Social Cognitive and Affective Neuroscience Advance Access published June 27, 2015

The Neural Basis of Conceptualizing the Same Action at Different Levels of Abstraction

[1]Robert P. Spunt, [2]David Kemmerer, [1]Ralph Adolphs

Word Count: 5,609
Number of Tables: 4
Number of Figures: 4

Author Affiliations:

[1]California Institute of Technology

[2]Purdue University

Corresponding Author:

Robert P. Spunt

California Institute of Technology

1200 E. California Blvd.

Pasadena, CA 91125

Phone: (818) 645-8188

Email: spunt@caltech.edu

Abstract

People can conceptualize the same action (e.g., "riding a bike") at different levels of abstraction (LOA), where higher LOAs specify the abstract motives that explain *why* the action is performed (e.g., "getting exercise"), while lower LOAs specify the concrete steps that indicate *how* the action is performed (e.g., "gripping handlebars"). Prior neuroimaging studies have shown that *why* and *how* questions about actions differentially activate two cortical networks associated with mental-state reasoning and action representation, respectively; however, it remains unknown whether this is due to the differential demands of the questions *per se* or to the shifts in LOA those questions produce. We conducted functional MRI while participants judged pairs of action phrases that varied in LOA and that could be framed either as a *why* question (Why ride a bike? Get exercise.) or a *how* question (How to get exercise? Ride a bike.). Question framing (*why vs. how*) had no effect on activity in regions of the two networks. Instead, these regions uniquely tracked parametric variation in LOA, both across and within trials. This suggests that the human capacity to understand actions at different levels of abstraction is based in the relative activity of two cortical networks.

Keywords: action understanding, semantic memory, abstraction, concepts, social cognition, fMRI

Introduction

People have a rich set of concepts for representing actions and a flexible capacity to use these concepts to understand both their own and other people's actions. This allows humans to not only understand the meaning of different actions, but also understand that the *same* action can carry different meanings. For instance, an action initially identified as "writing a manuscript" might be further conceptualized in terms of its abstract motives—i.e., *why* it is done (e.g., "sharing knowledge")—or its concrete implementation—i.e., *how* it is done (e.g., "typing words"). The present study attempted to clarify the neural basis of conceptualizing an action at varying levels of abstraction.

Several recent neuroimaging studies have shown that answering *why* versus *how* questions about action reliably differentiates activity in two left-lateralized cortical networks (Spunt et al. 2010; Spunt et al. 2011; Spunt and Lieberman 2012a; Spunt and Lieberman 2012b; Spunt and Adolphs 2014). Specifically, the Why > How contrast reliably reveals activation in the dorsomedial and ventromedial prefrontal cortex (PFC), the anterior superior temporal sulcus (STS), the temporoparietal junction (TPJ), and the posterior cingulate cortex (PCC)—regions that have been independently implicated in representing and reasoning about the mental states that typically drive actions, such as beliefs, desires, and intentions (Gallagher and Frith 2003; Saxe 2006; Carrington and Bailey 2009; Van Overwalle and Baetens 2009; Lieberman 2010; Mar 2011; Denny et al. 2012; Schurz et al. 2014). Conversely, the How > Why contrast reliably activates the dorsal and ventral premotor cortex (PMC), posterior middle temporal gyrus (MTG), rostral inferior parietal lobule (IPL), and dorsal precuneus—regions that have been independently implicated in representing the visual motion patterns and somatomotor features of actions when perceived and performed (Caspers et al. 2010; Molenberghs et al. 2012; Rizzolatti et al. 2014), conceptualized (Kemmerer et al. 2012; Watson et al. 2013; Urgesi et al. 2014), and verbally processed (Kemmerer et al. 2012; Pulvermuller 2013; Kemmerer 2015).

The reliable cortical dissociation of *why* and *how* questions is based on an attentional manipulation referred to as the Why/How Task (Freitas, Gollwitzer, & Trope, 2004), where participants are asked to answer (or evaluate answers to) *why* and *how* questions about the same stimuli (e.g., named or visually presented actions). This ensures that any observed effects are caused by the different cognitive demands of answering a *why* versus *how* question about the same action. Yet, the nature of these differential demands—and their relationship to the *why/how* distinction—remains unclear. Here, we evaluate two alternative accounts of the robust neural effects observed in the *why*/*how* contrast. These alternatives are suggested by the hierarchical structure of human representations of action (Vallacher and Wegner 1987; Kozak et al. 2006; see also Fujita et al. 2006; Trope and Liberman 2010). To illustrate, consider the role that *why* and *how* questions play in navigating the levels of the conceptual action hierarchy in **Figure 1A**. On the one hand, *why* and *how* questions elicit directional (signed) shifts up and down the levels of the hierarchy, respectively. Taking as examples the intermediate-level phrases “make a phone call” and “contact a friend”, a *why* question enables an upward movement (“Why make a phone call? Contact a friend”), while a *how* question enables the reverse (“How to contact a friend? Make a phone call”). This suggests that the cortical networks modulated by the *why*/*how* contrast may underlie the distinct “mindsets” (Gollwitzer et al., 1990) for thinking about actions in terms of their *motive* versus their *implementation*.

This account is complicated by the fact that the language used to represent an action becomes progressively less concrete, specific, imageable, embodied, and emotionally neutral as one ascends the levels of the hierarchy. Henceforth, we use the general term Level of Abstraction (LOA) to refer to this basic semantic dimension correlated with increasing levels of a conceptual action hierarchy. Because previous neuroimaging studies contrasted *why* and *how* questions about the same actions (e.g., “*Why*/*How* to contact a friend?)”, the answers yielded by *why* questions were likely conceptually higher in LOA (e.g., “Feel connected”) than those yielded by corresponding *how* questions (e.g., “Make a phone call”). For this reason, the effective ingredient in the *why*/*how* contrast may be the differential demands of conceptualizing the same action at relatively higher versus lower LOAs. In the present study, we designed a novel Why/How Task that separates the *why*/*how* question manipulation from the changing conceptual demands imposed by understanding actions at different LOAs. This task design exploits the fact that *why* and *how* questions can be posed at any level of a conceptual action hierarchy, making it possible for a *why* question ("Why dial numbers?") to yield an answer that is less abstract ("Make a phone call") than an answer ("Contact a friend") to a *how* question ("How to feel connected?"). This orthogonality allowed us to evaluate the alternative interpretations of the *why*/*how* contrast described above.

Method

Participants

Nineteen right-handed adults participated in the study in exchange for financial compensation. Participants were neurologically and psychiatrically healthy, had normal/corrected-to-normal vision, spoke English fluently, had IQ in the normal range (assessed with the Wechsler Abbreviated Scales of Intelligence), and were not pregnant. Participants provided written informed consent according to a protocol approved by the Institutional Review Board of the California Institute of Technology. Data from two participants were excluded due to poor task performance (no response to greater than 10% of trials). This left 17 participants for the analysis (11 males, 6 females; mean age = 29.47, age range = 21-46). A power analysis of the Why/How contrast from Spunt and Adolphs (2014) indicated the present study's sample size was sufficiently large to test our hypotheses (see Supplemental Methods).

Action Stimuli

Stimuli consisted of 128 syntactically similar phrases, each of which used a simple verb-complement construction (typically a transitive verb and its direct object) to describe a familiar type of human action (the complete set is provided in **Table S1**). As illustrated in **Figure 1A**, each phrase was part of a grouping of four phrases that collectively formed a conceptual action hierarchy, such that the action described by a phrase at one level (e.g., "Make a phone call") was both a commonly-accepted motive for the action described by a phrase at the level below it (e.g., "Dial numbers"), and a commonly-accepted means for the action described by the phrase at the level above it (e.g., "Contact a friend"). By pairing phrases at contiguous levels, each set of four phrases yielded three *why* question-answer pairs (e.g., "*Why* dial numbers? Make a phone call") and three *how* question-answer pairs (e.g., "*How to* make a phone call? Dial numbers").

The final group of 25 question-answer pairs was selected by identifying those that elicited the highest degree of yes/no response agreement in an independent sample of approximately 40 native English-speaking adults recruited online via Amazon.com's Mechanical Turk (see **Appendix A** for details). This resulted in 150 total test trials, 30 of which were foils in that they featured question-answer pairs that were commonly-rejected (e.g., "How to show ambition? Serve alcohol"). These foils were included to check task comprehension and were not of interest in the fMRI analyses.

We characterized each verb-complement phrase (without the *why* and *how* question markers) on several lexical dimensions, including the number of characters and words, and the average frequency of content words (per million words in the SUBTLEX database). In addition, we used independent groups of native English-speaking Mechanical Turk participants to norm each phrase on five semantic dimensions believed to covary with the LOA (see **Appendix A** for details): (1) abstract vs. concrete (*N*=150), (2) non-imageable vs. imageable (*N*=97), (3) broad vs. specific (*N*=97), (4) mind-dependent vs. body-dependent (*N*=132), and (5) emotionally valenced vs. neutral (*N*=96), which was calculated as the absolute deviation from the neutral point on a bipolar (negative/positive) scale. Interrater reliability for all measures was excellent (minimum intra-class correlation coefficient=.98; Shrout and Fleiss, 1979). Given their substantial inter-correlation (**Table 1)**, we subjected the five semantic dimensions to a principal component analysis (PCA) using the MATLAB Statistics Toolbox (version R2014b; MathWorks Inc., Natick, MA, USA). The PCA showed that a single component explains 91.01% of their total variance (first eigenvariate=8.12, second eigenvariate=.42). Consequently, we used the scores for this underlying dimension to quantify the LOA for each action phrase.

