

# Your Highness: Vertical Positions as Perceptual Symbols of Power

Thomas W. Schubert  
University of Jena

Metaphorically, power equals up. Drawing on embodied theories of cognition, the author argues that thinking about power involves mental simulation of space and can be interfered with by perception of vertical differences. Study 1 assessed image schemas for power and found a shared vertical difference metaphor. Studies 2, 3, and 4 showed that the judgment of a group's power is influenced by the group's vertical position in space and motor responses implying vertical movement. Study 5 ruled out that the influence of vertical position on power judgments is driven by valence differences. Study 6 showed that vertical position also influences the power judgment result itself. The evidence suggests that the concept of power is partly represented in perceptual form as vertical difference.

*Keywords:* power, embodiment, perceptual symbols, metaphors

When we talk about power, we often use metaphors about *up* and *down*. Examples abound: When someone has a *high* status, or is *up* in the hierarchy, he or she has control *over* and can *oversee* others who have *lower* status. One can look *up* to those who *rose* to the *height* of their power or look *down* on *underlings*. When a picture of a hierarchy is drawn, the most powerful person is usually at the top, and the *subordinates* are drawn below. All these metaphors are cases of the “control is up, lack of control is down” metaphor (Lakoff, 1987; Lakoff & Johnson, 1980). In short, power is metaphorically described as a vertical dimension in physical space.

In psychological research, however, power is usually defined as the potential to influence others and to promote one's own goals (Ellyson & Dovidio, 1985; Keltner, Gruenfeld, & Anderson, 2003; McClelland, 1975). Although spatial positions have been considered insofar as they are cues to the power of others in the environment (Argyle, 1988; Mehrabian, 1972), we do not know what role they play in thinking about power. Are we missing something important? Is it the case that metaphors about the vertical dimension are used only to talk about power but not to think about it? Should social psychology leave metaphors to linguistics? The following arguments and data will try to show that by taking metaphors literally, we can discover what thinking about seem-

ingly abstract concepts like power has to do with perception and experience.

## Perceptual Symbols

Perceptual content is usually neglected in current theories of mental representation. Instead, mental representations are typically described as nodes that are abstract and amodal, which are connected to each other via associative links. The idea is that knowledge is represented in the nodes and their connections. Perceptual content is only an early input to the representations of concepts in these models and not regarded as important for conceptual thinking itself. As an alternative to such amodal models of mental representation, recent theorizing on the nature of human knowledge, categories, and cognition proposes that concepts are a lot less abstract than previously thought. Embodied theories of cognition argue that concepts include a lot more perceptual content than a clear divide between perception, cognition, and behavior would suggest (Barsalou, 1999; Glenberg, 1997). Empirical work, both from cognitive psychology and neuropsychology, supports this idea. In a review of evidence from several domains, Barsalou (1999) found that mental representations of concepts are still tied to their perceptual basis. Barsalou concluded that knowledge consists not of amodal propositions but of modality-specific or modal representations, which he calls *perceptual symbols*. Perceptual symbols are thought of as schematized perceptual experiences involving all senses, including proprioception, introspection, and motor programs. These schematized perceptual representations are then used in cognitive processes, such as perception, categorization, and judgment. These processes are made possible by using the perceptual symbols to construct and run simulations, similar to mental models (Johnson-Laird, 1983). In other words, thinking is argued to involve perceptual simulation.

It is interesting to note that an embodied view of cognition is compatible with recent findings and theoretical developments in

---

This research was partly funded by a research grant awarded to Thomas W. Schubert by the Friedrich-Schiller-Universität Jena. I thank Markus Brauer, Steffen Gießner, Arthur Glenberg, Jonathan Haidt, Johann Jacoby, Thomas Kessler, Sander Koole, Amélie Mummendey, Kai Sassenberg, Beate Seibt, and Sven Waldzus for helpful discussions and comments on earlier versions of this article, and Babett Voigt, Tobias Raabe, Susann Schmidt, Lisa Schubert, and Anna Steudel for collecting data.

Correspondence concerning this article should be addressed to Thomas W. Schubert, Department of Psychology, University of Jena, Humboldtstraße 26, 07743, Jena, Germany. E-mail: schubert@igroup.org

social psychology (cf. Barsalou, Niedenthal, Barbey, & Ruppert, 2003; Niedenthal, Barsalou, Winkielman, Kraut-Gruber, & Ric, in press). For example, the activation of a social stereotype has been shown to cause the unintended mimicry of behavior that is strongly associated with the stereotyped group (e.g., mimicry of the slow walking associated with the category of the elderly; Bargh, Chen, & Burrows, 1996; Dijksterhuis & Bargh, 2001). Such automatic behavior is explained by assuming that the mental representation of a group stereotype contains sensory-motor representations. This theoretical framework complements quasi-verbal associative networks with modal content. Further examples are findings showing that the processing of valence information is hindered by performing behavior that does not fit the valence (e.g., approach vs. avoidance movements, or nodding vs. head shaking; Förster & Stepper, 2000; Förster & Strack, 1996; Neumann, Förster, & Strack, 2003). Such bodily feedback effects are explained by arguing that motor representations interact with the cognitive processing involved in the cognitive tasks, which can be explained only if cognitive processing involves perceptual simulation. In sum, both theoretical developments and empirical evidence in cognitive and social psychology suggest that modal content plays a role in conceptual thinking. This encourages the search for perceptual content in such social concepts as power.

#### *Perceptual Simulation of Space for Concrete Concepts*

As the current goal is to investigate the spatial metaphor of power, most interesting to the present discussion are results that illustrate perceptual simulation of space in thinking about concepts. Indeed, there is evidence that points to a simulation of space in conceptual thinking. In one instructive study, Borghi, Glenberg, and Kaschak (2003) asked their participants to verify whether a given object (e.g., a car) had a certain part (e.g., a roof). To answer, participants had to press one of two keys on a vertically mounted keyboard. Affirmative answers had to be given either with a key that required an upward movement of the participant's arm or with a key that required a downward movement. Results showed that if there was a fit between the position of the part relative to the object (i.e., roof is at the top of the car) and the movement of the arm (i.e., upward), reactions were quicker. This supports the idea that conceptual thinking involves perceptual simulation: When we want to verify the property of a category, the concept is simulated mentally on the basis of perceptual knowledge, and the result is read off. Actual motor movement can either interfere with or facilitate this simulation (cf. Glenberg & Kaschak, 2003).

If a visual simulation of the words' referents underlies these effects, then spatial information provided by vision should also be able to interfere with or facilitate the simulation. This is what Zwaan and Yaxley (2003) showed. They presented two words simultaneously, and the participants' task was to judge whether the two were related or not (e.g., *root* and *branch*). The words were presented above each other, and the crucial manipulation was that their order either followed the canonical arrangement (i.e., *branch* above *root*) or contradicted it (i.e., *root* above *branch*). As predicted by a simulation account, relatedness affirmations were quicker when the arrangement of the words followed the canonical arrangement of the objects. By providing additional perceptual input, either motor (Borghi et al., 2003) or visual (Zwaan &

Yaxley, 2003; see also Richardson, Spivey, Barsalou, & McRae, 2003), perception interacts with these mental representations.

#### *Perceptual Simulation of Space for Abstract Concepts*

These results certainly encourage embodied views of knowledge; however, they are all related to concrete objects that can actually be observed in the environment. It is interesting to note that there is also evidence for perceptual simulation of space in thinking about abstract concepts, especially about valence and time (for further concepts, see Talmy, 1988).

Studying the embodiment of valence, Meier and Robinson (2004) showed that a vertical spatial dimension underlies valence representations. They tested this by letting participants evaluate words that appeared at either the top or the bottom of a computer screen. As predicted from a perceptual symbols account, positive words were evaluated quicker when they appeared at the top of the screen compared with the bottom of the screen, whereas the opposite was true for negative words. Apparently, judging valence involves simulation of a vertical spatial dimension, on which good is up and bad is down.

The hypothesis that time is thought about by imagining it as a space was already put forward by Jaynes (1976) and is supported by the existence of metaphors that we use to talk about time. The best experimental evidence comes from a set of ingenious experiments by Boroditsky and colleagues (Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Gentner, Imai, & Boroditsky, 2002). Time is imagined as a horizontal, not vertical, space, at least in Western cultures. Consequently, Boroditsky (2001) could show that judgments about temporal facts (e.g., whether January comes before June) are made faster when a preceding judgment has to be made about a horizontal spatial dimension than when it has to be made about a vertical spatial dimension. Time, an abstract dimension, is thought about in terms of spatial—that is, perceptual—terms.

These data support the notion that abstract concepts, like concrete concepts, are at least partially represented by perceptual symbols that relate the concepts to perceptual content. The conclusion is that conceptual thinking involves the simulation of this content and can therefore be influenced by priming or concurrent presentation of perceptual input.

#### *Perceptual Simulation of Space for Power*

These examples show that thinking about both concrete and abstract concepts can be influenced by spatial information that is canonically included in the construction of these concepts—either in the real environment or in metaphorical thinking. This supports the idea that power, an abstract social concept, also includes spatial information about the vertical dimension, as the metaphors suggest. In other words, the hypothesis is that when we think of power differences, we actually think of spatial differences.

Barsalou (1999) proposed that a perceptual symbol is derived from multiple sources of direct experience. He exemplified this for anger: To represent it, experiences of goal blocking, intense affective states, and behavioral responses are schematized. The same reasoning can be applied to mental representations of power. Here, direct physical experience of vertical differences might be schematized into a perceptual symbol of power. Indeed, the linguists Lakoff and Johnson (1980) listed this image as one of the "met-

aphors we live by,” and argued that the experiential basis lies in the fact that “physical size typically correlates with physical strength, and the victor in a fight is typically on top” (p. 15). As they keenly observed, there are two different spatial correlates of power: vertical size (height) and vertical position. The association of power and height is important for most mammals that are involved in physical fights: The larger animal is typically more powerful, or as Freedman (1979) noted, “throughout nature the rule is the bigger, the more dangerous” (p. 29). For humans, size matters a lot for negotiating power relations, especially during childhood and adolescence. Children learn that their taller parents are more powerful and that taller siblings or other taller children are able to coerce them physically. Schwartz, Tesser, and Powell (1982) noted that a “universal association of statural superiority and parental dominance” exists, which leads to an “invariant use of elevation symbolism in the representation of social dominance as a generalization of this elementary facet of experience” (p. 119; cf. Argyle, 1967; Spiegel & Machotka, 1974). This carries over into adulthood, where it is still the rule that taller persons use their physical advantage to gain power: Big people hit little people (Felson, 2002), and taller persons enjoy gaining higher wages, reaching higher status occupations, and (at least on average) winning presidential elections (Young & French, 1996; for a review, see Judge & Cable, 2004). Finally, according to Keltner and Haidt (2003), size or vastness is a central feature of the emotion awe, which, in its prototypical form, is felt toward powerful others. Perceiving something that is “much larger than the self, or the self’s ordinary level of experience or frame of reference,” (Keltner & Haidt, 2003, p. 303), together with the experience that one needs to accommodate one’s concepts in response to the event, leads to experiences of awe (see also Haidt, 2003).

For a relation between power and vertical position, there is plenty of evidence from the anthropological literature, although the difference between vertical position and vertical size is sometimes fuzzy. Often, vertical position acts as a surrogate of physical size. People construct power as vertical difference in language, in nonverbal communication, and in physical manifestations. Fiske (1992, 2004) has described extensively how authority ranking, the social and cognitive construction of who has power over whom, relies on vertical spatial difference. He observed that in order to differentiate between persons with and without power, people typically use metaphors of spatial order and magnitude. Citing evidence from diverse cultures, Fiske concluded that virtually all cultures use vertical markers for authority ranks in their language (for further references, see Schwartz, 1981; Schwartz et al., 1982). Furthermore, verticality embodies power almost everywhere in the domains of posture, housing, and furniture (Hewes, 1955). Many of these manifestations use height and size simultaneously or interchangeably, such as when the powerful have the largest house with the highest tower, sit at an elevated seat during meals, and are addressed as “Your Highness.”

Taken together, extensive evidence from diverse sources suggests that vertical space easily affords the linear ordering necessary for the creation of a hierarchy. Human children are surrounded by a world in which power is over and over again correlated with vertical positions, between people, in language, and in artifacts. Humans may even be evolutionarily prepared to pick up associations of power and spatial positions (Fiske, 2004). These are ideal

preconditions for the schematization of a perceptual symbol out of a multitude of perceptual events and for the development of a strong association between vertical position and power (Barsalou, 1999).

