriptive Psychology Press.

Bergner, R. M. (2016). What is behavior? And why is it not reducible to biological states of affairs? *Journal of Theoretical and Philosophical Psychology, 36*(1), 41–55.

Brinkmann, S. (2010). *Psychology as a moral science: Perspectives on normativity*. Springer.

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS One, 9*(4), e94842.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67–80.

Dreisbach, G., Goschke, T., & Haider, H. (2007). The role of task rules and stimulus–response mappings in the task switching paradigm. *Psychological Research, 71*(4), 383–392.

Engelsted, N. (2017). *Catching up with Aristotle: A journey in quest of general psychology*. Springer.

Festinger, L. (1964). *Conflict, decision, and dissonance*. Stanford University Press.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3–4), 455–479.

Gozli, D. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology, 27*, 741–758.

Gozli, D. (2019). *Experimental psychology and human agency*. Springer.

Gozli, D., & Deng, W. (2018). Building blocks of psychology: On remaking the unkept promises of early schools. *Integrative Psychological and Behavioral Science, 52*, 1–24.

Harré, R. (1993). *Social being* (2nd ed.). Wiley.

Hilbig, B. E., & Thielmann, I. (2017). Does everyone have a price? On the role of payoff magnitude for ethical decision making. *Cognition, 163*, 15–25.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research, 55*, 270–279.

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica, 136*(2), 189–202.

Hommel, B., Colzato, L. S., & van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychological Science, 20*, 794–798.

Huffman, G., Rajsic, J., & Pratt, J. (2017). *Ironic capture: Top-down expectations exacerbate distraction in visual search*. Psychological Research.

Johns, G. (2006). The essential impact of context on organizational behaviour. *Academy of Management Review, 31*, 386–408.
Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science, 322*(5908), 1681–1685.
Kenward, B., Karlsson, M., & Persson, J. (2010). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20101399.
Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*(6), 661–668.
Korn, C. W., Ries, J., Schalk, L., Oganian, Y., & Saalbach, H. (2018). A hard-to-read font reduces the framing effect in a large sample. Psychonomic Bulletin & Review, 25, 696–703.
Kozasa, E. H., Sato, J. R., Lacerda, S. S., Barreiros, M. A., Radvany, J., Russell, T. A., et al. (2012). Meditation training increases brain efficiency in an attention task. *NeuroImage, 59*, 745–749.
Mammen, J. (2008). What is a concept. *Journal of Anthropological Psychology, 19*(2), 25–27.
Mammen, J. (2017). *A new logical foundation for psychology*. Springer.
Mammen, J. (2019). *A grammar of praxis: An exposé of "a new logical Foundation for Psychology", a few additions, and replies to Alaric Kohler and Alexander Poddiakov*. Integrative Psychological and Behavioral Science. Online first.
McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition, 134*, 1–10.
Pfister, R., Wirth, R., Schwarz, K. A., Foerster, A., Steinhauser, M., & Kunde, W. (2016a). The electrophysiological signature of deliberate rule violations. *Psychophysiology, 53*, 1870–1877.
Pfister, R., Wirth, R., Schwarz, K., Steinhauser, M., & Kunde, W. (2016b). Burdens of nonconformity: Motor execution reveals cognitive conflict during deliberate rule violations. *Cognition, 147*, 93–99.
Quesque, F., Lewkowicz, D., Delevoye-Turrell, Y. N., & Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Frontiers in Neurorobotics, 7*, 14.
Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Consciousness and Cognition, 18*, 766–772.
Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science, 27*, 1360–1370.
Searle, J. R. (1995). *The construction of social reality*. Simon & Schuster.
Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*(3), B11–B21.
Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Elsevier.
Simons-Morton, B. G., Bingham, R. C., Falk, E. B., Li, K., Pradhan, A. K., Ouimet, M., Almani, F., & Shope, J. T. (2014). Experimental effects of injunctive norms on simulated risky driving among teenage males. *Health Psychology, 33*(7), 616–627. https://doi.org/10.1037/a0034837
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662.
Székely, M., & Michael, J. (2018). Investing in commitment: Persistence in a joint action is enhanced by the perception of a partner's effort. *Cognition, 174*, 37–42.
Tenbrink, T., Ross, R. J., Thomas, K. E., Dethlefs, N., & Andonova, E. (2010). Route instructions in map-based human–human and human–computer dialogue: A comparative analysis. *Journal of Visual Languages & Computing, 21*(5), 292–309.
Ting, C. (2018). *The feedback loop of rule-breaking: Experimental evidence*. The Social Science Journal. Online first.
Toomela, A. (2008). Variables in psychology: A critique of quantitative psychology. *Integrative Psychological and Behavioral Science, 42*(3), 245–265.

Toomela, A. (2016). What are higher psychological functions? *Integrative Psychological and Behavioral Science, 50*(1), 91–121.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Valsiner, J., & Brinkmann, S. (2016). Beyond the "variables": Developing metalanguage for psychology. In S. H. Klempe & R. Smith (Eds.), *Centrality of history for theory construction in psychology* (pp. 75–90). Springer.

Wachtel, P. L. (1973). Psychodynamics, behavior therapy, and the implacable experimenter: An inquiry into the consistency of personality. *Journal of Abnormal Psychology, 82*(2), 324–334.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2417–2428.

Wirth, R., Pfister, R., Foerster, A., Huestegge, L., & Kunde, W. (2016). Pushing the rules: Effects and aftereffects of deliberate rule violations. *Psychological Research, 80*(5), 838–852.

Wirth, R., Foerster, A., Herbort, O., Kunde, W., & Pfister, R. (2018a). This is how to be a rule breaker. *Advances in Cognitive Psychology, 14*(1), 21–37.

Wirth, R., Foerster, A., Rendel, H., Kunde, W., & Pfister, R. (2018b). Rule-violations sensitise towards negative and authority-related stimuli. *Cognition and Emotion, 32*(3), 480–493.