As shown in **Table 2**, independent-samples t-tests showed that Why and How trials were matched not only on all lexical parameters, but also on the average LOA of the question and answer in each trial, which we henceforth call "Trialwise LOA." As described below, our first two analyses exploit the orthogonality of the Why/How manipulation and Trialwise LOA to identify their independent effects. In contrast, our third analysis exploits the fact that Why and How trials differed in what we henceforth call the "Signed LOA Shift", which refers to the within-trial change in LOA introduced by the answer phrase relative to the question phrase. This is calculated for each trial by subtracting the LOA of the question phrase (henceforth called the "Prepotent LOA") from the LOA of the answer phrase. In line with the view that *why* and *how* questions produce relative increases and decreases in LOA, respectively, Why trials were associated with positive Signed LOA Shifts while How trials were associated with negative Signed LOA Shifts (**Table 2**).

Experimental Task

The 150 test trials were presented to participants in an event-related design. The trial structure and timing are shown in **Figure 1B**. Trials were arranged in a pseudo-random order with a variable onset asynchrony drawn from a pseudo-exponential distribution with a mean of 6000ms (Min/Max=5000/10000ms). The order and onsets of trials were optimized for estimating the *why*/*how* contrast. This was achieved by generating the design matrices for one million pseudo-randomly generated designs, and for each summing the efficiencies of Why > How contrast estimation. The most efficient design was used for all participants.

Stimulus presentation and response recording were implemented with the MATLAB

Psychophysics Toolbox (version 3.0.9; Brainard, 1997). An LCD projector was used to present the task on a screen at the rear of the scanner bore that was visible to participants through a mirror positioned on the head coil. Participants were given a button box and made their yes/no responses with their right-hand index/middle fingers. Prior to the experimental task, participants performed a practice version featuring trials not used in the experimental task.

Image Acquisition

All imaging data were acquired at the Caltech Brain Imaging Center using a Siemens Trio 3.0 Tesla MRI scanner outfitted with a 32-channel phased-array headcoil. For the experimental task, we acquired 909 T2*-weighted echoplanar image volumes (EPIs; multi-band acceleration factor=4, slice thickness=2.5mm, in-plane resolution=2.5 x 2.5mm, 56 slices, TR=1000 ms, TE=30ms, flip angle=60°, FOV=200mm). Participants' in-scan head motion was minimal (max translation=1.93mm, max rotation=1.75°). We also acquired an additional 1,330 EPI volumes for each participant as part of a separate study. Finally, we acquired a high-resolution anatomical T1-weighted image (1mm isotropic) and fieldmaps used to estimate and correct for inhomogeneity-induced image distortion.

Image Preprocessing

Images were analyzed using Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Each participant's EPI timeseries was subjected to the following preprocessing steps: the first four volumes were discarded to account for T1-equilibration effects; the realign and unwarp procedure was used to perform distortion correction and motion correction; their T1 structural volume was co-registered to the mean of the corrected EPI volumes; group-wise DARTEL registration (Ashburner, 2007) was used to normalize the T1 volume to a group-specific template, with subsequent affine registration to Montreal Neurological Institute (MNI) space; and all EPI volumes were normalized to MNI space using the deformation flow fields from the previous step, which re-sampled volumes (2x2x2 mm) and applied spatial smoothing (6mm Gaussian kernel, full-width-at-half-maximum).

Single-Subject Regression Models

General linear models were used to estimate three models of the EPI timeseries for each participant. The following procedures were used in all three models. First, covariates of interest excluded foil trials and trials to which the participant gave either no response or the incorrect response. Excluded trials were modeled in a separate nuisance covariate. Second, the neural response to each trial was defined with variable epochs (Grinband et al., 2008) spanning the onset of the question phrase and the offset of the answer phrase (see **Figure 1B**). Third, all models included trialwise nuisance covariates for variability in response time (RT), character count, word count, and average content word frequency. The three lexical parameters pertain only to the simple verb-complement action phrases, without the *why* and *how* question markers. Each of these nuisance covariates was constructed by modulating the amplitude of the predicted neural response for each trial of interest by the de-meaned parameter value. Fourth, all models included scanwise nuisance covariates for the 6 motion parameters estimated from image realignment and a predictor for every timepoint where in-brain global signal change (GSC) exceeded 2.5 SDs of the mean GSC or where estimated motion exceed 0.5 mm of translation or 0.5° of rotation. Finally, all models used the canonical (double-gamma) response function and its temporal derivative to model the hemodynamic response; were high-pass filtered at 1/128 Hz; and were estimated using the SPM8 RobustWLS toolbox, which implements robust weighted least-squares estimation (Diedrichsen and Shadmehr, 2005).

*Estimating the Independent Effects of Why/How Questions and Trialwise LOA***.** The first two models were designed to assess the independent contributions of two factors: variation in the question in each trial (Why vs. How), and variation in the Trialwise LOA (average LOA for question and answer phrases). The models differed only in their representation of the factor corresponding to Trialwise LOA.

The first model included four covariates of interest corresponding to the cells created by crossing factors corresponding to the Question (Why vs. How) and Trialwise LOA binarized into a two-level factor (High vs. Low). We binarized LOA using Otsu's method (implemented using GRAYTHRESH in MATLAB), which selects the threshold that minimizes the variance within the two bins. Independent-samples t-tests confirmed that Trialwise LOA for trials binned as High LOA was significantly higher than that for trials binned as Low LOA, $t(109)=20.57$, $p<.001$. Following estimation, we computed the [(High-Why + Low-Why) > (High-How + Low-How)] contrast to identify regions independently modulated by the Why/How factor, and the [(High-Why + High-How) > (Low-Why + Low-How)] contrast to identify regions independently modulated by Trialwise LOA.

The second model included two covariates of interest corresponding to trialwise variability in the Question (Why vs. How) and in LOA. These covariates were constructed by modulating the height of the predicted neural response to each trial by a value representing the Question (Why=+1; How=-1) and Trialwise LOA. In addition to the nuisance covariates specified above, this model also included a covariate corresponding to the fixed-amplitude (time-invariant) response to each trial. For subsequent group analysis, we computed two contrast images, one for the modulator coding the Why/How contrast, and one for the modulator coding Trialwise LOA.

Estimating the Independent Effects of the Prepotent LOA and Signed LOA Shift. The third and final model exploits the fact that although Why and How trials were equated on Trialwise LOA, they differed in the direction of the Signed LOA Shift (**Table 2**). Moreover, the magnitude of these shifts varied across trials in a way uncorrelated with variation in Trialwise LOA (r = -.02). **Figure 4A** shows that the magnitude of such shifts showed a moderate negative correlation with the Prepotent LOA (r = -.49), defined earlier as the LOA of the just question phrase. If the active ingredients of *why*/*how* questions are *increases*/*decreases* in LOA, then brain regions associated with the *why*/*how* contrast in prior work should track the magnitude of the Within-Trial Shift in LOA relative to the Prepotent LOA. Hence, this model included two covariates of interest corresponding to trialwise variability in both of these measures. To ensure that any shift-related effects were due to *signed* shifts, we included an additional nuisance covariate for the *unsigned* (absolute) LOA shifts. As in the second model, this model also included a covariate of no interest corresponding to the fixed-amplitude response to each trial. For subsequent group analysis, we computed two contrast images, one for the modulator coding the Prepotent LOA, and one for the modulator coding the Signed LOA Shift.

Group Analyses

Contrasts were first interrogated using a set of independently-defined left-hemisphere regions of interest (ROI) based on the group-level Why/How contrasts from Study 1 and Study 3 reported in Spunt and Adolphs (2014). These images are publicly available on NeuroVault (http://neurovault.org/collections/445/). We selected five ROIs from both the Why > How and How > Why contrasts (**Figure 2**). The peak coordinate and spatial extent of each ROI is provided in **Table S2**. For each ROI, we tested our hypotheses with t-tests on the average parameter estimate across voxels. Confidence intervals (CIs) were estimated using the bias-corrected and accelerated percentile method (10,000 random samples with replacement; implemented using BOOTCI in MATLAB).

ROI analyses were complemented by whole-brain analyses. We conducted one-sample t-tests on single-subject contrast images for effects of interest. The resulting group-level tstatistic images were interrogated by applying a cluster-forming threshold of $p<.001$ followed by cluster-level correction at a family-wise error (FWE) of .05. Thresholded results were surface rendered using SurfPlot (http://mrtools.mgh.harvard.edu/index.php/SurfPlot).

Results

Behavioral Results

Table 2 displays the mean percentage correct and response time (RT) to correct responses across Why and How trials. Paired-sample t-tests yielded no evidence for Why/How effect on either outcome (*p*s > .295). Given that the LOA factors examined in the regression models were based only on trials with correct responses, we tested the effect of LOA only on RT to correct trials. We first tested this with the binarized LOA factor used in the factorial model. A paired-samples t-test showed that RTs to High LOA trials (M=1035 ms, SD=143) were longer than RTs to Low LOA trials (M=965 ms, SD=137), $t(16) = 5.52$, $p < .001$, 95% CI_{Boot} [-.096, -.048]. Next, we tested for a non-zero within-subject correlation of RT and Trialwise LOA. A one-sample t-test showed that the within-subject correlation between RT and Trialwise LOA ($r_{mean} = 0.14$, $r_{SD} = .08$) was reliably above zero, $t(16) = 6.89$, $p < .001$, 95% CI_{Boot} [.097, .172].

These effects of Trialwise LOA on RT underscore the importance of controlling for within-subject performance variability on the measured fMRI signal. Indeed, as shown in **Table S5** and **Figure S1,** longer RTs were indeed associated with activity in a distributed cortical network consistent with regions observed in previous studies examining the executive aspects of semantic memory use (Badre et al. 2005; Binder et al. 2005; Goldberg et al. 2007; Hoffman et al. 2010; Raposo et al. 2012; Satpute et al. 2014). Critically, the imaging results presented below are statistically independent of these RT effects.