The logic behind this schematizing of experiences can be made clearer by juxtaposing metaphors that refer to a perceptual symbol with those that do not. Not every metaphor is still “alive” and tied to actual experiences; some are learned and used without referring to perceptual content. One example of the latter type is the metaphor that conservative parties are “right,” whereas liberal parties are “left.” These labels originally referred to an actual seating order in the 1789 French National Assembly—that is, an actual experience of a spatial order. Although many modern parliaments maintain this seating order, laypersons today probably do not share this experience enough for it to be schematized, and therefore it is likely that today only the spatial labels “left” and “right” are associated with conservatism and liberalism, but not the actual spatial representation. If the current hypothesis on power is correct, however, the powerful = up and powerless = down perceptual symbol is different from the conservative = right and liberal = left metaphor in that in the case of power, everybody shares the underlying experience that then becomes schematized. If vertical positions are perceptual symbols of power, thinking about power should be influenced by the perception of vertical spatial differences. This hypothesis is tested in the following studies, examining judgments of power as one type of thinking about power.

### Overview of the Current Research

The following six studies were designed to test the hypothesis that the concept of power involves a perceptual simulation of vertical differences in space. Study 1 assessed image schemas held by laypersons for propositions that describe power and powerlessness. Study 1’s goal was to verify that these metaphors are schematic and shared. Studies 2, 3, and 4 investigated whether perceived vertical differences interfere with judgments of power. In these studies, predictions of the perceptual symbols hypothesis were tested by investigating whether judgmental speed and accuracy can be influenced by the spatial position of the groups to be judged. In Study 2, pairs of groups were used as stimuli, and both visual input and the motor response varied such that they were compatible or incompatible with the power relations. Study 3 isolated the motor response, and Study 4 isolated the visual input, with the common goal of testing whether each alone can interfere with power judgments. By providing pretest and manipulation check data on the valences of powerful and powerless groups, Studies 2–4 also addressed an alternative explanation, namely, that powerful agents are associated with *up* because they are positive and positive valence is associated with *up* (Meier & Robinson, 2004). In addition, Study 5 tested experimentally whether valence predicts interference judgments on power judgments or only on valence judgments. Finally, Study 6 went beyond influencing the speed and accuracy of judgments and shows that the amount of power attributed to an agent can be influenced by perceiving vertical differences.

## Analytic Strategy Concerning Response Times and Errors as Indicators of Interference

Stroop-like interference effects can typically be observed on both response latencies and accuracy. Although the latter index has received less empirical attention, both are equally good indicators, and both deserve equal attention in any two-choice reaction task, because both are outcomes of the same processes (Ratcliff & Smith, 2004). In many interference paradigms, the effect on response latencies is more pronounced because the stimuli are selected such that their judgment should be unambiguous and, as a consequence, the total number of errors is low. Effects on accuracy show up only when external factors lead to a higher number of errors (e.g., the requirement to answer within a short response window) or when the stimuli themselves become harder to judge (MacLeod, 1991). If this is not the case, many more trials are necessary to get reliable results on accuracy (Ratcliff & Smith, 2004). In the following studies, effects on both response latencies and error frequencies will be reported and combined across all studies in a meta-analysis at the end. Effects are primarily predicted to occur on response latencies, unless other factors lead to more difficult judgments.

### Study 1: Surveying Power

In the introduction, several examples for a vertical spatial metaphor of power were cited. Still, intuition notwithstanding, it has first to be shown that there really is a shared power = up schema. In Study 1, I sought to test this hypothesis. To do so, I based Study 1 on a method developed by Richardson and colleagues (Richardson, Spivey, Barsalou, & McRae, 2003; Richardson, Spivey, Edelman, & Naples, 2003). They identified both horizontal verbs (e.g., pull) and vertical verbs (e.g., sink) by asking participants to associate a verb with one of several tilted lines. Study 1 assessed whether being powerful and being powerless is associated with high and low positions in space, respectively.

### Method

#### Overview and Design

For 18 propositions, participants answered which one of eight pictures best represented the proposition. In 6 of the propositions, the agent (represented by a small black circle) was more powerful than the patient (represented by a small white circle): for example, “● has influence on ○.” In another 6 propositions, the agent was less powerful than the patient: for example, “● is weaker than ○.” Finally, 6 propositions described horizontal relations between agent and patient: for example, “● pulls ○.” The eight pictures depicted eight possible angles between agent and patient (see Figure 1). Because the agent’s color (black or white) was counterbalanced across participants, the study had a 2 (agent color) × 3 (proposition type: powerful vs. powerless vs. horizontal) design with repeated measures on the second factor. The following hypotheses were tested: Both powerful and powerless propositions were expected to be vertical rather than horizontal; that is, the angle of a powerful proposition should be larger than both 45° and the horizontal proposition’s angle, and the angle of a powerless proposition should be lower than –45° and the horizontal proposition’s angle.

#### Participants

Students of an introductory psychology course at the University of Jena (Jena, Germany) answered the questionnaire voluntarily. After deleting one

case with missing values, 78 participants, 12 of them male, remained in the sample. The mean age was 21.2 years ( $SD = 2.2$ ).

### Materials and Procedure

Each proposition (see Appendix), with the agent as a black circle and the patient as a white circle (or reversed), was depicted above a picture with the eight alternative positions (shown in Figure 1). On the paper-and-pencil questionnaire, participants were instructed to mark for each of the 18 propositions the picture that best fit their image of the proposition. The order of propositions was determined by random and kept constant across participants. At the end, participants indicated their age and gender and were debriefed and thanked.

### Results

Figure 2 provides an overview of the data by showing the frequencies of each of the eight angles summed separately for powerful, powerless, and horizontal propositions. The figure shows that for both powerful and powerless propositions, primarily vertical angles were chosen, whereas horizontal propositions were indeed primarily horizontal in their angles.

To compute the angle of the propositions, it was necessary to ignore whether the agent was left or right of the patient. Each answer was scored as seen in Figure 1: Horizontal angles were scored as 0°, vertical angles with the agent on top were scored as 90°, vertical angles with the agent at the bottom were scored as –90°, and the diagonal lines were scored as in between. Then, angles were averaged for each cell of the Proposition Type × Agent Color design. Collapsed across agent color, powerful propositions had a mean angle of 65.6° ( $SD = 17.7$ ); that is, the angle was vertical, with the agent over the patient. A one-sample  $t$  test confirmed that it was significantly larger than 45°,  $t(77) = 10.29$ ,  $p < .001$ . In contrast, powerless propositions had an angle of –56.4° ( $SD = 16.4$ ); that is, the angle was also vertical, but the agent was below the patient. This angle was significantly smaller than –45°,  $t(77) = 6.15$ ,  $p < .001$ . The horizontal propositions were indeed horizontal ( $M = 0.1^\circ$ ,  $SD = 15.0$ ).

The angles were submitted to a 2 (agent color) × 3 (proposition type) general linear model (GLM) with repeated measures on the second factor. Proposition type had a significant effect,  $F(2, 152) = 926.90$ ,  $p < .001$ ,  $\eta_p^2 = .92$ .<sup>1</sup> Both powerful and powerless proposition angles differed significantly from the horizontal propositions angle,  $F(1, 76) = 510.13$ ,  $p < .001$ ,  $\eta_p^2 = .87$ , and  $F(1, 76) = 598.61$ ,  $p < .001$ ,  $\eta_p^2 = .89$ , respectively. No other effect reached significance.<sup>2</sup>

### Discussion

Study 1 confirmed that there is a shared metaphor that links power relations to a vertical schema, in which the powerful agent is on top of the powerless one. At first sight, this result might seem

<sup>1</sup>  $\eta_p^2$  denotes the effect size estimate partial  $\eta$ -squared as computed by SPSS Version 12.

<sup>2</sup> Figure 2 also suggests that powerful agents are more frequently located left of powerless patients and that powerless agents are more frequently located right of powerful patients. It may be that power is also represented as left of powerlessness, which might be a more general case of a schema of causation that flows from left to right (Maass & Russo, 2003).

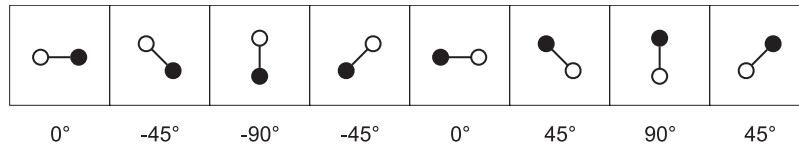


Figure 1. Depictions of eight different angles between agent and patient, and angle values scored for each picture, as used in Study 1 (values were not shown in the original questionnaires).

trivial. But in fact, power is defined as having influence on others' outcomes, although the results support the idea that in addition to this definition, it is understood as a vertical schema, or in the form of a power = up perceptual symbol. Of course, the results do not allow strong conclusions about the mental representations. It could be that the participants merely succeeded in applying a metaphor and aligning it with the graphical representation but that this schema is not usually evoked when one thinks about power. Study 2 was designed to collect more conclusive evidence that power involves the perceptual simulation of vertical difference.

### Study 2: Vertical Difference Between Two Group Names Influences Power Judgments

How spontaneously is this power = up schema used, and how deeply are thinking about power and thinking about vertical spatial difference intertwined? In other words, is the power = up schema indeed a perceptual symbol or just a metaphor that can be reproduced when asked for (as in Study 1)? One way to show that it is indeed a perceptual symbol is to show that thinking about power interacts with available perceptual content on the vertical dimension. This follows directly from the interference paradigm studies cited above, in which compatibility between motor schema and cognitive tasks (Förster & Strack, 1996) or between visual schema and cognitive tasks (Zwaan & Yaxley, 2003) served as evidence for the inclusion of perceptual symbols in the cognitive representation.

In Study 2, I implemented such an interference paradigm for the verticality of power. To do so, I adapted the task developed by Zwaan and Yaxley (2003). In each trial of a reaction time task, participants saw two group labels together on the screen (e.g., *master* and *servant*). The group pairs were selected such that the

two groups were related and that a clear power difference between the groups existed. An additional pretest made sure that high power was not confounded with positive valence. For each pair, participants had to decide as quickly as possible which of the two groups was the more powerful one. The crucial point was that the groups were presented above each other on the screen, such that the powerful one was at either the top or the bottom. It was expected that participants would more quickly identify a powerful group when it was at the top rather than when it was at the bottom of a pair. As this confounds the answering key with power, in a second condition, participants had to find the less powerful group. Here, it was expected that participants would more quickly identify the powerless group when it was at the bottom, compared with when it was at the top. In sum, an interaction of the group type and the vertical position was predicted.

### Method

#### Overview and Design

In each trial, participants saw labels of two social groups above each other on a screen and had to decide either which one was the powerful or which one was the powerless group; this was manipulated between participants and formed the first factor task. Each of the pairs was presented twice, once with the powerful group at the top and once with the powerful group at the bottom. This formed the second factor position (top vs. bottom), which was manipulated within subjects. Answers had to be given with the cursor up and down keys. Thus, whatever group the participants had to find, they had to press the up key when it was at the top and the down key when it was at the bottom. In sum, the experiment had a 2 (task: find powerful vs. find powerless, between) × 2 (position: top vs. bottom, within) design.

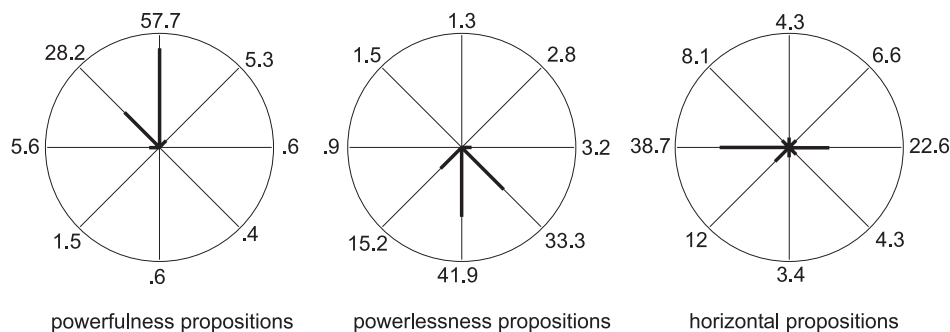


Figure 2. Frequencies and percentages of angles chosen for propositions with powerful and powerless agents, and horizontal relations (Study 1). Each line's length depicts frequency of its angle if the patient is at the center of the circle and the agent is at the outer end of the line. Percentages are given in the corresponding numbers. The outer circle corresponds to 65%.

## Participants

Ninety-one participants took part in the study. Eleven of them were excluded because they answered wrong or too slowly in more than 10 trials, leaving 80 in the sample. Of these, 47 were female and 28 were male (missing data for 5 cases). The average age was 21.5 years ( $SD = 2.7$ ).

## Materials

Altogether, 24 pairs of social groups were used. For each pair, a pretest had shown that the one group was almost unanimously judged to be more powerful than the other group (pairs are listed in the Appendix). In a further pretest, 34 participants were presented with the group pairs and were asked to decide which of the two groups they liked more on a scale from 1 (*more liking for the powerful group*) to 5 (*more liking for the powerless group*).<sup>3</sup> The mean of the averaged ratings ( $M = 3.21$ ,  $SD = .29$ ) differed significantly from the scale midpoint (3),  $t(33) = 4.23$ ,  $p < .001$ , indicating that the powerful groups were on average liked less than the powerless groups.