Imaging Results

Effects of Why/How Questions and Trialwise LOA. **Table 3** shows ROI-specific results for the first and second models that examine the independent contributions of Question (Why vs. How) and binarized Trialwise LOA. Remarkably, neither model produced evidence for a Why/How effect in any of the 10 ROIs examined. In contrast, the Trialwise LOA factor in both models showed a reliable effect on the same 8/10 ROIs, including 5/5 of the ROIs selected for showing an effect in the Why > How contrast in prior studies.

These findings are largely reproduced in the whole-brain analysis (**Table S3**). Neither model yielded evidence for a reliable Why/How effect in any region. In contrast, the Trialwise LOA factor modulated a largely left-lateralized cortical network that was consistently localized across both models. As shown in **Figure 3,** higher LOAs were consistently associated with regions of the medial PFC, anterior STS, temporal pole, TPJ, and precuneus, while lower LOAs were consistently associated with the posterior MTG, rostral IPL, and regions of the IFG around the pars triangularis. Several clusters were only observed in the second model examining parametric variation in Trialwise LOA. This included bilateral regions of middle occipital cortex that were associated with increasing Trialwise LOA. Although not predicted, this finding is consistent with recent work showing a left middle occipital association with abstraction (Gilead, Liberman & Maril, 2014). Moreover, decreasing Trialwise LOA yielded more extensive regional associations in the second model, including regions of the ventral PMC and presupplementary motor area.

*Effects of the Prepotent LOA and Signed LOA Shift***. Table 4** shows ROI-specific results for the third model examining the Prepotent LOA and the Signed LOA Shift. Parametric variation in Prepotent LOA showed largely the same effects observed for the Trialwise LOA parameters examined by the first two models, and this observation was reproduced by the whole-brain analysis (**Figure 4B; Table S4**). More interestingly, however, when examining the Signed LOA Shift, we found evidence for a positive association with 4 of the 5 ROIs from the Why > How contrast (the effect in the TPJ ROI was also in the expected direction, p=.062). And although none of the *a priori* ROIs based on the How > Why contrast showed a significant negative association with the Signed LOA Shift, 4 of 5 were in the expected direction. In the whole-brain analysis, the dorsomedial PFC and the anterior STS showed a positive association with the Signed LOA Shift (**Figure 4C; Table S4**). No regions, however, were found to show a negative association with the Signed LOA Shift.

Discussion

We investigated the neural basis of conceptualizing the same actions at different LOAs. This was directly motivated by previous neuroimaging studies showing that *why* and *how* questions about actions differentially activate cortical networks associated with mental-state reasoning and action representation, respectively. Since these studies always asked *why* and *how* questions about the same action stimuli (named or visually presented), they confounded the task manipulation (*why*/*how* questions) with the use of action concepts that varied in LOA. We deconfounded these two factors in order to evaluate two alternative functional accounts of the cortical networks known to be modulated by the *why*/*how* contrast. We found no support for the account that *why* and *how* questions per se elicit distinct and content-free cognitive sets for conceptualizing action. Instead, the evidence supported the alternative account that the distinct effects of *why* and *how* questions are caused by the relative increases and decreases in the conceptual LOA they tend to produce, respectively.

Brain Regions for Conceptualizing an Action at Different LOAs

Increasing LOAs were associated with the set of left hemisphere regions reliably associated with the Why > How contrast in previous studies (**Figure 2**; Spunt et al. 2010; Spunt et al. 2011; Spunt and Lieberman 2012a; Spunt and Lieberman 2012b; Spunt and Adolphs 2014). These regions partially overlap with several meta-analytically defined functional networks, including: (1) the so-called "theory-of-mind" or "mentalizing" network associated with tasks of mental-state reasoning (Gallagher and Frith 2003; Amodio and Frith 2006; Saxe 2006; Carrington and Bailey 2009; Van Overwalle and Baetens, 2009; Schurz et al. 2014); (2) the default mode network (DMN), especially its dorsomedial PFC component (Raichle et al. 2001; Buckner et al. 2008; Andrews-Hanna et al. 2010; Andrews-Hanna et al. 2014); (3) the network associated with mentally simulating episodes both past and future (Hassabis and Maguire 2007; Spreng et al. 2009; Schacter et al. 2012); (4) the network associated with comprehending narrative discourse (Ferstl and von Cramon 2001; Ferstl et al. 2008; Mar 2011; Nijhof and Willems 2015); (5) the network associated with transmodal semantic processing (Binder et al. 2009; Binder and Desai 2011); and (6) the network associated with comprehending abstract compared to concrete words (Binder et al. 2009; Wang et al. 2010; Binder and Desai 2011).

Decreasing LOAs were associated with the majority of the left hemisphere regions reliably associated with the How > Why contrast in previous work. These regions fall within a broader functional network responsive to cognitive tasks devoid of meaningful socio-emotional content (Van Overwalle 2011; Jack et al. 2012). But they are more frequently regarded as forming a subset of the functional network thought to enable representation of the visual and somatomotor features of actions when they are perceived, performed, conceptualized, and verbally processed (Caspers et al. 2010; Kemmerer et al. 2012; Molenberghs et al. 2012; Pulvermuller 2013; Watson et al. 2013; Rizzolatti et al. 2014; Urgesi et al. 2014; Kemmerer 2015). This included a region of the left posterior MTG that has been associated with encoding the visual motion components of action concepts (Chen et al. 2008; Deen and McCarthy 2010; Saygin et al. 2010; Wallentin et al. 2011; Humphreys et al. 2013; Watson et al. 2013). Interestingly, this seems to contrast with the view that the left posterior MTG represents more schematic aspects of the event structures encoded by both action and non-action verbs/sentences (Bedny et al. 2008; Bedny et al. 2012).

Does the LOA construct help integrate this diverse set of findings? Several proposals have attempted to integrate at least a subset of these findings, but these have primarily regarded the regions associated with increasing LOA in the present study, which we have recently demonstrated map well on to the dorsomedial PFC subsystem of the DMN (Spunt et al. 2015). For instance, Buckner and Carroll (2007) suggest that mental-state reasoning, perspective-taking, episodic memory, and prospection all depend on a process they call "self-projection", which involves the mental simulation of events and experiences that transcend the immediate environment. Mar and Oatley (2008) suggest a similar process of simulation to describe the intense social and emotional experiences that literary narratives are capable of evoking (see also Nijhof and Willems 2015). Finally, Liberman and Trope (2014; Trope and Liberman 2010) assert that mental representations can all be characterized by how "psychologically distant" they are from "a common zero-distance point, which is the experienced reality of me here and now" (Liberman and Trope 2014, p. 1). While these theories have been primarily focused on the process of abstracting away from sensorimotor experiences (corresponding to increasing LOA in the present study), others have been primarily concerned with understanding the neural bases of relevant conceptual dualisms, such as between embodied and disembodied semantics (Pulvermuller 2013) or between social and physical knowledge domains (Van Overwalle and Baetens 2009; Jack et al. 2012).

What these accounts all seem to share is a concern with understanding the human brain's ability for abstract thought, and distinguishing that ability from the one enabling concrete thought. The present study shares this concern in the important domain of action understanding, taking as its point of departure the idea that people naturally think about their actions as having a hierarchical structure. Such a point of departure overcomes a number of theoretical limitations of existing proposals. One such limitation is that prior proposals have typically focused on just one side of the conceptual coin, namely, on mental processes involved in abstraction. The present account parsimoniously handles both sides of the coin by identifying upward movements with abstraction and downward movements with the human ability to ground their abstract ideas in sensory experience and motor acts. Proposals that are concerned with both sides of the coin are limited by a tendency to frame the distinction as a conceptual dichotomy or dualism. The present account naturally permits representations to vary on a continuous dimension corresponding to the levels of a conceptual hierarchy. Finally, prior accounts have paid very little attention to identifying the specific task conditions and mental processes by which people shift their level of understanding. The present account is grounded in an ecologically-valid method for eliciting such shifts: The Why/How Task.

Limitations and Future Directions

We identify several limitations of the present study that offer worthwhile directions for future research. First, we acknowledge the preliminary status of our empirical definition of LOA. The five semantic dimensions used to define the LOA factor were not intended to provide a complete and definitive list of semantic dimensions constituting LOA. There are likely many additional dimensions that could be included in an expanded definition of LOA. For instance, compare "sharing knowledge" to "typing words" as descriptions of the same act of writing a scientific paper. Compared to the latter, the former description gives the act a place in a pursuit that is more difficult, long-term, and socially relevant than does the latter description. Lin et al. (2014) recently showed that some of the brain regions tracking increasing LOA in the present study showed enhanced activation for action verbs that typically refer to social interactions (e.g., embrace), relative to verbs for individual actions (e.g, walk). They did not, however, measure any of the five dimensions of LOA featured in the present study, nor did they include RT as a nuisance covariate, making it difficult to ascertain to what extent a "sociality" dimension adds anything distinctive to the concept of LOA as presented here. Similarly, it is likely that dimensions (e.g., mind vs. body) that are useful for describing LOA in one conceptual domain (e.g., human action) will prove less useful for describing LOA in other conceptual domains (e.g., mathematical knowledge).

Second, we emphasize that we are not proposing that each of the five dimensions included in this study does not possess a unique and useful meaning on its own. The decision to focus on their shared variance was primarily motivated by the nature of our research question, which was to examine action conceptualization along a single dimension of hierarchical representation. Importantly, this decision was also empirically supported by a PCA on ratings of the five dimensions, which showed that a single factor could explain >91% of their total variance. Hence, we ultimately restricted our analyses to the derived LOA factor and did not examine any of the dimensions individually. While these dimensions were highly correlated in the present study and may naturally correlate in language about actions, there is at least some evidence that they can be dissociated experimentally. For instance, recent work provides evidence for unique effects of concreteness, imageability, and valence in the neural responses to action verbs (Skipper and Olson 2014; Vigliocco et al. 2014).

Third, we consider how our experimental task design could be adjusted to allow examination of novel questions regarding the neural bases of hierarchical action representation. In particular, we chose to present each question-answer pair in rapid succession, with the question offset and answer onset divided by a 250 msec blank screen (see **Figure 1B**). This had the benefit of minimizing working memory load and spontaneous answer production during the interval between question and answer presentation. This benefit came with a cost in that it prevented us from being able to separate the responses to question and answer presentation. Given that the magnitude of the Signed LOA Shift is ultimately established by presentation of the answer phrase, future studies may be able to more precisely model the onset of the within-trial shift by introducing an optimal amount of jitter between the question and answer onsets. However, modeling the onset of the shift at answer onset would still lack precision since the question marker (i.e., why vs. how) provides unambiguous information about the sign of the imminent shift. This coupled with the high likelihood of spontaneous answer production in response to question presentation makes it likely that, in fact, the Signed LOA Shift of interest in our final model actually begins prior to answer presentation.

Finally, we identify two points of clarification regarding our claim that the primary function of *why* and *how* questions is to motivate increases or decreases in LOA, respectively. The first regards the possibility that, even when controlling for the stimulus content, the brain states evoked by *why* and *how* questions may be dissociable in subtle ways not detectable using univariate methods. Future studies could test for such effects using multivariate methods such as pattern-information analysis (Kriegeskorte and Kievit 2013). The second regards the misinterpretation, briefly discussed above, that this means that *why* versus *how* is not to be considered a meaningful cognitive distinction. We believe the contribution of the present study is to more precisely specify the distinct effects that these two questions have on cognition. As discussed above, these effects require that the two questions are motivated by the same action stimulus. When this condition is met, it will almost always be the case that a *why* question will yield an answer that is at a higher LOA than will a *how* question. If this condition is not met, the effects will be unreliable and may sometimes reverse entirely. This is because, as noted in the Introduction, each question can be posed at any LOA, making it possible for a *why* question to yield an answer that rests at a lower LOA (Q: *Why* grip handlebars? A: Ride a bike.) than an answer to a *how* question (Q: *How to* stay healthy? A: Get exercise.). By way of summarizing, we offer the simple analogy:

asking *why* versus *how* is akin to pressing the up versus down buttons when calling an elevator. As long as you're starting from the same floor, you can rest assured that pressing up will almost always put you on a higher floor than will pressing down.

Conclusion

We used a novel action understanding task with functional MRI to examine the neural basis of the well-documented effects of answering *why* and *how* questions about actions. Our data conclusively demonstrate that these effects can be attributed to the fact that *why* and *how* questions—when asked of the same action—produce systematic changes in action understanding, and do so on what appears to be a single hierarchical dimension, which we refer to as the LOA (Vallacher and Wegner 1985; Vallacher and Wegner 1987). Increases and decreases in LOA tracked dissociable brain networks, consistent with prior work using the Why/How contrast. Our data particularly highlight the role of the dorsomedial PFC and anterior STS in upward shifts in LOA. Such shifts make it possible for people to conceive the "here-and-now" of physical reality—including their own bodies—in abstract terms. This, in turn, gives us the power to appreciate that even the simplest of motor actions can carry information about who we are and what we care about most.

Acknowledgments

The Authors would like to acknowledge Emily Ellsworth for help with data collection; Mike Tyszka and the Caltech Brain Imaging Center for help with the neuroimaging; and the Della Martin Foundation for postdoctoral fellowship support to R.P.S. This work was supported in part by the National Institutes of Health (R01 MH080721-03 to R.A.).

References

Amodio, D, Frith, C. 2006. Meeting of minds: The medial frontal cortex and social cognition. Nat Rev Neurosci 7:268–277.

Andrews-Hanna, J, Reidler, J, Sepulcre, J, Poulin, R, Buckner, R. 2010. Functional-anatomic fractionation of the brain's default network. Neuron 65:550–562.

Andrews-Hanna, J, Smallwood, J, Spreng, R. 2014. The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. Ann N Y Acad Sci 1316:29–52.

Ashburner, J. 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38:95–113.

Badre, D, Poldrack, R, Pare-Blagoev, E, Insler, R, Wagner, A. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron 47:907–918.

Bedny, M, Caramazza, A, Grossman, E, Pascual-Leone, A, Saxe, R. 2008. Concepts are more than percepts: The case of action verbs. J Neurosci 28:11347–11353.

Bedny, M, Caramazza, A, Pascual-Leone, A, Saxe, R. 2012. Typical neural representations of action verbs develop without vision. Cereb Cortex 22:286–293.

Binder, J, Desai, R. 2011. The neurobiology of semantic memory. Trends Cogn Sci 15:527–536.

Binder, J, Desai, R, Graves, W, Conant, L. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767–2796.

Binder, J, Westbury, C, Mckiernan, K, Possing, E, Medler, D. 2005. Distinct brain systems for processing concrete and abstract concepts. J Cogn Neurosci 17:905–917.

Brainard, DH. 1997. The psychophysics toolbox. Spatial Vision 10:433–436.

Buckner, R, Andrews-Hanna, J, Schacter, D. 2008. The brain's default network: Anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buckner, R, Carroll, D. 2007. Self-projection and the brain. Trends Cogn Sci 11:49–57.

Carrington, S, Bailey, A. 2009. Are there theory of mind regions in the brain? A review of the neuroimaging literature. Hum Brain Mapp 30:2313–2335.

Caspers, S, Zilles, K, Laird, A, Eickhoff, S. 2010. Ale meta-analysis of action observation and imitation in the human brain. NeuroImage 50:1148–1167.

Chen, E, Widick, P, Chatterjee, A. 2008. Functional-anatomical organization of predicate metaphor processing. Brain Lang 107:194–202.

Deen, B, McCarthy, G. 2010. Reading about the actions of others: biological motion imagery and action congruency influence brain activity. Neuropsychologia 48:1607–1615.

Denny, B, Kober, H, Wager, T, Ochsner, K. 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J Cogn Neurosci 24:1742–1752.

Diedrichsen, J, Shadmehr, R. 2005. Detecting and adjusting for artifacts in fmri time series data. NeuroImage 27:624–634.

Ferstl, EC, Neumann, J, Bogler, C, von Cramon, DY. 2008. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Hum Brain Mapp 29:581–593.

Ferstl, EC, von Cramon, DY. 2001. The role of coherence and cohesion in text comprehension: an event-related fMRI study. Brain Res Cogn Brain Res 11:325–340.

Freitas, AL, Gollwitzer, P, Trope, Y. 2004. The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. J Exp Soc Psychol 40:739–752.

Fujita, K, Trope, Y, Liberman, N, Levin-Sagi, M. 2006. Construal levels and self-control. J Pers Soc Psychol 90:351–367.

Gallagher, HL, Frith, CD. 2003. Functional imaging of 'theory of mind'. Trends Cogn Sci 7:77–83.

Gilead M, Liberman N, Maril A (2014) From mind to matter: neural correlates of abstract and concrete mindsets. Soc Cogn Affect Neurosci 9(5):638–645.

Goldberg, R, Perfetti, C, Fiez, J, Schneider, W. 2007. Selective retrieval of abstract semantic knowledge in left prefrontal cortex. J Neurosci 27:3790–3798.

Gollwitzer, PM, Heckhausen, H, Steller, B. 1990. Deliberative and implemental mind-sets: Cognitive tuning toward congruous thoughts and information. J Personality and Social Psychology 59:1119–1127.

Grinband, J, Wager, T, Lindquist, M, Ferrera, V, Hirsch, J. 2008. Detection of time-varying signals in event-related fmri designs. NeuroImage 43:509–520.

Hassabis, D, Maguire, E. 2007. Deconstructing episodic memory with construction. Trends Cogn Sci 11:299–306.

Hoffman, P, Jefferies, E, Lambon Ralph, M. 2010. Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: Convergent neuropsychological and repetitive tms evidence. J Neurosci 30:15450–15456.

Humphreys, GF, Newling, K, Jennings, C, Gennari, SP. 2013. Motion and actions in language: semantic representations in occipito-temporal cortex. Brain Lang 125:94–105.

Jack, A, Dawson, A, Begany, K, Leckie, R, Barry, K, Ciccia, A, Snyder, A. 2012. Fmri reveals reciprocal inhibition between social and physical cognitive domains. NeuroImage 66C:385–401.

Kemmerer, D. 2015. Visual and Motor Features of the Meanings of Action Verbs: A Cognitive Neuroscience Perspective. In: de Almeida, RG, Manouilidou, C, editors. Cognitive science perspectives on verb representation and processing. New York:

Springer. p. 189–212.

Kemmerer, D, Rudrauf, D, Manzel, K, Tranel, D. 2012. Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. Cortex 48:826–848.

Kozak, M, Marsh, A, Wegner, D. 2006. What do I think you're doing? Action identification and mind attribution. J Pers Soc Psychol 90:543–555.

Kriegeskorte, N, Kievit, R. 2013. Representational geometry: Integrating cognition, computation, and the brain. Trends Cogn Sci 17:401–412.

Liberman, N, Trope, Y. 2014. Traversing psychological distance. Trends Cogn Sci 18:364–369.

Lieberman, M. 2010. Social cognitive neuroscience. In: Fiske, ST, Gilbert, DT, Lindzey, G, editors. Handbook of Social Psychology. 5th ed. New York: McGraw-Hill. p. 143–193.

Lin, N, Bi, Y, Zhao, Y, Luo, C, Li, X. 2014. The theory-of-mind network in support of action verb comprehension: Evidence from an fMRI study. Brain Lang 141C:1–10.