In the reaction time task, each pair was presented twice, resulting in 48 trials. Order of presentation was randomized, and it was counterbalanced for each pair whether it first appeared with the powerful at the top or at the bottom. Each trial started with a blank screen for 750 ms and a fixation cross (+) in the middle of the screen, which disappeared after 250 ms, followed by a pair of groups above each other, centered vertically and horizontally on the screen, with five blank lines between them. (Thus, the fixation cross was located directly in the middle between the two labels.) When no answer was given after 2 s, the program continued. Feedback on too long or wrong answers was not given. Words appeared in black 10-point Arial letters on a white background. The study was run on laptops with 14-in. displays (resolution  $1024 \times 768$ ) and programmed in DMDX (Forster & Forster, 2003). The laptops stood on tables, at which the participants were seated. Thus, the participants had to look down at the laptop screen.

## Procedure

Participants were approached at the campus of the University of Jena and asked to participate in a computer study in exchange for chocolate (a value of about \$1). When they agreed, they completed the experiment in groups of up to 3 persons in cubicles set up at the campus. The first screen informed them that the study investigated reactions to verbal stimuli, explained the task to which they were randomly assigned (find powerful vs. find powerless), and asked them to work both as quickly and as accurately as possible. The instructions simply asked participants to find the more powerful (or powerless) group, without defining power more explicitly. (None of the participants complained about a lack of clarity.) After they completed the task, they were debriefed, thanked, and given their chocolate.

## Results

### Response Latencies

The grand mean of all response latencies was 1,095 ms. Following the recommendations of Bargh and Chartrand (2000), I planned to exclude all response latencies longer than three standard deviations above the mean; however, this criterion equaled 1,980 ms in this study, probably because each trial was terminated automatically after 2,000 ms anyway. Consequently, no response latencies were excluded in this study, and the maximum response latency was increased in the following studies.

Latencies of responses in which the to-be-found group was at the top were averaged to one score, and reactions in which the to-be-found group was at the bottom were averaged to a second

score. Note that top answers were required when the task was to find the powerful group and it was at the top (compatible trials), but they were also required when the task was to find the powerless group and it was at the top (incompatible trials). Likewise, bottom answers were required when the task was to find the powerless group and it was at the bottom (compatible trials), but they were also required when the task was to find the powerful group and it was at the bottom (incompatible trials). It was hypothesized that for both tasks, the reactions in the compatible trials would be faster than in the incompatible trials.

To test this, the two scores were entered into a 2 (task: find powerful vs. find powerless)  $\times$  2 (position: top vs. bottom) GLM with repeated measures on the second factor. The means in Table 1 show the expected pattern: When the task was to find the powerful group, reactions were indeed faster when it was at the top, compared with when it was at the bottom. Simple effects analyses confirmed that this difference was significant,  $F(1, 78) = 11.91$ ,  $p = .001$ ,  $\eta_p^2 = .13$ . The opposite was the case when the task was to find the powerless group. Here, reactions were faster when it was at the bottom, compared with when it was at the top. This difference was less pronounced but still significant,  $F(1, 78) = 4.12$ ,  $p = .046$ ,  $\eta_p^2 = .05$ . Together, these differences resulted in a significant interaction,  $F(1, 78) = 14.82$ ,  $p < .001$ ,  $\eta_p^2 = .16$ . In addition, there was a marginal main effect of task: Answers were slower when the task was to find the powerless group than when it was to find the powerful group,  $F(1, 78) = 3.71$ ,  $p = .058$ ,  $\eta_p^2 = .05$ .

### Error Frequencies

The same analytic procedure was repeated for the number of errors (not counting too long or not given answers). Each participant committed on average 2.95 ( $SD = 2.09$ ), or 6.1%, errors. Although the means showed the predicted pattern, a GLM following the above design found no Task  $\times$  Position interaction,  $F(1, 78) = 0.95$ ,  $p = .330$ ,  $\eta_p^2 = .01$ .

## Discussion

The goal of Study 2 was to test whether vertical differences actually play a role in mental representations of power and thereby influence power judgments. To test this, participants had to decide which of two groups presented together on the screen was, in the first condition, the more powerful or, in the second condition, the powerless group. Furthermore, whether the group that was to be found was above or below the other group varied. The results show that this position on the screen influenced how quickly the task could be solved. Finding the powerful group was faster when it was at the top than when it was at the bottom. Finding the powerless group was faster when it was at the bottom than when it was at the top. In sum, answers were facilitated when the groups were where the metaphor suggests: the powerful at the top and the powerless at the bottom.

The simple effects analyses showed that the effect was somewhat stronger when the powerful group had to be found compared

<sup>3</sup> The original German question asked which of the two groups participants had more "Sympathie" for. This term differs in its meaning from the English *sympathy* and is best translated as *liking*.

Table 1  
*Mean Response Latencies (in ms) to Find the Powerless or Powerful Group, Depending on Its Position on Screen (Study 2)*

Task	Position of target group on screen			
	Top		Bottom	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Find powerful	1050	134	1093	138
Find powerless	1140	142	1113	124

with the condition when the powerless group had to be found. However, this effect has to be interpreted with caution because the answers use different keys: the up key when the powerful group was at the top and the down key when the powerful group was at the bottom. In principle, it is possible that the up key can be hit faster or that the group appearing at the top can always be found more quickly because it is read first. These two reasons might have resulted in a slight overall increase in the speed of top answers and might thus have distorted the simple comparisons. The interaction itself, which is not affected by these concerns, should be interpreted as the main result of the study. A curious and unexpected result was that reactions of participants whose task was to find the powerful group were faster; this will be discussed later.

Independent ratings of the used group pairs indicated that the powerful groups were liked less than the powerless groups, although the reaction time data showed that reactions were faster when they were at the top. This makes it unlikely that the pattern found in Study 2 is due to the confounding of high power with positive valence. Note, however, that there was another, deliberate, confound: The paradigm manipulated the position of the key used for the answer and the position on the screen in parallel: The two keys used for the answer were the cursor-up and the cursor-down keys, which for regular computer users are associated with up and down movements, in addition to the position of the target group, which varied between top and bottom. Thus, it is impossible to say whether the association of the response keys with a vertical movement (i.e., it is easier to hit the up key for a powerful group), or the visual representation (i.e., the powerful group is where it is expected to be), or a combination of the two led to the pattern observed in Study 2. The data from Borghi et al. (2003) cited above suggest that the motor schema alone can result in this pattern, whereas the data from Zwaan and Yaxley (2003) and the strong effects of Study 1 make it plausible that visual representation alone might also be sufficient. These two hypotheses will be tested in Studies 3 and 4, respectively.

### Study 3: Motor Responses Associated With a Vertical Difference Influence Power Judgments

Study 2 found that powerful groups are judged more quickly as powerful when they are at the top and when this is acknowledged with an up movement (pressing the up key on the keyboard). The question is whether the effect is due only to the visual component (being up) or whether the motor component alone can influence thinking about power.

As cited above, evidence from Borghi et al. (2003) that supports this hypothesis shows that judgments about concrete objects are influenced by the motor response necessary to make the judgment. Apparently, imagining an object is facilitated by compatible movement and interfered with by an incompatible movement. Given the parallel evidence for thinking about concrete objects and metaphorical mappings cited so far, it seems likely that thinking about the abstract domain of power is also influenced by vertical movements. This hypothesis is also compatible with the theory of event coding (Hommel, Müsseler, Aschersleben, & Prinz, 2001), which argues that both perception and action codes are represented at a common, commensurable level—that of external events. Up and down motor responses would therefore be represented as vertical movements of an external object (hand or the affected object) and thus at the same level as external visual stimuli (for a similar line of reasoning, see Neumann & Strack, 2000).

The up and down cursor keys used in Study 2 are not actually above each other in space. Thus, it would be more precise to say that they imply different imagined movements or motor imagery of an up and down movement. Given most students' extensive use of computer software in which a press of the up cursor results in an up movement of a cursor or another object, it seems likely that the movement is actually represented as upward. Study 3 therefore tested whether judgments of groups' power could be made more quickly when a compatible movement was necessary to make the judgment. To do so, the second component of the combined manipulation of Study 2—namely, using the up and down cursor keys on a regular computer keyboard—was isolated and manipulated without confounding it with visual representation. To design a simple task, I judged only one group at a time for its power. If powerful groups are represented as vertically on top, as the powerful = up perceptual symbol hypothesis postulates, then movements that imply an upward movement will facilitate the judgment as powerful and interfere with judgments as powerless. The reverse is hypothesized for downward movements.

As in the previous study, pretest data were collected to show that high power was not confounded with valence. Because the one-item measure from Study 2 might be questioned because it referred only to likability, a multiple-item measure was adopted.

## Method

### Overview and Design

Participants judged groups as powerful or powerless in a reaction time task. This was the first factor, Group Status, which was varied within participants. To make judgments, participants used the up cursor key and the down cursor key on their keyboard. In two consecutive blocks, each of the two keys was paired with each of the two judgments: In one block, judgments as powerful had to be given with the up key, and judgments as powerless had to be given with the down key (the compatible block). In the other block, judgments as powerful had to be given with the down key, and judgments as powerless had to be given with the up key (the incompatible block). This formed the second factor, compatibility, which was also varied within participants. The order of these blocks was counterbalanced. This resulted in a 2 (block order: compatible vs. incompatible block first, between)  $\times$  2 (group status: powerful vs. powerless, within)  $\times$  2 (compatibility: compatible vs. not compatible, within) design.

## Participants

From the total sample of 59 participants, 9 were excluded: 1 because she was not a native speaker of German, 1 because of more than 15 wrong or too slow answers, and 7 because they guessed the hypothesis of the study correctly.<sup>4</sup> Of the remaining 50 participants, 30 were female; the average age was 24.4 years ( $SD = 3.1$ ).

## Materials

Participants judged single groups. For both powerless and powerful group status, 8 groups were selected from the Study 2 stimuli, and 8 new groups were generated (see Appendix). This resulted in a total of 16 for each group status condition, or 32 in sum, plus 5 practice items that appeared only once in the beginning. Groups were selected such that it was as clear as possible that a group was powerful even in the absence of a comparison group. To ensure that this was the case, an independent sample ( $N = 24$ ) rated all groups on six 7-point bipolar scales (anchored from 1 to 7). The scales were constructed following Aronoff, Barclay, and Stevenson (1988) and Aronoff, Woike, and Hyman (1992). Power was measured on two scales, with the anchors *powerful–powerless* and *strong–weak*. Furthermore, valence was measured on four scales, with the anchors *positive–negative*, *good–bad*, *pleasant–unpleasant*, and *kind–unkind*. The two power items correlated well across targets, average correlation  $r = .46$ . Likewise, the valence scale had a satisfying internal consistency, with an average of .70. Thus, average scores for power and valence were computed separately for powerful and powerless target groups. For ease of presentation, the scores were reversed; high values thus indicate high power and positive valence. A paired-samples  $t$  test on the power ratings confirmed that powerful groups were judged as more powerful ( $M = 5.62$ ,  $SD = .74$ ) than powerless groups ( $M = 2.84$ ,  $SD = .54$ ),  $t(23) = 11.93$ ,  $p < .001$ . Furthermore, powerful groups did not differ from powerless groups in their valence ( $M = 4.03$ ,  $SD = .56$ ; and  $M = 4.01$ ,  $SD = .49$ , respectively),  $t(23) = .14$ ,  $p = .888$ .

In the reaction time task, all group labels were judged twice, once in the compatible block and once in the incompatible block. Each block was preceded by an instruction explaining the key assignment and five practice items. Each trial began with a blank screen (500 ms), followed by a horizontally and vertically centered fixation cross (300 ms) and then the group label, appearing at the place of the fixation cross, in 10-point Arial font. Group labels remained until an answer was given but not longer than 3 s. Hardware and software were identical to that used in Study 2.

## Procedure

The main procedure was identical to that used in Study 2. After the reaction time task, participants were asked to write down what they thought the purpose of the study was.

## Results

### Response Latencies

Response times more than three standard deviations longer than the grand mean ( $M = 801$ ,  $SD = 292$ ) and wrong response times were excluded (6.13%). To ease comparison with the other studies, latencies for up and down responses were averaged. The following four average response times were computed for each participant: judgments of both powerful and powerless groups, each made once with the up response key and once with the down response key. Means are represented in Table 2.

These four scores were entered into a 2 (block order)  $\times$  2 (group status)  $\times$  2 (response) mixed-model GLM with repeated measures on the last two factors. The predicted interaction of Response and

Table 2  
Average Response Latencies (in ms) and Average Number of Errors (and SDs) for Power Judgments of Groups, Depending on Group Status and Response Key (Study 3)

Group status	Response key			
	Up		Down	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Response latencies				
Powerful	707	114	776	161
Powerless	836	161	793	131
Error frequencies				
Powerful	0.30	0.65	0.64	0.94
Powerless	1.02	1.44	0.64	0.85

Group Status was significant,  $F(1, 48) = 17.09$ ,  $p < .001$ ,  $\eta_p^2 = .26$ . Judgments of powerful groups were significantly faster when made with the up key than when made with the down key,  $F(1, 48) = 18.76$ ,  $p < .001$ ,  $\eta_p^2 = .28$ . Conversely, judgments of powerless groups were significantly slower when made with the up key than when made with the down key,  $F(1, 48) = 8.79$ ,  $p = .005$ ,  $\eta_p^2 = .16$ .