Mar, R. 2011. The neural bases of social cognition and story comprehension. Annu Rev Psychol 62:103–134.

Mar, RA, Oatley, K. 2008. The function of fiction is the abstraction and simulation of social experience. Perspect on Psych Sci 3:173–192.

Molenberghs, P, Cunnington, R, Mattingley, J. 2012. Brain regions with mirror properties: A meta-analysis of 125 human fmri studies. Neurosci Biobehav Rev 36:341–349.

Nijhof, AD, Willems, RM. 2015. Simulating fiction: individual differences in literature comprehension revealed with FMRI. PLoS One 10:e0116492.

Pulvermuller, F. 2013. How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. Trends Cogn Sci 17:458–470.

Raichle, M, Macleod, A, Snyder, A, Powers, W, Gusnard, D, Shulman, G. 2001. A default mode of brain function. Proc Natl Acad Sci U S A 98:676–682.

Raposo, A, Mendes, M, Marques, J. 2012. The hierarchical organization of semantic memory: Executive function in the processing of superordinate concepts. NeuroImage 59:1870–1878.

Rizzolatti, G, Cattaneo, L, Fabbri-Destro, M, Rozzi, S. 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. Physiol Rev 94:655–706.

Satpute, A, Badre, D, Ochsner, K. 2014. Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. Cereb Cortex 24:1269–1277.

Saxe, R. 2006. Uniquely human social cognition. Curr Opin Neurobiol 16:235–239.

Saygin, AP, McCullough, S, Alac, M, Emmorey, K. 2010. Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. J Cogn

Neurosci 22:2480–2490.

Schacter, D, Addis, D, Hassabis, D, Martin, V, Spreng, R, Szpunar, K. 2012. The future of memory: Remembering, imagining, and the brain. Neuron 76:677–694.

Schurz, M, Radua, J, Aichhorn, M, Richlan, F, Perner, J. 2014. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. Neurosci Biobehav Rev 42C:9–34.

Shrout, PE, Fleiss, JL. 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol bulletin 86:420.

Skipper, L, Olson, I. 2014. Semantic memory: Distinct neural representations for abstractness and valence. Brain Lang 130:1–10.

Spreng, R, Mar, R, Kim, A. 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. J Cogn Neurosci 21:489–510.

Spunt, R, Adolphs, R. 2014. Validating the why/how contrast for functional mri studies of theory of mind. Neuroimage 99:301–311.

Spunt, R, Falk, E, Lieberman, M. 2010. Dissociable neural systems support retrieval of how and why action knowledge. Psychol Sci 21:1593–1598.

Spunt, R, Lieberman, M. 2012a. An integrative model of the neural systems supporting the comprehension of observed emotional behavior. NeuroImage 59:3050–3059.

Spunt, R, Lieberman, M. 2012b. Dissociating modality-specific and supramodal neural systems for action understanding. J Neurosci 32:3575–3583.

Spunt, RP, Meyer, ML, Lieberman, MD. 2015. The default mode of human brain function primes the intentional stance. J Cogn Neurosci 27:1116–1124.

Spunt, R, Satpute, A, Lieberman, M. 2011. Identifying the what, why, and how of an observed action: An fmri study of mentalizing and mechanizing during action observation. J Cogn Neurosci 23:63–74.

Trope, Y, Liberman, N. 2010. Construal-level theory of psychological distance. Psychol Rev 117:440–463.

Urgesi, C, Candidi, M, Avenanti, A. 2014. Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. Front Hum Neurosci 8:344.

Vallacher, RR, Wegner, DM. 1985. A theory of action identification. New York: Psychology Press.

Vallacher, RR, Wegner, DM. 1987. What do people think they're doing? Action identification and human behavior. Psychol Rev 94:3–15.

Van Overwalle, F. 2011. A dissociation between social mentalizing and general reasoning. NeuroImage 54:1589–1599.

Van Overwalle, F, Baetens, K. 2009. Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. NeuroImage 48:564–584.

Vigliocco, G, Kousta, S, Della Rosa, P, Vinson, D, Tettamanti, M, Devlin, J, Cappa, S. 2014. The neural representation of abstract words: The role of emotion. Cereb Cortex 24:1767–1777.

Wallentin, M, Nielsen, AH, Vuust, P, Dohn, A, Roepstorff, A, Lund, TE. 2011. Bold response to motion verbs in left posterior middle temporal gyrus during story comprehension. Brain and language 119:221–225.

Wang, J, Conder, J, Blitzer, D, Shinkareva, S. 2010. Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. Hum Brain Mapp 31:1459–1468.

Watson, CE, Cardillo, ER, Ianni, GR, Chatterjee, A. 2013. Action concepts in the brain: an activation likelihood estimation meta-analysis. J Cogn Neurosci 25:1191–1205.

Table 1
Pearson correlations among all phrase-level parameters.

Parameter	1	2	3	4	5	6	7	8	9	10
1 Level of Abstraction	-	0.99	0.97	0.96	0.92	0.77	0.16	-0.25	-0.03	0.14
2 Abstract/Concrete	0.99	-	0.96	0.95	0.89	0.74	0.17	-0.26	-0.06	0.14
3 Non-Imageable/Imageable	0.97	0.96	-	0.93	0.83	0.76	0.14	-0.24	-0.07	0.15
4 Broad/Specific	0.96	0.95	0.93	-	0.82	0.74	0.16	-0.28	-0.07	0.13
5 Mind/Body	0.92	0.89	0.83	0.82	-	0.64	0.14	-0.15	0.05	0.1
6 Valenced/Neutral	0.77	0.74	0.76	0.74	0.64	-	0.13	-0.2	0.13	0.1
7 Number of Characters	0.16	0.17	0.14	0.16	0.14	0.13	-	0.04	-0.35	0.02
8 Number of Words	-0.25	-0.26	-0.24	-0.28	-0.15	-0.2	0.04	-	0.45	0
9 Content Word Frequency	-0.03	-0.06	-0.07	-0.07	0.05	0.13	-0.35	0.45	-	-0.06
10 Response Time (secs)	0.14	0.14	0.15	0.13	0.1	0.1	0.02	0	-0.06	-

Note. The first parameter - Level of Abstraction - represents the first eigenvariate from a principal components analysis on parameters 2-6. Parameters 7-10 were included as parametric nuisance covariates in all single-subject fMRI analyses (represented at the trial-level by averaging the parameter value for the question and answer phrases). Note that because RT is a trial-level parameter and is contingent on participant behavior, all RT correlation values were computed as the mean within-subject correlation between RT to each trial and the average parameter value across the question and answer phrases within each trial.

Table 2
Means for performance, lexical, and Level of Abstraction (LOA) parameters for both Why and How trials.

Parameter Group / Parameter Name	Why	How	$P_{difference}$
Performance			
Percent Correct	94.72	94.14	.463
Response Time (s)	1.00	1.02	.296
Lexical			
Number of Characters	11.90	12.02	.712
Number of Words	2.46	2.42	.700
Content Word Frequency	3.80	3.81	.957
Level of Abstraction			
Question-and-Answer	.44	.44	.964
Question	.34	.59	<.001
Answer	.57	.32	<.001
Answer - Question	.23	-.27	<.001

Note. The final column lists the p-value from a t-test comparing the Why and How trials on each parameter. Paired samples t-tests were used to compare the performance parameters while independent samples t-tests were used for the remaining parameters. To facilitate comparability, the LOA parameters were computed after rescaling the LOA across phrases to 0-1.

Table 3
Results of two-tailed paired-sample t-tests on percent signal change in the set of a priori regions of interest (ROI) for the two models examining the independent effects of Why/How manipulation and Trialwise Level of Abstraction (LOA).

	Model 1						Model 2					
ROI Source	Question: Why > How			LOA: High > Low			Question: Why > How			Increasing LOA		
Region Name	*t*	*p*	95% CI	*t*	*p*	95% CI	*t*	*p*	95% CI	*t*	*p*	95% CI
Why > How												
Dorsomedial PFC	.54	.595	[-.12 .17]	3.84	.001	[.21 .65]	.63	.535	[-.1 .15]	4.00	.001	[.04 .11]
Ventromedial PFC	.24	.816	[-.11 .15]	4.31	<.001	[.17 .42]	.57	.577	[-.06 .18]	3.32	.004	[.02 .08]
Anterior STS	.93	.366	[-.02 .07]	5.67	<.001	[.09 .19]	1.21	.242	[-.02 .07]	4.49	<.001	[.02 .04]
TPJ	.77	.454	[-.06 .13]	4.06	<.001	[.1 .27]	.41	.688	[-.08 .12]	2.77	.014	[.01 .05]
PCC	.24	.811	[-.07 .09]	4.01	.001	[.08 .24]	.21	.840	[-.08 .09]	2.95	.009	[.01 .05]
How > Why												
Dorsal PMC	-.57	.575	[-.09 .06]	-1.11	.283	[-.16 .03]	-1.15	.269	[-.11 .04]	-1.19	.252	[-.06 0]
Ventral PMC	-.08	.939	[-.07 .07]	-2.42	.028	[-.27 -.05]	-.65	.526	[-.09 .04]	-2.99	.009	[-.07 -.02]
Posterior MTG	.05	.960	[-.07 .08]	-3.57	.003	[-.36 -.12]	-.35	.734	[-.09 .06]	-3.99	.001	[-.09 -.03]
Rostral IPL	.62	.543	[-.03 .08]	-2.52	.023	[-.27 -.05]	.24	.814	[-.04 .07]	-2.65	.018	[-.06 -.01]
Dorsal Precuneus	-.21	.837	[-.11 .07]	1.57	.135	[-.01 .2]	-.62	.547	[-.13 .06]	1.57	.136	[-.01 .04]

Note. Both models featured the same two covariates of interest: variation in the question posed in each trial (Why vs. How) and trialwise variation in the LOA, computed as the mean LOA for the Question and Answer phrases appearing within the trial. Whereas Model 1 examines the trialwise LOA as a categorical variable with two levels (High vs. Low), the Model 2l examines it as a continuous parametric modulator of the response to each trial (see Methods for further details regarding model construction and contrast calculations). The examined ROIs are displayed in Fig. 2 and further details can be found in the main text and in Table S2. Confidence intervals (CIs) for the effect size in each comparison were estimated using the bias corrected and accelerated percentile method (10,000 random samples with replacement; implemented using the BOOTCI function in MATLAB). TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus.

Table 4
Results of two-tailed paired-sample t-tests on percent signal change in the set of a priori regions of interest (ROI) for the model examining the independent effects of the Prepotent Level of Abstraction (LOA) and the Signed Shift in LOA for each trial.

	Effect Name					
ROI Source	Prepotent LOA			Signed LOA Shift		
Region Name	*t*	*p*	95% CI	*t*	*p*	95% CI
Why > How						
Dorsomedial PFC	4.53	<.001	[.04 .10]	2.87	.011	[.01 .07]
Ventromedial PFC	3.10	.007	[.02 .08]	2.81	.013	[.02 .07]
Anterior STS	4.47	<.001	[.02 .04]	4.02	<.001	[.01 .03]
TPJ	3.06	.008	[.01 .06]	2.00	.062	[.00 .05]
PCC	2.75	.014	[.01 .05]	2.63	.018	[.01 .04]
How > Why						
Dorsal PMC	-1.15	.266	[-.06 .00]	-0.96	.351	[-.03 .01]
Ventral PMC	-2.78	.014	[-.06 -.01]	-1.83	.086	[-.05 -.00]
Posterior MTG	-4.17	<.001	[-.08 -.03]	-1.88	.078	[-.04 -.00]
Rostral IPL	-2.70	.016	[-.06 -.01]	-0.95	.358	[-.03 .01]
Dorsal Precuneus	1.66	.117	[-.01 .05]	1.99	.064	[.00 .04]

Note. The model featured two covariates of interest: the LOA of the Question phrase (i.e., Prepotent LOA) and the signed shift in LOA introduced in the Answer phrase, computed by subtracting the LOA of the Question phrase from that of the Answer phrase (see Methods for further details regarding model construction and contrast calculations). The examined ROIs are displayed in Fig. 2 and further details can be found in the main text and in Table S2. Confidence intervals (CIs) for the effect size in each comparison were estimated using the bias corrected and accelerated percentile method (10,000 random samples with replacement; implemented using the BOOTCI function in MATLAB). TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus.

Figure Legends

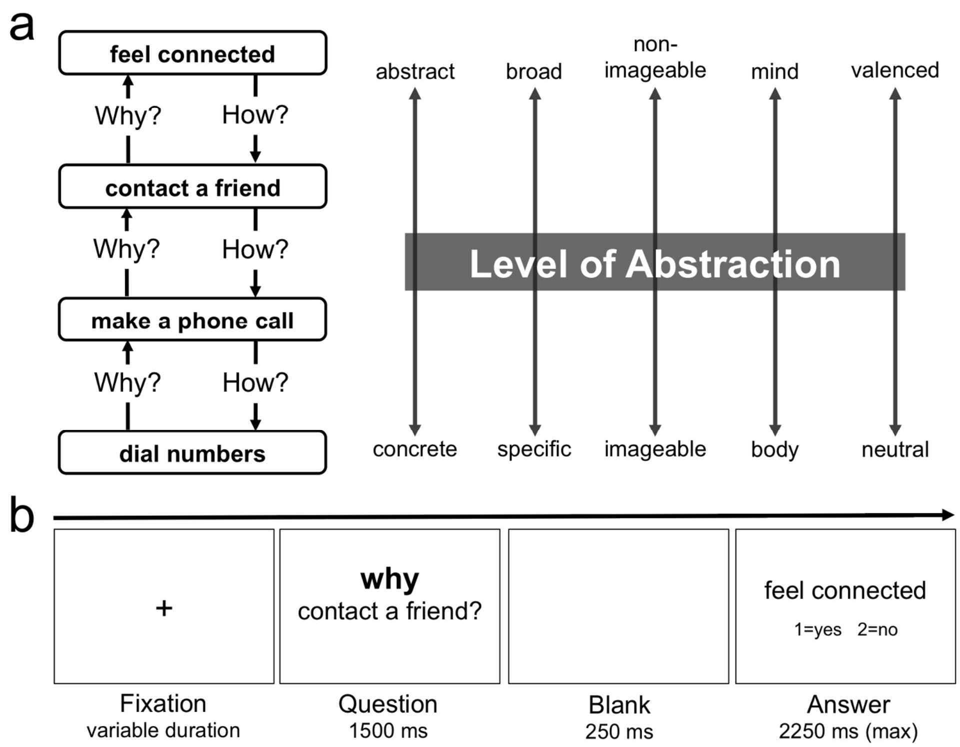
Figure 1. Experimental Design. (A) One of 25 four-level action hierarchies featured in the experimental task. For each four-level hierarchy the action described by a phrase at one level (e.g. "make a phone call") was both a commonly accepted motive for performing the action described by the phrase at the level immediately below it (e.g., "dial numbers") and a commonly accepted means for performing the action described by the phrase at the level immediately above it (e.g., "contact a friend"). As shown to the right of the example and described further in the text all phrases were normed on five dimensions used to derive a single factor describing each phrase's level of abstraction (LOA). (B) Structure of one of the trials formed by pairing phrases at contiguous levels of the hierarchy. Trials began with an action phrase embedded in either a *why* (shown) or *how* question and concluded with a different action phrase presented as a possible answer. Answer phrases were presented for a maximum duration of 2250 ms. Once the participant responded, the screen was replaced by a fixation cross until the onset of the next trial.

Figure 2. Left hemisphere cortical regions reliably modulated by the Why/How contrast in prior work. See *Methods* and Table S2 for further details. TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus.

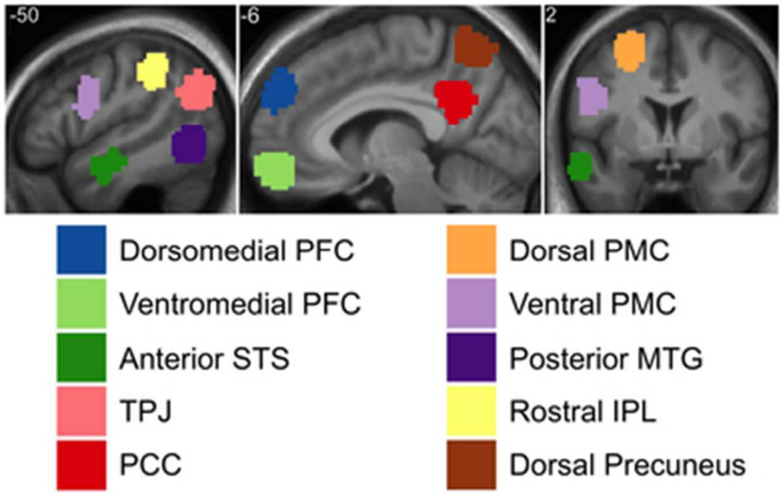
Figure 3. Whole-brain surface renderings of regional activity associated with Trialwise Level of Abstraction (LOA), computed as the mean of the question and answer phrases appearing in each trial. The results in (A) show regions modulated in the categorical High > Low LOA contrast from Model 1, while the results in (B) show regions modulated by the continuous Trialwise LOA parameter from Model 2. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of .05, and their locations are reported in Table S3. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.

Figure 4. (A) Scatterplot showing the relationship between the Prepotent LOA and the signed Within-Trial Shift in LOA across Why (black markers) and How (white markers) trials. The Prepotent LOA refers to the LOA of the action phrase appearing in the question at the beginning of each trial. The signed Within-Trial Shift in LOA is computed for each trial by subtracting the

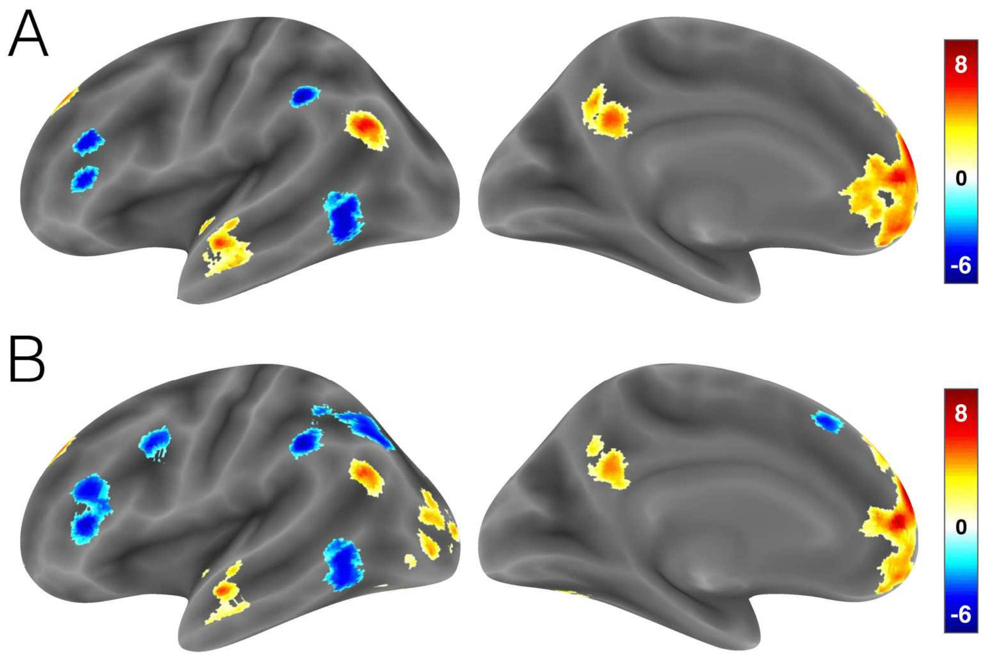
Prepotent LOA from the LOA of the action phrase appearing in the presented answer. Positive shift trials induced an upward change in LOA, while negative shift trials induced a downward change in LOA. To facilitate comparability, the LOA dimension has been rescaled to 0-1. (B) Regions uniquely associated with the Prepotent LOA parameter when controlling for the signed Within-Trial Shift in LOA. (C) Regions uniquely associated with the signed Within-Trial Shift in LOA parameter. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of .05, and their locations are reported in Table S4. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.