The analysis revealed a number of additional significant effects. Judgments of powerless groups were again slower than judgments of powerful groups,  $F(1, 48) = 49.43$ ,  $p < .001$ ,  $\eta_p^2 = .51$ . Furthermore, there were significant interactions of Block Order  $\times$  Response Key,  $F(1, 48) = 5.98$ ,  $p = .018$ ,  $\eta_p^2 = .11$ ; and of Block Order  $\times$  Response Key  $\times$  Group Status,  $F(1, 48) = 44.56$ ,  $p < .001$ ,  $\eta_p^2 = .48$ . This three-way interaction needs further exploration. To make it easier, I computed difference scores by subtracting average response latencies in compatible blocks from those in incompatible blocks. This was done for both powerful and powerless groups, resulting in two difference scores. These can be interpreted as compatibility effects, similar to other interference paradigms (e.g., the implicit association test effect; Greenwald, McGhee, & Schwartz, 1998). The two scores were entered into a 2 (block order)  $\times$  2 (group status) GLM with repeated measures on the second factor. Replicating the three-way interaction above, I found that Block Order had a significant effect,  $F(1, 48) = 44.56$ ,  $p < .001$ ,  $\eta_p^2 = .48$ . This main effect was further qualified by an interaction with Group Status,  $F(1, 48) = 5.98$ ,  $p = .018$ ,  $\eta_p^2 = .11$ . The pattern was as follows: When the incompatible trials came first, the difference between incompatible and compatible blocks was positive for both powerful and powerless targets ( $M = 160$ ,  $SD = 137$ ; and  $M = 177$ ,  $SD = 127$ , respectively), whereas it was negative when the compatible trials came first for both powerful and powerless targets ( $M = -7$ ,  $SD = 111$ ; and  $M = -71$ ,  $SD = 124$ , respectively). The interaction indicates that this difference was larger for powerful target groups. Does this mean that the hypothesis was supported only when the incompatible trials came

<sup>4</sup> Note, though, that the unconsciousness of the hypothesized effect is not a critical claim. Identical results were found when these participants were retained in the analyses.



first? It does not, because it has to be kept in mind that all group labels had to be judged twice and that the second judgment of a group label is probably easier and therefore faster. Thus, when the incompatible trials came first, the computed differences are actually the sum of compatibility and learning: The second block was easier because it was compatible and participants knew the group labels already. In contrast, when the compatible trials came first, the difference scores are the difference between compatibility and learning: The first block was easier because it was compatible, but the interference by the incompatible key assignment was obscured by the fact that the group labels were now easier to judge because they were already familiar. The negative difference in the second case reveals that the learning effect was stronger than the compatibility effect itself, and the interaction with group status shows that this was less so for the powerful targets.

### *Error Frequencies*

On average, participants committed 2.6 ( $SD = 2.52$ ) errors (not counting not given or too long answers), or 4.1%. Errors were summed for the same four categories as used in the response latencies analysis and submitted to the GLM. Table 2 shows that judgments of powerful groups were more accurate when the necessary key was up rather than down,  $F(1, 48) = 4.37, p = .042, \eta_p^2 = .08$ , whereas judgments of powerless groups were marginally less accurate when the necessary key was up rather than down,  $F(1, 48) = 3.33, p = .074, \eta_p^2 = .07$ . This resulted in a significant Response  $\times$  Group Status interaction,  $F(1, 48) = 6.48, p = .014, \eta_p^2 = .12$ . In addition, judgments of powerless groups were less accurate than those of powerful groups,  $F(1, 48) = 9.24, p = .004, \eta_p^2 = .16$ .

### *Discussion*

Results of Study 3 suggest that the judgment of a group as powerful or powerless is easier when the motor response used to answer fits the perceptual symbol of powerful groups as up and powerless groups as down. Judgments of a group as powerful were faster and more accurate when the up cursor key had to be used for the answer than when the down cursor key had to be used. The reverse was true for judgments of a group as powerless: Here, answers were faster and more accurate when the down cursor key had to be used for the answer, compared with when the up cursor key had to be used.

In line with the theory of event coding (Hommel et al., 2001), I interpret this result as showing that action codes (up and down) are mentally represented as external events (top vs. bottom) and that these external event representations on the vertical spatial dimension interact with the power = up perceptual symbol. When a powerful group is perceived and the judgment has to be made, the vertical position top is activated. If the necessary movement is up, it fits the activated vertical position, and the response is facilitated. If the necessary movement is down, it does not fit the activated vertical position, and thus it interferes with the response. The reverse applies for judgments of powerless groups.

Compatibility had an effect not only on judgments of powerful groups but also on judgments of powerless groups. However, in line with Study 2, the effect was smaller for judgments of powerless groups. Second, the present results again show a main effect

of group status such that judgments of powerful groups were faster than judgments of powerless groups. Both findings are consistent across the studies and will be discussed later in more detail. As in Study 2, it is unlikely that the effect is due to the confounding of high power with positive valence, because a pretest found no valence difference between high and low power groups.

It is noteworthy that Study 3 found significant effects not only on response latencies but also on accuracy. As suggested above, effects on accuracy are most likely when the judgment is made difficult by external factors. The external factor at work in Study 3 might have been the blockwise manipulation of the compatible and incompatible response key assignment, in contrast to the randomized order of compatibility and incompatibility in the previous study. Such a blocking might make it especially easy to judge in the compatible condition and especially hard when it must be unlearned, leading to stronger effects on accuracy (although the total number of errors was not markedly increased).

### Study 4: Vertical Position of a Group Influences Power Judgments

With the term *perceptual symbol*, Barsalou (1999) denoted not only modal representations in the narrow meaning of the word *perceptual* (e.g., visual, auditory, tactile), but also motor representations. Moreover, Glenberg (1997) argued that motor representations or patterns of possible actions are primary and central for conceptual thinking. Combining visual representations and motor response, as Study 2 did, is consistent with this previous theorizing. However, showing that the visual experience alone is schematized into a perceptual symbol of power would make an even stronger argument for the perceptual symbol hypothesis. Thus, Study 4 tried to demonstrate that visual spatial input alone can influence judgments of power. It followed the paradigm developed in Study 2, but with several modifications. First, as in Study 3, single groups had to be judged. Second, the answer format was changed such that the answering keys did not replicate the visual representation but were horizontal to each other. In addition, to demonstrate clearly that the effects were independent of valence, every participant rated all target groups concerning their valence and power.

### *Method*

#### *Overview and Design*

In each trial of a reaction time task, participants had to decide whether the presented group was powerful or powerless. Sixteen powerful groups and 16 powerless groups were presented; this formed the first factor, Group Status, which varied within participants. Each of the groups had to be judged twice, once when it appeared in an upper position on the screen and once when it appeared in a lower position. This formed the second factor, Position (top vs. bottom), varied within participants. To make the judgments, participants used the left and the right cursor keys, which are not above each other. Assignment of the keys to the judgments (powerful vs. powerless) was counterbalanced and formed the third factor, Key Assignment. Thus, the study had a 2 (key assignment, between)  $\times$  2 (group status: powerful vs. powerless, within)  $\times$  2 (position: up vs. down, within) design. In addition, awareness of the metaphor being tested in the study was assessed.

### Participants

In total, 44 participants took part in the study. Three of them never answered in less than 3 s (probably due to a misunderstanding about the assigned keys) and were excluded, as was 1 additional participant with 56 errors (all others had fewer than 15 errors). Thirteen of the final 40 participants were male; the mean age was 22 years ( $SD = 3.0$ ). None of them guessed what the purpose of the study was.

### Materials

Participants judged the same groups as in Study 3, except that the item *officer* was replaced by *warder*. Order of appearance was randomized, and whether a group label appeared first in the top or bottom position was counterbalanced across participants. Each trial started with a blank screen (750 ms) and a fixation cross (+) in the middle of the screen (400 ms). Then, the cross disappeared and a group label appeared five lines above or below the place where the fixation cross was located. Labels were set in black 10-point Arial font on a white background. Group labels remained until an answer was given, but not more than 3 s. Hardware and software were identical to that used in Study 2.

### Procedure

The procedure was identical to that used in Study 2. In addition, after they completed the reaction time task, participants were asked to rate each of the 32 groups for how much they liked the group and how much power it had. Both scales ranged from 1 (*low power, low liking*) to 5 (*high power, high liking*).

## Results

### Manipulation Checks

Data from 2 participants are missing for the ratings of each group's power and valence. The power ratings served as a manipulation check and confirmed that powerful groups were rated as more powerful ( $M = 4.25$ ,  $SD = .32$ ) than powerless groups ( $M = 1.99$ ,  $SD = .42$ ),  $t(37) = 24.17$ ,  $p < .001$ . The valence ratings showed that powerful groups were rated as less positive ( $M = 2.94$ ,  $SD = .47$ ) than powerless groups ( $M = 3.45$ ,  $SD = .40$ ). This difference was also significant,  $t(37) = 5.17$ ,  $p < .001$ .

### Response Latencies

Latencies belonging to wrong answers or those three standard deviations longer than the grand mean ( $M = 853$ ,  $SD = 249$ ) were excluded from the analyses (7.5%). Then, reaction times were averaged separately for powerful and powerless groups appearing in the upper or lower position, resulting in four scores (see Table 2). These four scores were then submitted to a 2 (key assignment)  $\times$  2 (group status)  $\times$  2 (position) mixed-model GLM with repeated measures on the last two factors. This analysis revealed the predicted interaction of Group Status and Position on the screen,  $F(1, 38) = 4.81$ ,  $p = .034$ ,  $\eta_p^2 = .11$ .

Table 3 shows that judgments of powerful groups as powerful were faster when the groups appeared at the top position compared with when they appeared at the bottom position. This difference was significant,  $F(1, 38) = 16.96$ ,  $p < .001$ ,  $\eta_p^2 = .31$ . On a descriptive level, and contrary to expectations, judgments of powerless groups as powerless were also faster when they appeared at the top position compared with when they appeared at the bottom

Table 3  
Mean Response Latencies (in ms) and Standard Deviations for Power Judgments of Groups, Depending on Group Status and Position on Screen (Study 4)

Group status	Position of target group on screen			
	Top		Bottom	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Powerful	785	85	825	70
Powerless	871	96	881	93

position, but this difference was not significant,  $F(1, 38) = 1.80$ ,  $p = .188$ ,  $\eta_p^2 = .045$ . In addition, main effects of Group Status,  $F(1, 38) = 74.10$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , and of Position,  $F(1, 38) = 14.67$ ,  $p < .001$ ,  $\eta_p^2 = .28$ , were found. None of the other effects was significant.

### Error Frequencies

Errors (not counting ungiven or too long answers) were summed for the same four categories and submitted to the same analysis. On average, participants committed 3.43 errors ( $SD = 2.54$ ), or 5.4%. In line with the latencies data, judgments of powerful groups were more correct and judgments of powerless groups were less correct when they appeared at the top, but this interaction did not reach significance,  $F(1, 38) = 2.21$ ,  $p = .145$ ,  $\eta_p^2 = .06$ .

## Discussion

Results of Study 4 suggest that the vertical spatial position of a group label can influence judgments of the group's power, even when effects of the response movement are excluded. Powerful groups were judged more quickly as powerful when they appeared in the upper part of the screen compared with when they appeared in the lower part of the screen. For powerless groups, there was no significant difference between the two spatial positions, resulting in a significant interaction. Note that this interaction is the main finding and more meaningful than the simple main effects. Nevertheless, it is noteworthy that for powerless groups, there was no compatibility effect, in contrast to Studies 2 and 3. It is possible that targets in the upper screen position are always judged more quickly, irrespective of their status, maybe because of reading habit or attention biases. This could have resulted in a main effect that distorted the interaction. However, it cannot be ruled out that the perceptual symbol is stronger for power = up than for powerless = down. It could also be that the symbol for powerlessness is not so much related to a down position as to a position at the middle or a "ground level" position. Because in the current studies only top and bottom positions or up and down movements were compared, this question is beyond the scope of the present data.

Two more details deserve notice: Powerless groups were again on average judged more slowly than powerful groups. The alternative valence account, namely, that powerful groups are at the top because they are more positive, can be ruled out on the basis of the manipulation checks, which confirmed that the more powerful groups are actually seen as less positive than the powerless groups.