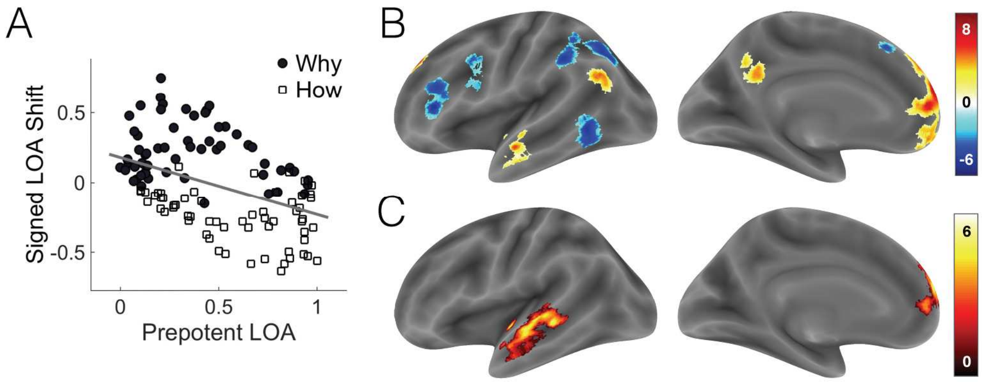
Experimental Design. (A) One of 25 four-level action hierarchies featured in the experimental task. For each four-level hierarchy the action described by a phrase at one level (e.g. "make a phone call") was both a commonly accepted motive for performing the action described by the phrase at the level immediately below it (e.g., "dial numbers") and a commonly accepted means for performing the action described by the phrase at the level immediately above it (e.g., "contact a friend"). As shown to the right of the example and described further in the text all phrases were normed on five dimensions used to derive a single factor describing each phrase's level of abstraction (LOA). (B) Structure of one of the trials formed by pairing phrases at contiguous levels of the hierarchy. Trials began with an action phrase embedded in either a why (shown) or how question and concluded with a different action phrase presented as a possible answer. Answer phrases were presented for a maximum duration of 2250 ms. Once the participant responded, the screen was replaced by a fixation cross until the onset of the next trial.
169x127mm (300 x 300 DPI)



Left hemisphere cortical regions reliably modulated by the Why/How contrast in prior work. See Methods and Table S2 for further details. TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus.
33x21mm (300 x 300 DPI)



Whole-brain surface renderings of regional activity associated with Trialwise Level of Abstraction (LOA), computed as the mean of the question and answer phrases appearing in each trial. The results in (A) show regions modulated in the categorical High > Low LOA contrast from Model 1, while the results in (B) show regions modulated by the continuous Trialwise LOA parameter from Model 2. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of .05, and their locations are reported in Table S3. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.
184x123mm (300 x 300 DPI)



(A) Scatterplot showing the relationship between the Prepotent LOA and the signed Within-Trial Shift in LOA across Why (black markers) and How (white markers) trials. The Prepotent LOA refers to the LOA of the action phrase appearing in the question at the beginning of each trial. The signed Within-Trial Shift in LOA is computed for each trial by subtracting the Prepotent LOA from the LOA of the action phrase appearing in the presented answer. Positive shift trials induced an upward change in LOA, while negative shift trials induced a downward change in LOA. To facilitate comparability, the LOA dimension has been rescaled to 0-1. (B) Regions uniquely associated with the Prepotent LOA parameter when controlling for the signed Within-Trial Shift in LOA. (C) Regions uniquely associated with the signed Within-Trial Shift in LOA parameter. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of .05, and their locations are reported in Table S4. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.
227x93mm (300 x 300 DPI)

Supplementary Materials for

The Neural Basis of Conceptualizing the Same Action at Different Levels of Abstraction

[1]Robert P. Spunt, [2]David Kemmerer, [1]Ralph Adolphs

Author Affiliations:

[1]California Institute of Technology

[2]Purdue University

Corresponding Author:

Robert P. Spunt

California Institute of Technology

1200 E. California Blvd.

Pasadena, CA 91125

Phone: (818) 645-8188

Email: spunt@caltech.edu

Supplementary Methods

Sample Size Determination

A power analysis of the Why > How contrast from Study 3 of Spunt and Adolphs (2014) indicated the present study's sample size was sufficiently large to test our primary hypotheses. We based our power analysis on that study because it was conducted using the same scanner, sequence, and preprocessing pipeline as those used in the present study. The power analysis was performed using the open-source software fmripower (http://mumford.fmripower.org/) operating in MATLAB (version 2014b; MathWorks Inc., Natick, MA, USA). Across the 10 regions used in the region-of-interest anlayses reported in the main text (see Figure 2 in main text and Table S2 below for ROI details), 90% detection power for the Why > How contrast required an average of only 8.90 subjects (SD = 3.84, Maximum = 15).

Table S1

Action phrases included in the study listed in the order of increasing Level of Abstraction (LOA). To facilitate comparability, all semantic dimensions (columns 1-6) have been rescaled to 0 – 1.

Phrase	LOA	Abstract	Non-imageable	Broad	Mind	Arousing	N Chars	N Words	Freq
turn a doorknob	0.00	0.00	0.01	0.00	0.00	0.00	13	3	2.99
click a mouse	0.02	0.00	0.02	0.01	0.08	0.05	11	3	2.62
grab a pencil	0.03	0.02	0.04	0.02	0.05	0.01	11	3	3.12
crack eggs	0.04	0.02	0.00	0.04	0.06	0.13	9	2	3.09
clap hands	0.04	0.00	0.01	0.03	0.03	0.45	9	2	3.16
raise a hand	0.05	0.04	0.02	0.06	0.03	0.13	10	3	3.76
fasten seatbelt	0.05	0.01	0.02	0.01	0.08	0.36	14	2	1.96
brush one's hair	0.05	0.03	0.05	0.05	0.05	0.13	14	3	3.25
cut paper	0.07	0.04	0.05	0.10	0.08	0.03	8	2	3.87
take aspirin	0.07	0.03	0.05	0.03	0.14	0.16	11	2	3.79
wash one's hands	0.07	0.07	0.04	0.04	0.06	0.32	14	3	3.66
open a bottle	0.07	0.04	0.04	0.14	0.05	0.11	11	3	3.72
drink coffee	0.07	0.03	0.02	0.08	0.09	0.35	11	2	3.83
use a comb	0.08	0.03	0.05	0.15	0.06	0.09	8	3	3.18
sign one's name	0.08	0.06	0.04	0.05	0.15	0.06	13	3	4.02
use scissors	0.08	0.04	0.06	0.14	0.10	0.04	11	2	3.36
throw away trash	0.08	0.06	0.04	0.11	0.06	0.30	14	3	3.74
open a magazine	0.09	0.05	0.06	0.15	0.06	0.16	13	3	3.64
lift weights	0.09	0.04	0.07	0.18	0.00	0.45	11	2	2.57
go jogging	0.10	0.08	0.08	0.10	0.04	0.47	9	2	3.52
clip coupons	0.10	0.04	0.05	0.14	0.13	0.20	11	2	2.05
drink wine	0.10	0.04	0.05	0.10	0.15	0.36	9	2	3.64
lift a glass	0.11	0.02	0.05	0.30	0.03	0.12	10	3	3.25
use a remote	0.11	0.08	0.04	0.16	0.14	0.04	10	3	3.52
press buttons	0.12	0.05	0.08	0.25	0.07	0.06	12	2	2.83
point a camera	0.12	0.09	0.08	0.12	0.14	0.19	12	3	3.24
write a check	0.12	0.06	0.04	0.14	0.22	0.18	11	3	3.65
use a brush	0.12	0.08	0.06	0.24	0.08	0.12	9	3	3.39
wear ties	0.13	0.06	0.16	0.20	0.07	0.18	8	2	3.11
stir ingredients	0.13	0.07	0.09	0.25	0.07	0.16	15	2	2.21
use chalk	0.13	0.08	0.06	0.24	0.12	0.08	8	2	3.22
enter a classroom	0.13	0.10	0.10	0.22	0.07	0.16	15	3	2.77
open a browser	0.14	0.10	0.08	0.13	0.20	0.09	12	3	2.52
dial numbers	0.14	0.09	0.06	0.25	0.14	0.03	11	2	2.91
take pills	0.15	0.08	0.08	0.26	0.12	0.29	9	2	4.04
pack a bag	0.17	0.10	0.11	0.28	0.18	0.14	8	3	3.38
consume alcohol	0.17	0.08	0.14	0.22	0.18	0.42	14	2	2.48
make a phone call	0.19	0.12	0.06	0.18	0.35	0.16	14	4	4.30
use a camera	0.20	0.11	0.09	0.24	0.27	0.26	10	3	3.82
give applause	0.20	0.19	0.09	0.21	0.14	0.67	12	2	3.77
take a picture	0.20	0.14	0.06	0.24	0.27	0.34	12	3	4.40
consume medicine	0.20	0.12	0.15	0.34	0.16	0.31	15	2	2.63
consume caffeine	0.21	0.14	0.27	0.28	0.13	0.29	15	2	2.10
mix a cocktail	0.21	0.09	0.19	0.36	0.17	0.37	12	3	2.73
play piano	0.22	0.11	0.12	0.16	0.40	0.46	9	2	3.63
make breakfast	0.22	0.12	0.12	0.36	0.20	0.42	13	2	4.18
use credit cards	0.23	0.14	0.12	0.29	0.29	0.37	14	3	3.64
serve alcohol	0.23	0.13	0.22	0.39	0.15	0.27	12	2	3.10
watch television	0.27	0.17	0.04	0.23	0.52	0.32	15	2	3.70
dress formally	0.27	0.22	0.20	0.37	0.19	0.35	13	2	2.54
write an email	0.28	0.16	0.15	0.23	0.53	0.11	12	3	2.88