### Study 5: The Task Moderates Whether Vertical Position Embodies Power or Valence

One impressive piece of evidence for the embodied representation of concepts comes from the previously cited results of Meier and Robinson (2004), who showed that positive valence is represented as up in space, whereas negative valence is down. Although this evidence supports the embodiment account in general, it also questions the current results by providing an alternative explanation for the presented results: In reality, valence and power may indeed be often confounded in their symbolization as up in space, as it is the case for gods and depictions of admired parental figures (e.g., statues). Is it possible that powerful groups are represented as up because they are evaluated positively? Studies 2, 3, and 4 already addressed this question by collecting explicit valence judgments of the group stimuli. Study 2 showed in a pretest, by using a single item measure of likability, that powerful groups were rated as less positive. Study 3 showed in a pretest, by using multiple items, that powerful and powerless groups were rated as having equal valence. Study 4 showed in a manipulation check, by using the single likability item, again that powerful groups were rated as less positive. Nevertheless, although these data did not support the assumption that powerful groups are evaluated more positively, it could be that explicit assessments fail to capture an implicit valence that might drive the effects on power judgments—participants might be unwilling to publicly express their positive evaluation of powerful groups, although they implicitly favor them, or their negative evaluation of powerless groups, although they implicitly dislike them, just like the well-known difference between implicit and explicit prejudice (Devine, 1989). To address this alternative explanation more conclusively, it is necessary to pit power and valence against each other experimentally and to test which dimension drives the effect of interest here.

To ensure a fair test of the valence account, I had to choose clearly positive and negative target groups. In the following study, this was ensured by adopting a stimulus set that had been assembled for a study on evaluative priming by Crusius and Wentura (2005). This set comprised 16 positive and 16 negative target groups. Furthermore, this set took into account an important dimension for valence of social groups—namely, who is profiting from it. Peeters and Czapinski (1990) and Wentura, Rothermund, and Bak (2000) have pointed out that there are two questions for which the valence of a social actor can be judged: “Is it good or bad for me that Person X possesses the Characteristic Y? . . . and Is it good or bad for Person X him- or herself to possess the Characteristic Y?” (Wentura et al., 2000, p. 1024). The first type of valence is called *other-relevant*; the second type is called *self-relevant*. Acknowledging this important difference, I included 8 primarily self-relevant and 8 primarily other-relevant groups for each valence in the stimulus set (see Appendix). Pretests by Crusius and Wentura (2005) ensured the categorization of groups into valences and relevance types.

Further pretests (see below) revealed that groups that had a negative self-relevant valence were typically judged as powerless, whereas all other groups were typically judged as powerful. Thus, the stimulus set allowed a test of the power hypothesis. As in Study 4, in the present study the target groups were presented at the top and at the bottom of the screen, and participants had to judge their power. If the previous results were driven by a confounding of

valence and power (the valence account), judgments of negative groups would be facilitated when they were at the bottom, and judgments of positive groups would be facilitated when they were at the top, irrespective of relevance. If the power hypothesis is correct, different predictions follow: Only judgments of self-relevant negative groups would be facilitated when they were at the bottom, but judgments of all other groups would be facilitated when they were at the top. Of central interest is the distinction between negative self-relevant and negative other-relevant groups, because their valence is identical, but the former are powerless and the latter are powerful, thus pitting valence and power against each other.<sup>5</sup> Thus, Study 5a tested whether the power or the valence of the groups included in the stimulus set would drive interference effects of vertical position on judgments of power.

The current argument is that because power is mentally represented as a vertical dimension, thinking about power involves a mental simulation of space and will be influenced by vertical spatial information. Meier and Robinson (2004) proposed the same for valence. When one acknowledges that associations between space and a concept can be acquired for both dimensions, these two hypotheses are not at all mutually exclusive. In contrast, it seems that space can very well serve as a dimension for both concepts. The task itself, that is which concept is currently mentally simulated, should moderate which feature of a given stimulus—power or valence—is used to construct the mental simulation. If thinking focuses on power, a powerful group should be construed as up, even if it is negative. If thinking focuses on valence, however, the same group, if it is negative, should be construed as down. To integrate the current results with the findings on spatial representation of valence, Study 5b repeated the same study but tested interference effects of spatial position on an evaluative decision, thus attempting to replicate the finding by Meier and Robinson (2004) for the present stimuli.<sup>6</sup>

## Method

### Overview and Design

*Study 5a.* In a reaction time task, both negative and positive group names were presented on the screen. Their valence was either self-relevant or other-relevant, and they were presented at either the top or the bottom of the screen. In sum, the study had a 2 (valence: positive vs. negative, within)  $\times$  2 (relevance: self vs. other, within)  $\times$  2 (position: top vs. bottom, within) design. Participants decided whether the presented group was powerful or powerless.

*Study 5b.* The same study was repeated, but this time participants decided whether the presented group was positive or negative.

### Participants

*Study 5a.* After discarding data from nonnative speakers, data from 40 native speakers of German remained. One of them gave no answer at all,

<sup>5</sup> An ideal stimulus set would also include powerless but positive group labels. Because pretesting revealed that such a stimulus set could not easily be assembled, the existing stimulus set that was designed for test of valence effects was preferred.

<sup>6</sup> Meier and Robinson's (2004) stimulus set seemed to include both self-relevant (e.g., active, ambitious, leisure) and other-relevant (e.g., ethical, generous, loyal) stimuli. It is interesting to note that it also included power-related words like *power*, *hero*, *champion*, *victory*, and *defeat*.

leading to the exclusion of his data. Of the 39 participants remaining, 25 were female. The average age was 23.1 years ( $SD = 3.3$ ).

*Study 5b.* Data from 35 native speakers of German, 24 of them female, were collected and used in the analyses (and data from one nonnative speaker were discarded). The average age was 22.0 years ( $SD = 2.3$ ).

## Materials

*Target groups.* In a pretest, 31 participants (16 female, mean age = 24.2 years,  $SD = 3.4$ ) rated the 32 target groups concerning valence and power on six 7-point bipolar scales. The scales were the same as in Study 3 and were answered for all 32 target groups by every participant. The valence scale had a satisfying internal consistency, with an average alpha of .75 (averaged across target groups). Likewise, the two power items correlated well across targets, average correlation  $r = .50$ . Thus, two average scores for valence and power were computed for every target group. As in Study 3, both scores were reversed for ease of interpretation. Higher scores denote positive valence and powerful, respectively. To learn more about the valence and power characteristics of the stimulus set, I averaged ratings for all groups of each of the four types (Positive vs. Negative  $\times$  Self- vs. Other-Relevant), to gain one power and one valence score for each group type. On these scores, 2 (Valence)  $\times$  2 (Relevance) GLMs with repeated measures on both factors were conducted separately for Power and Valence.

The analysis on the Valence ratings confirmed that negative targets were indeed rated as more negative ( $M = 2.39$ ,  $SE = .08$ ) than positive targets ( $M = 5.78$ ,  $SE = .07$ ),  $F(1, 30) = 797.30$ ,  $p < .001$ ,  $\eta_p^2 = .96$ . The analysis also revealed that this difference was stronger for other-relevant targets than for self-relevant targets, resulting in a significant interaction,  $F(1, 30) = 111.20$ ,  $p < .001$ ,  $\eta_p^2 = .79$ , but it was significant for both nonetheless.

An interesting pattern emerged for the power ratings. There were significant main effects for both valence and relevance of the targets,  $F_s(1, 30) > 130.00$ ,  $ps < .001$ , both  $\eta_p^2 > .80$ , but these were qualified by a significant interaction,  $F(1, 30) = 310.30$ ,  $p < .001$ ,  $\eta_p^2 = .91$ . Only negative self-relevant targets were rated as powerless ( $M = 2.41$ ,  $SD = .48$ ), but negative other-relevant targets ( $M = 5.04$ ,  $SD = .67$ ), positive self-relevant targets ( $M = 5.25$ ,  $SD = .61$ ), and positive other-relevant targets ( $M = 5.07$ ,  $SD = .54$ ) were not. In fact, a look on the power ratings of the individual targets revealed that all negative self-relevant targets were rated as significantly less powerful than the midpoint of the scale and that all of the remaining targets were rated as more powerful than the midpoint of the scale. Thus, a contrast between the negative self-relevant targets and all other targets equals the distinction between powerless and powerful targets and will be used in the subsequent analyses of Study 5a.

*Reaction time tasks.* In Study 5a, the design of trials and the task instructions were identical to those used in Study 4. Ten practice trials with clearly powerful and powerless targets preceded the actual trials. In Study 5b, trial and trial design were identical to those used in Study 5a, but the participants were asked to judge whether the displayed groups were typically positive or negative. Ten practice trials with clearly positive and negative self- and other-relevant target groups preceded the actual trials. In both studies, feedback was given only during practice trials. Participants had to press the right key for powerful or positive groups and the left key for powerless or negative groups, respectively.

## Procedure

Procedures were the same for both Studies 5a and 5b and identical to Study 4, except that groups were not rated explicitly.

## Results

### Study 5a

*Response latencies.* The grand mean of all response latencies equaled 1,014 ms ( $SD = 383$ ). This grand mean is much longer than those in the previous studies in which only one group had to be judged (Studies 3 and 4). The long latencies and the large variance suggest that at least for some of the targets, the power judgments were rather difficult and time-consuming. After excluding latencies three standard deviations above the mean, I averaged response latencies separately for Valence, Relevance, and Position. A GLM with these factors as repeated measures revealed only three significant main effects but no interactions (all  $F_s < 1$ ). For a second analysis, response latencies to powerful groups (combining all positive and the negative other-relevant groups) were combined (see below). A 2 (status)  $\times$  2 (position) repeated-measures GLM revealed a main effect only of position but no interaction ( $F < 1$ ). Apparently, the long judgment latencies obscured any possible interference effects.

*Error frequencies.* Of all responses, 18% were errors; that is, the judgment differed from the average pretest judgment (not counting not given answers or too long answers). Of course, some of these errors probably just reflect that the participant did not agree with the average pretest sample. The crucial questions are whether these errors were influenced by the vertical position on the screen and whether such an influence is driven by the status or by the valence of the targets.

First, the data were analyzed with respect to the underlying valence and relevance factors. Errors (not counting ungiven answers) were counted separately for each of the four target types (positive and negative valence and self- and other-relevance) and whether these appeared at the top or at the bottom (see Table 4). The resulting eight scores were entered in a 2 (valence)  $\times$  2 (relevance)  $\times$  2 (position) repeated-measures GLM. If valence is the driving cause behind the effects found in the previous analysis, a Valence  $\times$  Position effect should be found. This interaction did not emerge ( $F < 1$ ). Instead, there was a weak Valence  $\times$  Position  $\times$  Relevance interaction,  $F(1, 38) = 2.44$ ,  $p = .127$ ,  $\eta_p^2 = .06$ . Simple effects analyses on the differences between top and bottom representation for each target type separately revealed more details. A top position led to more errors in the judgment of negative self-relevant targets,  $F(1, 38) = 4.17$ ,  $p = .048$ ,  $\eta_p^2 = .10$ , but to less errors in the judgment of negative other-relevant targets,  $F(1, 38) = 7.84$ ,  $p = .008$ ,  $\eta_p^2 = .17$ , and positive other-relevant targets,  $F(1, 38) = 4.26$ ,  $p = .046$ ,  $\eta_p^2 = .10$ . Vertical position had no effect on positive self-relevant targets ( $F < 1$ ).<sup>7</sup>

In sum, differences in valence cannot account for the effect of vertical position on accuracy of power judgment. But can the differences in power? To test this, I also analyzed the data according to the power of the target groups. Remember that all negative self-relevant groups were rated as powerless (and only these), whereas all others (including negative other-relevant groups) were rated as rather powerful. Following this categorization, four sep-

<sup>7</sup> In addition, there were two significant interactions that are theoretically irrelevant for the present purposes, one of Valence  $\times$  Relevance,  $F(1, 38) = 15.96$ ,  $p < .001$ ,  $\eta_p^2 = .30$ , and one of Position  $\times$  Relevance,  $F(1, 38) = 12.42$ ,  $p = .001$ ,  $\eta_p^2 = .25$ .

Table 4  
Average Number of Errors (and SDs) Made While Judging the Power of Groups, Depending on Their Valence and Self-Relevance and Their Vertical Position on the Screen (Study 5a)

Relevance	Vertical Position			
	Top		Bottom	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Negative				
Self (powerless)	0.62	1.07	0.38	0.82
Other (powerful)	3.15	2.41	3.67	2.41
Positive				
Self (powerful)	0.46	1.00	0.49	1.00
Other (powerful)	1.23	1.39	1.56	1.62
Powerful combined	1.62	1.13	1.91	1.20

Note. The power of each category is given in parentheses, and effects on powerful groups are combined in the last row.

arate scores were computed for powerful groups appearing at the top or the bottom and for powerless groups appearing at the top or the bottom. Note that this way of combining the data is equal to testing a +1 +1 -3 +1 contrast on the Valence (positive vs. negative)  $\times$  2 Relevance (self vs. other) design. To adjust for the number of targets, I divided the number of errors for the powerful targets by 3. These four scores were then submitted to a 2 (group status)  $\times$  2 (position) repeated-measures GLM. More errors were made in the judgment of powerful groups,  $F(1, 38) = 25.50, p < .001, \eta_p^2 = .40$ , but this main effect was moderated by an interaction of Position and Group Status,  $F(1, 38) = 11.07, p = .002, \eta_p^2 = .23$ . Table 4 shows that fewer errors occurred in the judgment of powerful groups when they were displayed at the top compared with when they were at the bottom,  $F(1, 38) = 8.14, p = .007, \eta_p^2 = .18$ . The reverse was true for powerless groups; here more errors occurred when they were displayed at the top compared with when they were at the bottom,  $F(1, 38) = 4.17, p = .048, \eta_p^2 = .10$ .

### Study 5b

*Response latencies.* The grand mean of all response latencies for the evaluative decision was much lower than for the power decision ( $M = 863, SD = 239$ ). Consequently, an analysis of the reaction-time data was possible. Response latencies three standard deviations above the mean were discarded. The remaining latencies were separately averaged to eight scores for their valence (positive vs. negative), the relevance of their valence (self vs. other), and their position (top vs. bottom) and were submitted to a GLM with these factors as repeated measures. This analysis yielded several significant effects. First, a main effect of position emerged,  $F(1, 34) = 4.97, p = .033, \eta_p^2 = .127$ . This main effect of position was qualified by an interaction with valence,  $F(1, 34) = 8.33, p = .007, \eta_p^2 = .20$ . Table 5 shows that although the valence of positive groups was evaluated more quickly when they appeared at the top than when they appeared at the bottom,  $F(1, 34) = 9.54, p = .004, \eta_p^2 = .22$ , there was no difference for

negative groups ( $F < 1$ ). The Position  $\times$  Valence interaction was not further qualified by relevance ( $F < 1$ ). Further analyses revealed that the top position decreased evaluation latencies for both self-relevant positive groups and other-relevant positive groups,  $F(1, 34) = 8.96, p = .005, \eta_p^2 = .21$ , and  $F(1, 34) = 4.42, p = .043, \eta_p^2 = .12$ , respectively. No differences were found for either type of negative group names (both  $F$ s  $< 1$ ). In addition, there was a marginal main effect of valence,  $F(1, 34) = 3.05, p = .090, \eta_p^2 = .08$ , and a theoretically less interesting interaction of valence and relevance,  $F(1, 34) = 22.48, p < .001, \eta_p^2 = .40$ .

*Error frequencies.* The average frequency of errors was low ( $M = 2.63, SD = 2.73$ ), or 4.1%. When the above analysis was repeated for error frequencies, no significant effects emerged besides a marginal main effect of valence,  $F(1, 34) = 3.32, p = .077, \eta_p^2 = .09$ , which indicated that there were slightly more errors for negative groups.

### Discussion

The main goal of Study 5 was to test whether differences in valence of powerful and powerless groups could account for interference effects of spatial information on power judgments. To test this, a stimulus set with extremely evaluated groups was used. Several findings rule out the possibility that valence is behind the effects on power judgments: First, valence did not predict effects of vertical position on power judgments, but power did. Second, within the negatively evaluated groups, there was a clear difference between those for which the negative valence was self-relevant and those for which the negative valence was other-relevant. Self-relevant negative groups, which in a pretest were rated as powerless, were judged more accurately as such when they were presented at the bottom. Other-relevant negative groups, which in a pretest were rated as powerful, were judged more accurately as such when they were presented at the top. Thus, the power of a group, which is implied by the relevance of its negative valence, and not its valence, determined effects of spatial positions on power judgments. Third, when groups were combined according to power instead of valence, there was a clear interaction of power and spatial position on accuracy.

However, the findings did not provide a full replication of the previous findings, because the predicted effects were found on accuracy rather than on response latencies, as in Study 4. The reason probably lies in the fact that the stimulus groups were selected for extreme valence to provide a fair test of the valence

Table 5  
Mean Response Latencies (in ms) and Standard Errors for Valence Judgments of Groups, Depending on Group Valence and Position on Screen, Collapsed Over the Factor Relevance (Study 5b)

Group valence	Position of target group on screen			
	Top		Bottom	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Positive	817	16	853	17
Negative	856	19	856	19

hypothesis. The side effect of this selection was that the power or powerlessness of the groups was less extreme and clear. This led to longer response latencies and more errors. In longer response latencies, any effects of automatic processes are obscured by other effects, leading to null results for this measure. Instead, the less clear statuses of the stimulus groups gave room for errors and thus allowed effects of the vertical position on accuracy. This reasoning is also supported by the pattern in Study 5b. For the evaluative decisions made there, the average response time was much lower, and there were much fewer errors. Consequently, the effects were found on the response times.

Going beyond power, these results of Study 5b show embodiment effects in the domain of valence as well. The same stimuli used in Study 5a showed markedly different interference effects with spatial position when participants had to judge valence, not power. Now, judgments of negative groups were facilitated when they were at the bottom regardless of their relevance, and judgments of positive groups were facilitated when they were at the top. These findings have interesting implications in two ways. First of all, they apply to the question of whether valence is an alternative explanation of the interference effects on power judgment by showing that there are indeed embodiment effects of valence but only when the decision is about valence as well. Thus, the findings provide indirect evidence for the validity of the stimulus selection because they show that the stimulus's valences were extreme enough to produce embodiment effects. But second, these findings show clearly that embodiment effects are due to the mental simulation process and thus to the judgment that has to be made. A vertical dimension in space is probably part of many different representations of abstract concepts besides power and valence—for instance, abstractness (Schnall & Clore, in press) and time for native speakers of Chinese (Boroditsky, 2001). Interference effects of spatial information can be expected when the task itself calls for a simulation of the concept or when the simulation of this dimension is automatically initiated by the stimuli themselves (e.g., in a relatedness decision for concrete well-known objects; Zwaan & Yaxley, 2003).

#### Meta-Analytic Combination of Response Times and Error Frequencies

For Studies 2 to 5b, I reported results concerning both response latencies and error frequencies. Whereas in Studies 2, 3, and 4, the pattern was always in the predicted direction for both latencies and errors, the effects were stronger for latencies, and, only for Study 3, a significant effect emerged for errors as well. In Studies 5a and 5b, an interesting pattern was revealed: In Study 5a, response latencies were much longer than in Studies 3 and 4, in which participants also had to judge single groups concerning their power (Study 2 is not comparable because two groups had to be judged at once). At the same time, participants made many more errors than before. This resulted in a null effect on response latencies and a significant effect on response errors. The likely reason is that in contrast to the previous studies, in which target groups were selected to have a clear and extreme status, in Study 5a, target groups were selected to have a clear and extreme valence. For Study 5b, this relation was reversed, because now responses had to be made for valence, the feature for which the groups were selected. The evaluative decisions were quick and accurate, lead-

ing to a significant effect on response latencies but to a null effect on errors. In sum, the picture suggested by these studies is that quick and accurate judgments can be accelerated or decelerated by vertical positions but only slightly influenced in their direction. The speed of slow and insecure judgments, on the other hand, cannot be influenced in a detectable way, but vertical positions of the group names will actually influence whether the groups are judged to be powerful or powerless.

Even though these relations among speed, accuracy, and effects of vertical position are already suggestive, it would be assuring to see whether the combination of both indices across studies provides a coherent picture. To combine and compare effects on the two indices (response latencies and error frequencies), I computed a meta-analysis. Using the formulas proposed by Morris and DeShon (2002; cf. Rosenthal, 1994), I computed effect sizes,  $d$ , for both response latencies and error frequencies within each study for the predicted interaction effects. Effect sizes belonging to the reversed interaction pattern (i.e., the two nonsignificant effects in Studies 5a and 5b) were multiplied by  $-1$ . Within each study, these effect sizes were averaged to give a single combined effect size. For Studies 2, 3, 4, 5a, and 5b, these are .25, .47, .29, .26, and .18, respectively. A meta-analysis of these effect sizes based on the weighted integration method developed by Hedges and Olkin (1985) showed that the set of effect sizes was judged homogenous ( $Q = .54, ns$ ) and that the average effect size equaled .29 and was significant ( $p = .014$ ). (As one would expect, the same result was found when the meta-analysis was performed on  $r$ s computed from the  $d$ s). It can thus be concluded that vertical position interfered with judgment performance, jointly indexed by response latencies and error frequencies, such that judgments on a dimension (power or valence) were influenced by a fit of vertical position and attributes of the judgment objects (amount of power or valence, respectively). When only the studies on power judgments were included in the meta-analysis, the mean effect size was slightly higher ( $d = .31, p < .001$ ).

#### Study 6: Influencing Judgments of Power

In the previous studies, it was seen that vertical differences function as a shared metaphor for power and influence how quickly and how accurately power can be judged. One remaining question is whether in addition to these characteristics of the judgment process, the outcome of the judgment itself can also be influenced. In other words, do we attribute more power to an agent just because he or she is on top? The final study addresses this question.

The literature on nonverbal communication of power already contains findings that are pertinent to this question. Schwartz et al. (1982) showed drawings of person dyads, and varied differences within the dyads on several dimensions. One of them was elevation, manipulated such that one of the persons stood on a pedestal above the other. Participants judged which of the persons was dominant. Elevation had by far the largest effect of all dimensions, and in fact, participants judged the elevated person as dominant in 73% of all trials (see also Spiegel & Machotka, 1974). Although this evidence is instructive, it could be explained by inferences that are drawn on the meaning of the pedestal and likely reasons for standing higher. Thus, this is not yet clear evidence that perceptual symbols influence judgment outcomes.

How could vertical position exert such an influence on the judgment result itself? In the previous studies, I have argued that judging power involves the simulation of a vertical spatial dimension and that perceptual content interferes with this judgment. Vertical position enters the judgment process as an additional cue. This reasoning suggests that if the judgment process (erroneously) incorporates vertical position into the judgment, and does not correct for it, then the judgment result itself could also be influenced. This hypothesis is tested in the following study, which differed in three further respects from the previous studies: First, participants did not judge groups presented by labels but animals presented by pictures on the screen. Second, participants did not explicitly judge the power of the animals but how much respect they would feel for the animals.<sup>8</sup> The goal of these changes was to extend the range of evidence from word processing to stimulus displays that are more similar to the real environment and from rather unusual judgments (explicitly judging power) to a more frequently occurring judgment, respect, which is, however, closely related to power (Keltner & Haidt, 2003).

### Method

#### Overview and Design

Participants judged their respect for both powerful and powerless animals, which was the first within-subjects factor: Status. To make it possible that each animal was judged only once by each participant, two lists of animals were created. One half of the participants saw eight of the powerful animals and eight of the powerless animals at the top of the screen, and the other eight powerful and eight powerless animals at the bottom of the screen (List A). For the other half of the participants, assignments of animals to screen position were reversed (List B). This resulted in a between-subjects factor: List. In addition, screen position was varied within participants by showing the animals either at the top or at the bottom of the screen. In sum, the study had a 2 (Status: powerful vs. powerless, within)  $\times$  2 (Position: up vs. down, within)  $\times$  2 (List, between) design.

#### Participants

Data from 113 participants were collected, but the data from 4 of them had to be excluded because they guessed the correct hypothesis. One more participant was excluded because almost all answers were missing. Of the remaining 108 participants, 55 were female; the mean age was 22.1 years ( $SD = 2.6$ ).

#### Materials

The pictures of the 32 animals were from a Web site selling animal replica toys. The pictures showed natural-looking animals in front of a white background, with a light shadow, all of them standing in their typical positions (for a list, see the Appendix). Pictures were scaled to a height of 150 pixels and displayed on the 14-in. laptop screens, set to a resolution of  $1024 \times 768$ . Pictures in the top position appeared with a vertical offset of 50 pixels from the top, and pictures in the bottom position appeared with a vertical offset of 50 pixels from the bottom.

#### Procedure

As in the previous studies, participants were recruited at the campus and seated in front of laptop computers. They were told that the study concerned effects of time pressure on impression formation and judgment and that they were in the group that had only a little time to form an impression but plenty of time to make a judgment. They were further told that they

would see a number of animals one by one on the screen and that they had to judge how much they would be in awe of this animal if they met it in the wild. To record their ratings, participants pressed a number key between 1 (*not at all*) and 9 (*very much*). Each trial started with a blank screen for 600 ms, followed by the question "How much respect do you have for this animal?", centered horizontally and vertically on the screen for 700 ms. Next, the question disappeared and the animal was shown for 800 ms. After it disappeared, the rating scale was shown in the middle of the screen ("*not at all* 1–2 . . . 8–9 *very much*") until an answer was given or until 7 s had elapsed. After four practice trials with animals not used in the actual lists, trials appeared in a random order. The experiment was again programmed in DMDX (Forster & Forster, 2003).

### Results

Ratings were averaged separately for both powerful and powerless animals that had appeared either at the top or at the bottom of the screen. These four scores were submitted to a 2 (Status)  $\times$  2 (Position)  $\times$  2 (List) GLM with repeated measures on the first two factors (cf. Pollatsek & Well, 1995). The predicted main effect of Position failed to reach significance,  $F(1, 106) = 1.43$ ,  $p = .234$ ,  $\eta_p^2 = .01$ . However, there was a significant Position  $\times$  Status interaction,  $F(1, 106) = 5.39$ ,  $p = .022$ ,  $\eta_p^2 = .05$ . Closer inspection of this interaction (see Table 6) showed that powerful animals were indeed judged as more powerful when they appeared at the top of the screen, compared with the bottom position,  $F(1, 106) = 7.70$ ,  $p = .007$ ,  $\eta_p^2 = .07$ . However, there was no difference for powerless animals ( $F < 1$ ).

In addition, there was the trivial main effect of Status,  $F(1, 106) = 1.415$ ,  $p < .001$ ,  $\eta_p^2 = .93$ , showing that powerful animals indeed elicited more respect than did powerless animals. Finally, interactions of List  $\times$  Position and List  $\times$  Position  $\times$  Status emerged. The latter needs explanation because it moderated the central prediction. Closer inspection revealed that it was due to the fact that the powerless animals in one list were slightly more powerful than the powerless animals in the other list, although the animals were randomly assigned to the two lists. Remember that one half of the powerless animals appeared at the bottom for one half of the participants but at the top for the other half of participants. Because one half of the powerless animals were slightly more powerful and because position had no effect on powerless animals in total, a strong disordinal interaction of Position  $\times$  List emerged. When the means for powerless animals were analyzed in a separate 2 (Position)  $\times$  2 (List) GLM with repeated measures on Position, this interaction was significant,  $F(1, 106) = 52.16$ ,  $p < .001$ ,  $\eta_p^2 = .33$ .

The three-way interaction of List  $\times$  Position  $\times$  Status, however, showed that the pattern was different for powerful animals. There position did have a total effect, causing animals appearing at the top to always elicit more respect than those at the bottom. Nevertheless, randomized assignment of animals to the lists seems to have resulted in a less than perfect balancing also for powerful animals, because a separate Position  $\times$  List GLM on respect for powerful animals found an interaction as well,  $F(1, 106) = 5.14$ ,  $p = .025$ ,  $\eta_p^2 = .05$ . This interaction was ordinal, showing that

<sup>8</sup> The German original asked how much "Respekt" participants would feel toward the animal. The German *Respekt* denotes a mixture of respect, admiration, fear, and awe, which is typically felt for authoritative figures.

Table 6  
*Ratings of Respect for Powerful and Powerless Animals (and SEs), Depicted at the Top or Bottom of the Screen (Study 6)*

Animal status	Position of the animal on the screen			
	Top		Bottom	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Powerful	7.68	.90	7.51	.96
Powerless	2.86	.13	2.91	.13

animals appearing at the top elicited more respect than those appearing at the bottom both when the former were from List A and the latter from List B and when the former were from List B and the latter from List A, but that one of the differences was larger than the other. However, because it is due only to the assignment of animals to lists, this interaction is theoretically irrelevant, in contrast to the also emerging significant main effect of position on respect for powerful animals (which reproduced the simple main effect found above).

### Discussion

Study 6 investigated whether in addition to speed and accuracy of power judgments, the result of the judgment could also be influenced by vertical position. The results show that this is true for powerful agents but not for powerless agents. The effect is quite subtle, but it should be noted that the manipulation was very subtle as well: Merely showing pictures of animals at a different place on the screen changed the judgments. The screen position was not confounded with any depicted reason for being higher—the pictures appeared on a white background. This differentiates these findings from earlier demonstrations (e.g., Schwartz et al., 1982) in which elevated positions were created by drawing pedestals, thus providing additional stimuli for inference processes.

Participants indicated that they felt more respect for typically powerful animals (e.g., a lion) when these appeared at the top of the screen, compared with when they appeared at the bottom of the screen. For powerless animals, in contrast, there was no effect of vertical position, leading to a significant interaction. The most likely explanation for this moderation is that the powerless animals used as stimuli were too clearly powerless. This interpretation is supported by the much smaller standard errors for powerless animals, compared with those for powerful animals. In other words, the vertical position cue had no chance to influence the judgment because of the well-defined status of the target. This would be parallel to the well-known fact that priming of concepts works best if the target behaves in a vague or ambiguous manner (Higgins, 1996). If this interpretation is correct, it suggests that perceptual cues can have effects especially on somewhat vague judgment targets.

This effect on judgment is best understood as a variant of a priming effect, in which vertical position primes a certain status. Because a higher vertical position eases and accelerates the categorization of a powerful animal as powerful, the powerful animal is then more likely to be judged as actually very powerful (Bruner, 1957). Interpreting the effect in this way allows additional predic-

tions to be derived from the extensive work on priming and judgment. It could, for instance, be expected that a high accuracy motivation reduces such influences (e.g., Kruglanski, 1989) and that mental load increases such effects (e.g., Gilbert & Hixon, 1991).

### General Discussion

In six studies, it was discovered that the social concept of power is embodied in vertical spatial positions. Study 1 showed that people possess a shared spatial metaphor for power and associate power and height difference. Study 2 showed that vertical positions of group labels on a screen interfere with judgments of which group has power in a dyad. Powerful and powerless groups were identified more quickly when they were in the correct spatial positions, that is, in the positions implied by the perceptual symbol power = up. Study 3 showed that this effect was partly due to the fact that reactions implying an upward or downward movement in space interfere with power judgments: Powerful groups were identified more quickly and more accurately with up movements, whereas powerless groups were identified more quickly and more accurately with down movements. However, the effect in Study 2 was not fully due to the answering movements, as Study 4 showed. It isolated the visual input and showed that the mere vertical position of a group label, in the absence of vertical movements, interfered with power judgments: Powerful groups were identified more quickly when they were at the top of the screen, whereas this was not the case for powerless groups. Study 5 showed that these interference effects are not due to the fact that powerful groups are evaluated more positively, and positive valence is associated with a top position (Meier & Robinson, 2004). The most important finding in Study 5a was that negatively evaluated powerless groups were identified more accurately as powerless when they were at the bottom, whereas negatively evaluated yet powerful groups were identified more accurately as powerful when they were at the top. Across all stimuli, the power of a target group, but not its valence, predicted which vertical position interfered with the power judgment. Study 5b showed that on the other hand, the valence of the used stimuli could very well predict interference of vertical positions, but only when the valence had to be judged, replicating the findings of Meier and Robinson (2004). A meta-analysis of all five interference studies (2, 3, 4, 5a, and 5b) demonstrated that they showed a robust average effect even when both indicators of interference, speed, and accuracy were combined. Finally, Study 6 provided evidence that not only the speed of power judgments but also the judgment result itself is influenced by irrelevant vertical positions: Powerful animals elicited even more respect when they appeared at the top of the screen, whereas there was no effect for powerless animals.

The results of Studies 2–5 can be readily interpreted within theories that understand Stroop-like interference effects as decision processes that gather evidence (e.g., Logan, 1980; see MacLeod, 1991; Ratcliff & Smith, 2004). The response that a presented group is powerful or powerless has a certain threshold, and evidence is accumulated until that threshold is reached. The evidence comes from multiple dimensions, and vertical position of the group label and vertical motor images are among these dimensions.



Analyzing this dimension is not controllable and, in this sense, automatic (Bargh & Chartrand, 2000), which leads to the longer response time in the case of conflicting evidence. The same basic logic applies to Study 6: The vertical position of the animals provides input for the power judgment, but here the vertical input results not only in a speeded or delayed response but in a higher or lower judgment itself.

### *Caveats and Open Questions*

Naturally, the current findings leave a number of questions open and even present new ones. Of special importance are questions of the asymmetry of powerful and powerless targets in the evidence, whether the effects reflect facilitation or interference, and the relation to other notions of semantic association.

In addition to supporting the main hypothesis, the data from Studies 2–5a revealed two unexpected findings. First, the judgments of groups as powerless were less clearly influenced by vertical position than judgments of groups as powerful. Second, the judgments of the powerless were always slower than judgments of the powerful. Before the first finding is interpreted, it is necessary to reiterate that it should be interpreted with caution, as it might be due to simple confounds related to the procedural details and might not have a psychological reason—the interactions are the main results of each study. Nevertheless, it is interesting to speculate what psychological reason might underlie this effect. One reason might be a general tendency to look for the powerful groups in the environment first (Brauer & Bourhis, in press). Another plausible explanation is that the power = up symbol is stronger than the powerless = down symbol. It might be that powerful groups are more important than powerless groups and that they therefore receive more attention and are represented by clearer symbolic representations. Another plausible explanation is that the two findings (smaller effects and slower answers in general for powerless groups) are related. The powerful groups used as stimuli (or even powerful groups in general) might have been easier to judge than the powerless groups. If powerlessness judgments were more complicated, then they involved more elaborate thinking, and consequently the spatial input had comparatively less impact. Increased difficulty of the powerlessness judgments could be due to a lower familiarity with this judgment or to the fact that the power dimension is asymmetric, with a marked end point (powerful) and an unmarked and negated end point (powerless).

An important question that cannot be answered with the present data is whether the patterns found in Study 2–5 are due to facilitation or interference: Does spatial input that fits the power of the group speed up the judgment, or does conflicting spatial input delay judgments, or both? In fact, the current paradigms may not allow a strict test of this question because the judged stimuli always have to be presented somewhere. Presenting the word in the screen's center would allow one to test whether a powerful judgment is delayed by a lower position compared with the middle position. But whether a faster judgment when the word is in an upper position compared with the middle position is due to facilitation (by the higher position) or interference (by the only middle position) would be difficult to say. Different paradigms might be necessary to follow up on this question.

In order to integrate the current findings with other research on knowledge activation, an important question that needs to be discussed is how these effects relate to what is typically called *semantic priming*. The current results were predicted on the basis of Barsalou's (1999) perceptual symbol systems theory, which argues that all conceptual knowledge is based on modal representations. This theory is at odds with common notions of semantic networks, in which quasi-verbal nodes are thought to be associated by links along which activation spreads. These simplistic spreading activation accounts have been called into serious question by findings showing that semantic priming often depends on other contextual variables (e.g., McKoon & Ratcliff, 1992) and does not need to depend on permanent links (McKoon & Ratcliff 1986; cf. Glenberg, 1997). Nevertheless, it is important to note that the current results do not necessarily contradict these traditional notions. As Barsalou (1999) has argued repeatedly, such amodal models of human knowledge can accommodate almost any finding. From the perspective of these models, one could still argue that the perception of a group at the top of the screen activates the quasi-verbal and amodal node *up*, from which activation spreads to the node *powerful*, which then primes the answer. This explanation cannot be ruled out on the basis of the present data, but it should be noted that the embodiment framework predicted the findings a priori. But apart from the theoretical dispute, it becomes a fascinating question about how perceptual interference effects and other measures of associative strength are related. Future studies on their correlation could shed light on the relation between the different accounts. Is priming the word *up* equivalent to showing a word in the top position? Are the two effects moderated by the same factors? Is there a difference between activating the lexical representation of a word and activating the perceptual symbol to which it refers? Perceptual symbol systems theory argues that this is the case. Given that most social cognition theories implicitly equate concepts and lexical representations (words), these are intriguing questions (cf. Gilbert & Hixon, 1991). Taking seriously perceptual symbols as a form of mental representations at least helps to sharpen the understanding of the traditional, and rarely questioned, social cognition model of human knowledge as a semantic network structure and to sharpen our definition of what semantic meaning actually is.

### *Multimodal Mental Representations of Power*

Vertical difference is certainly not the only experience that can become part of the concept of power via schematization. There are probably a number of additional experiences that play a role. Fiske (2004) noted,

[People] constitute Authority Ranking . . . relationships primarily by arranging persons in space, time, magnitude, and force. Higher rank is constituted by being above and in front, being more, coming earlier in time, and having greater 'power or force.' This physics of social relations is iconic, creating social relations by metaphorically mapping people onto position, quantity, and temporal order. (p. 63)

For the association of power and bodily force, supportive evidence is already available from Schubert (2004). This evidence shows that when a gesture of bodily force (making a fist) is induced unobtrusively, it activates the concept of power and leads

to changes in perceptions of power affordances in the environment and to different interpretations of assertive acts performed by other persons. In other words, perception of power is facilitated by a forceful gesture similar to the present results on the vertical spatial dimension. Further evidence (Schubert & Koole, 2004) shows that making a fist also changes the self-concept such that men conceptualize themselves as more powerful in an implicit self-concept task when they make a fist. These findings point to a multimodal mental representation of power. The rich literature on perception of power suggests a number of additional candidates for perceptual symbols (Argyle, 1988; Mehrabian, 1972; Tiedens & Fragale, 2003).

Recently, research on power has emphasized the subjective side of having (or not having) power, the so-called *sense or experience of power*, and its difference from objective power. Sense of power mediates effects of objective power (Anderson & Berdahl, 2002), overrides objective power when it contradicts it (Bugental, Lyon, Krantz, & Cortez, 1997), and has direct behavioral effects (Galinsky, Gruenfeld, & Magee, 2003). This evidence prompts the question of how this sense of power is actually represented. Multimodal perceptual symbols that are used to represent the self might be part of these representations: Feeling powerful might involve the simulation of the self in line with the perceptual symbols one holds for power. Keltner and Haidt (2003) proposed a fourfold model of emotions related to power: pride, which is felt by the powerful who focus on themselves vis-à-vis the powerless; contempt, which is felt by the powerful who focus on the powerless; shame, which is felt by the powerless who focus on themselves vis-à-vis the powerful; and awe, which is felt by the powerless who focus on the powerful. In each of these emotions, one of the perceptual symbols of power may play a role as the eliciting stimulus, as Keltner and Haidt argued, "Humans are prepared to respond to awe-inducing stimuli (e.g., large stature and displays of strength and confidence)" (2003, p. 306).

#### *Embodiment of Another Basic Social Relation: Communal Sharing*

The present evidence supports Fiske's (2004) hypothesis about the embodiment of the basic social relation *authority ranking*. It is interesting that Fiske presented similar arguments for another basic social relation, *communal sharing*. People relating to each other on the basis of communal sharing focus on what they have in common (and what differentiates them from others); the concept is similar to what is called *social identity* or the *social self* (Brewer & Brown, 1998; Tajfel, 1981). Fiske argued that this concept is embodied in variants of physical connection, which can be direct (body to body) or indirect (e.g., through feeding). Physical connection among people embodies communal sharing because it is apprehended as "transferring fundamental social properties, making people alike" (Fiske, 2004, p. 70). What the difference in size between children and parents may be to the embodiment of power by verticality, breast-feeding may be to communal sharing. And just like people have created a multitude of different implementations of vertical difference to signify authority, people have created a multitude of different implementations of direct or indirect bodily contact (and transmission of a substance) to signify communal sharing (see Fiske, 2004). Not surprisingly, social psychologists have been using physical distance as measures of social distance and attitude

toward other persons (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994). Furthermore, measures that schematically represent physical distance are useful to assess identification with another person or an in-group (Aron, Aron, & Smollan, 1992; Schubert & Otten, 2002; Tropp & Wright, 2001). Given these parallels between authority ranking and communal sharing, it would be interesting for future research to apply the methods developed to identify perceptual symbols to the embodied representation of communal sharing as physical contact.

#### *Reification of Perceptual Symbols*

Barsalou (1999) proposed that perceptual symbols develop through a schematization of everyday experience. Of central importance is the experience children make with the correlation of power and size of their parents. But for social concepts, it is important to note that experiences of bodily affordances are not the only possible source for learning a perceptual symbol, because humans structure and create their environments themselves—and they may do so by reference to the perceptual symbols they hold. There are many everyday experiences in which the powerful are also above the others but in which this order of things is made by humans and is part of culture: When charts of athletes, songs, or books are listed, the winners stand on top; when athletes stand on the podium to receive their medals, the winner stands higher than the others; bosses of organizations often get a larger office in the upper floors of a building. All of these height differences fit the power = up perceptual symbol; however, they are not caused by inherent properties of our bodies, but are cultural products. One could say that they are reified perceptual symbols, which are used to constitute, communicate, and yield confirmations of power relations. Again, the work of Fiske (2004) is pertinent: In his conformation theory, he proposed that the association of vertical position and power is an innate proclivity, which is then linked to socially transmitted complements, the reifications. As a result, our perceptual symbols of power are a mixture of an innate core, direct experience, and schematized constitutive and communicative acts. Reifications of perceptual symbols can then themselves become new experience that is schematized into perceptual symbols or that strengthens existing ones. Thus, because humans design their environment by reifying perceptual symbols that were schematized from direct experience of the environment and their bodily interactions with it, perceptual symbols perpetuate themselves.

#### References

- Anderson, C., & Berdahl, J. L. (2002). The experience of power: Examining the effects of power on approach and inhibition tendencies. *Journal of Personality and Social Psychology*, *83*, 1362–1377.
- Argyle, M. (1967). *The psychology of interpersonal behavior*. Baltimore: Penguin.
- Argyle, M. (1988). *Bodily communication* (2nd ed.). New York: Methuen.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612.
- Aronoff, J., Barclay, A. M., & Stevenson, L. A. (1988). The recognition of threatening facial stimuli. *Journal of Personality and Social Psychology*, *54*, 647–655.
- Aronoff, J., Woike, B., & Hyman, L. M. (1992). Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion

- recognition. *Journal of Personality and Social Psychology*, 62, 1050–1066.
- Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle. A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 17–39). New York: Cambridge University Press.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Barsalou, L. W., Niedenthal, P., Barbey, A., & Ruppert, J. (2003). Social embodiment. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 43–92). San Diego, CA: Academic Press.
- Borgh, A. M., Glenberg, A. M., & Kaschak, M. P. (2003). Putting words in perspective. *Memory & Cognition*, 32, 863–873.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75, 1–28.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43, 1–22.
- Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, 13, 185–189.
- Brauer, M., & Bourhis, R. Y. (in press). Powerful and powerless individuals as perceivers, targets, and actors. *European Journal of Experimental Social Psychology*.
- Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In D. T. Gilbert & S. T. Fiske (Eds.), *Handbook of social psychology* (pp. 554–594). New York: McGraw-Hill.
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64, 123–152.
- Bugental, D. B., Lyon, J., Krantz, J., & Cortez, V. (1997). Who's the boss? Differential accessibility of dominance ideation in parent–child relationships. *Journal of Personality and Social Psychology*, 72, 1297–1309.
- Crusius, J., & Wentura, D. (2005). *Assessment of implicit attitudes using the affective matching paradigm*. Unpublished manuscript, Saarland University, Saarbrücken, Germany.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1–40). San Diego, CA: Academic Press.
- Ellyson, S. L., & Dovidio, J. F. (1985). Power, dominance, and nonverbal behavior: Basic concepts and issues. In S. L. Ellyson & J. F. Dovidio (Eds.), *Power, dominance, and nonverbal behavior* (pp. 1–27). New York: Springer-Verlag.
- Felson, R. B. (2002). *Violence & gender reexamined*. Washington, DC: American Psychological Association.
- Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, 99, 689–723.
- Fiske, A. P. (2004). Four modes of constituting relationships: Consubstantial assimilation; space, magnitude, time, and force; concrete procedures; abstract symbolism. In N. Haslam (Ed.), *Relational models theory. A contemporary overview* (pp. 61–146). Mahwah, NJ: Erlbaum.
- Forster, K. I., & Forster, J. C. (2003). DMDX [Computer software]. Retrieved September 1, 2003, from University of Arizona Web site: <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>
- Förster, J., & Stepper, S. (2000). Compatibility between approach/avoidance stimulation and valenced information determines residual attention during the process of encoding. *Journal of Personality and Social Psychology*, 30, 853–871.
- Förster, J., & Strack, F. (1996). Influence of overt head movements on memory for valenced words: A case of conceptual–motor compatibility. *Journal of Personality and Social Psychology*, 71, 421–430.
- Freedman, D. G. (1979). *Human sociobiology*. New York: Free Press.
- Galinsky, A. D., Gruenfeld, D., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, 85, 453–466.
- Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space → time metaphors. *Language and Cognitive Processes*, 17, 537–565.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509–517.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.
- Glenberg, A. M., & Kaschak, M. P. (2003). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–565.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Haidt, J. (2003). Elevation and the positive psychology of morality. In C. L. M. Keyes & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 275–289). Washington, DC: American Psychological Association.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hewes, G. W. (1955). World distribution of certain postural habits. *American Anthropologist*, 57, 231–244.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–937.
- Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind*. Boston: Houghton Mifflin.
- Johnson-Laird, P. N. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income: Preliminary test of a theoretical model. *Basic and Applied Social Psychology*, 89, 428–441.
- Keltner, D., Gruenfeld, D., & Anderson, C. P. (2003). Power, approach, and inhibition. *Psychological Review*, 110, 265–284.
- Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition & Emotion*, 17, 297–314.
- Kruglanski, A. W. (1989). The psychology of being “right”: The problem of accuracy in social perception and cognition. *Psychological Bulletin*, 106, 395–409.
- Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12, 523–553.
- Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events: Nature or culture? *Psychological Science*, 14, 296–301.
- MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight. Stereotypes on the rebound. *Journal of Personality and Social Psychology*, 67, 808–817.
- McClelland, D. C. (1975). *Power: The inner experience*. New York: Wiley.

- McKoon, G., & Ratcliff, R. (1986). Automatic activation of episodic information in a semantic memory task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 108–115.
- McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1155–1172.
- Mehrabian, A. (1972). *Nonverbal communication*. Chicago: Aldine-Atherton.
- Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up. *Psychological Science*, *15*, 243–247.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105–125.
- Neumann, R., Förster, J., & Strack, F. (2003). Motor compatibility: The bidirectional link between behavior and evaluation. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation* (pp. 371–391). Mahwah, NJ: Erlbaum.
- Neumann, R., & Strack, F. (2000). Approach and avoidance: The influence of proprioceptive and exteroceptive cues on encoding of affective information. *Journal of Personality and Social Psychology*, *79*, 39–48.
- Niedenthal, P., Barsalou, L. W., Winkielman, P., Kraut-Gruber, S., & Ric, F. (in press). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review*.
- Peeters, G., & Czapinski, J. (1990). Positive–negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*, 33–60.
- Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research. A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785–794.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models of two-choice reaction time. *Psychological Review*, *111*, 333–367.
- Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*, 767–780.
- Richardson, D. C., Spivey, M. J., Edelman, S., & Naples, A. D. (2003). “Language is spatial”: Experimental evidence for image schemas of concrete and abstract verbs. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society* (pp. 873–878). Mahwah, NJ: Erlbaum.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York: Russell Sage Foundation.
- Schnall, S., & Clore, G. L. (in press). Emergent meaning in affective space: Congruent conceptual relations and spatial relations produce positive evaluations. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Schubert, T. W. (2004). The power in your hand: Gender differences in bodily feedback from making a fist. *Personality and Social Psychology Bulletin*, *30*, 757–769.
- Schubert, T. W., & Koole, S. (2004). *Making a fist causes an autonomous implicit self concept in men*. Unpublished data, University of Jena, Germany.
- Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, *1*, 535–576.
- Schwartz, B. (1981). *Vertical classification: A study in structuralism and the sociology of knowledge*. Chicago: University of Chicago Press.
- Schwartz, B., Tesser, A., & Powell, E. (1982). Dominance cues in nonverbal behavior. *Social Psychology Quarterly*, *45*, 114–120.
- Spiegel, J., & Machotka, P. (1974). *Messages of the body*. New York: Free Press.
- Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge, England: Cambridge University Press.
- Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100.
- Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*, 558–568.
- Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*, 585–600.
- Wentura, D., Rothermund, K., & Bak, P. M. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology*, *78*, 1024–1037.
- Young, T. J., & French, L. A. (1996). Height and perceived competence of U.S. presidents. *Perceptual and Motor Skills*, *82*, 1002.
- Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*, 945–958.

## Appendix

## Stimuli Used in Studies 1–6

The following propositions were used in Study 1:

Powerful: rules over, has influence on, is stronger than, is superior to, exerts power on, defeats.

Powerless: defers to, loses against, gives in to, is weaker than, succumbs to, obeys.

Horizontal: gives something to, wants, pushes, pulls, points toward, runs away from.

The following group pairs were used in Study 2: employer–employee, boss–staffer, coach–athlete, Russia–Chechnya, perpetrator–victim, pope–bishop, officer–soldier, master–servant, doctor–patient, West Germans–East Germans, chief editor–journalist, captain–sailor, boss–secretary, judge–defendant, professor–students, chancellor–minister, policeman–criminal, U.S.A.–Iraq, government–opposition, parents–child, teacher–pupil, master–apprentice, winner–loser, warder–prisoner.

The following groups were used in Study 3 and 4:

Powerful: boss, judge, professor, chancellor, policeman, U.S.A., government, entrepreneur, general, politician, chief, king, sovereign, president, head physician, officer. In Study 4, *officer* was replaced by *warder*.

Powerless: secretary, defendant, students, child, pupil, apprentice, loser, prisoner, arrestee, aide, Lithuania, assistant, menial, worker, sick person, slave.

The following group names, from Crusius and Wentura (2005), were used in Study 5:

Negative other-relevant: thief, troublemaker, enemy, gangster, attacker, cheater, dictator, villain.

Negative self-relevant: fool, nonworker, dilettante, failure, wimp, cripple, pessimist, loser.

Negative other-relevant: friend, paramedic, brother, admirer, nurse, comrade, benefactor, sister.

Positive self-relevant: victor, lucky fellow (translated from the German *Glückskind*), winner, genius, optimist, athlete, adventurer, connoisseur.

Pictures of the following animals were used in Study 6:

Powerful animals: gorilla, tiger, lion, bison, European bison, wolf, white bear, elk, deer, cheetah, elephant, wild boar, grizzly, reindeer, rhinoceros, musk oxen.

Powerless animals: hare, donkey, sheep, horse, hind, squirrel, foal, boar, llama, armadillo, ibex, shoat, roebuck, doe, young doe, fawn.

All original German group labels and original pictures are available from the author.

Received April 5, 2004

Revision received March 8, 2005

Accepted March 14, 2005 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!