prepare meals	0.29	0.13	0.24	0.50	0.25	0.36	12	2	2.92
go shopping	0.32	0.24	0.23	0.47	0.25	0.42	10	2	4.09
make a list	0.33	0.18	0.17	0.39	0.52	0.20	9	3	4.22
paint a picture	0.34	0.22	0.19	0.39	0.44	0.50	13	3	3.41
build muscles	0.34	0.24	0.47	0.53	0.03	0.65	12	2	2.98
attend a school	0.35	0.26	0.27	0.38	0.40	0.41	13	3	3.54
make a purchase	0.37	0.23	0.22	0.61	0.35	0.27	13	3	3.54
ingest nutrients	0.38	0.26	0.60	0.57	0.06	0.37	15	2	1.47
provide food	0.41	0.26	0.40	0.59	0.26	0.62	11	2	3.42
follow recipes	0.42	0.33	0.30	0.41	0.56	0.27	13	2	2.93
hold a party	0.43	0.33	0.35	0.47	0.41	0.61	10	3	4.19
read articles	0.43	0.28	0.20	0.42	0.75	0.19	12	2	3.29
arrive on time	0.45	0.42	0.53	0.27	0.41	0.68	12	3	3.99
kill germs	0.45	0.39	0.62	0.50	0.19	0.49	9	2	3.25
contact a friend	0.45	0.35	0.36	0.40	0.51	0.66	14	3	3.71
send a message	0.46	0.38	0.31	0.59	0.48	0.15	12	3	3.82
make a donation	0.47	0.34	0.36	0.48	0.51	0.73	13	3	3.55
satisfy hunger	0.48	0.47	0.40	0.56	0.29	0.70	13	2	2.46
give a lecture	0.48	0.36	0.38	0.46	0.66	0.26	12	3	3.69
write a poem	0.49	0.35	0.39	0.38	0.74	0.47	10	3	3.33
give to charity	0.50	0.36	0.51	0.47	0.53	0.76	13	3	3.85
get fit	0.51	0.44	0.52	0.72	0.17	0.84	6	2	4.16
teach a class	0.52	0.42	0.36	0.50	0.64	0.45	11	3	3.67
write a song	0.52	0.36	0.48	0.41	0.73	0.51	10	3	3.74
participate in class	0.53	0.45	0.37	0.51	0.64	0.40	18	3	3.06
get a job	0.54	0.41	0.58	0.53	0.51	0.71	7	3	4.85
show cleanliness	0.57	0.55	0.58	0.64	0.35	0.60	15	2	2.85
finish a job	0.58	0.46	0.68	0.62	0.43	0.75	10	3	3.98
get a degree	0.61	0.40	0.59	0.60	0.71	0.84	10	3	4.12
make art	0.64	0.55	0.39	0.85	0.55	0.67	7	2	4.19
help a person	0.65	0.57	0.56	0.77	0.47	0.92	11	3	4.31
save money	0.66	0.52	0.65	0.55	0.71	0.86	9	2	4.21
get a promotion	0.66	0.53	0.77	0.50	0.67	0.98	13	3	4.06
live a long life	0.67	0.63	0.81	0.63	0.41	0.95	13	4	4.32
obey laws	0.69	0.63	0.73	0.58	0.69	0.40	8	2	2.80
entertain others	0.72	0.71	0.56	0.81	0.54	0.64	15	2	3.10
teach others	0.73	0.66	0.54	0.77	0.70	0.69	11	2	3.64
reduce pain	0.73	0.68	0.87	0.74	0.46	0.62	10	2	3.02
make a friend	0.73	0.70	0.73	0.62	0.63	0.85	11	3	4.59
avoid diseases	0.75	0.71	0.99	0.77	0.40	0.58	13	2	2.69
seek advice	0.76	0.72	0.67	0.75	0.76	0.25	10	2	3.18
support a cause	0.77	0.71	0.73	0.74	0.68	0.71	13	3	3.60
show an ability	0.77	0.70	0.74	0.91	0.52	0.64	13	3	3.59
stay alert	0.77	0.76	0.71	0.68	0.76	0.44	9	2	3.19
express a talent	0.78	0.74	0.71	0.90	0.54	0.71	14	3	2.93
show courtesy	0.79	0.80	0.62	0.75	0.69	0.79	12	2	3.41
show appreciation	0.81	0.82	0.62	0.81	0.67	0.82	16	2	3.31
entertain oneself	0.82	0.75	0.70	0.96	0.68	0.58	16	2	2.27
express doubt	0.84	0.84	0.74	0.76	0.82	0.42	12	2	2.88
satisfy a need	0.85	0.82	0.83	0.96	0.56	0.64	12	3	2.99
impress others	0.85	0.85	0.85	0.89	0.59	0.56	13	2	3.21
plan ahead	0.85	0.78	0.76	0.76	0.90	0.57	9	2	2.14
feel well	0.86	0.87	0.76	0.87	0.65	0.87	8	2	3.85
pass the time	0.86	0.88	0.85	0.99	0.60	0.14	11	3	4.30
show competence	0.87	0.84	0.85	0.82	0.71	0.77	14	2	2.92
show humility	0.87	0.84	0.81	0.81	0.81	0.57	12	2	3.14
share ideas	0.88	0.85	0.74	0.85	0.81	0.65	10	2	3.36
avoid mistakes	0.89	0.83	0.99	0.83	0.74	0.53	13	2	3.07
express an idea	0.89	0.88	0.74	0.88	0.80	0.62	13	3	3.50
spread joy	0.90	0.87	0.78	0.95	0.68	0.95	9	2	3.08

spread knowledge	0.90	0.86	0.82	0.87	0.79	0.77	15	2	3.11
show creativity	0.91	0.86	0.77	1.00	0.75	0.76	14	2	3.09
avoid boredom	0.91	0.88	0.92	0.93	0.76	0.37	12	2	2.57
achieve success	0.92	0.86	0.87	0.96	0.68	1.00	14	2	2.86
feel sociable	0.92	0.94	0.81	0.84	0.83	0.63	12	2	3.10
show ambition	0.92	0.89	0.96	0.85	0.78	0.68	12	2	3.32
show intelligence	0.93	0.87	0.83	0.84	0.89	0.78	16	2	3.59
feel productive	0.93	0.91	0.86	0.85	0.78	0.87	14	2	3.29
improve oneself	0.93	0.92	0.83	0.99	0.65	0.93	14	2	2.33
gain knowledge	0.94	0.87	0.76	0.91	0.93	0.87	13	2	2.91
keep a memory	0.94	0.91	0.93	0.72	1.00	0.58	11	3	3.97
show intellect	0.96	0.88	0.85	0.94	0.89	0.76	13	2	3.12
reduce worry	0.96	0.92	1.00	0.84	0.86	0.62	11	2	2.40
feel nostalgia	0.96	1.00	0.84	0.79	0.96	0.54	13	2	3.04
feel secure	0.97	0.95	0.86	0.88	0.88	0.86	10	2	3.62
feel connected	0.99	0.97	0.92	0.95	0.85	0.72	13	2	3.21
increase wisdom	1.00	0.94	0.90	0.91	0.95	0.88	14	2	2.62

Table S2
Details for Regions of Interest (ROI) used in the ROI analyses presented in the main text.

Source Name / Region Name	Extent	Peak MNI Coordinates x	y	z
Why > How				
TPJ	917	-48	-66	30
Ventromedial PFC	1121	0	57	-12
Anterior STS	1102	-57	-9	-18
PCC	1024	-3	-48	30
Dorsomedial PFC	964	-6	57	36
How > Why				
Rostral IPL	1055	-42	-39	42
Dorsal PMC	817	-24	0	54
Posterior MTG	923	-54	-60	-3
Ventral PMC	677	-51	6	27
Dorsal Precuneus	975	-6	-60	60

Note. Please see the Methods for details on ROI definition. ROIs are visually depicted in Figure 2 in the main text. TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus; x, y, and z = Montreal Neurological Institute (MNI) coordinates in the left-right, anterior-posterior, and inferior-superior dimensions, respectively.

Table S3
Peaks from clusters observed in whole-brain analyses for the two models examining trialwise LOA and the Question (Why vs. How) at each trial.

Analysis Name			MNI		
Region Label	Extent	*t*	x	y	z
Factorial Model					
[(High-Why + High-How) > (Low-Why + Low-How)]					
L Middle Temporal Gyrus	254	8.66	-60	-12	-14
L Superior Medial Gyrus	1850	7.29	-12	52	8
L Rectal Gyrus	-	6.50	-2	44	-18
R Mid Orbital Gyrus	-	5.62	12	56	-4
R Superior Medial Gyrus	-	4.75	10	66	22
R Rectal Gyrus	-	4.65	12	24	-6
L Superior Frontal Gyrus	324	6.84	-18	46	44
L Angular Gyrus	214	6.82	-42	-58	28
R PCC	234	6.07	4	-54	32
[(Low-Why + Low-How) > (High-Why + High-How)]					
L Inferior Parietal Lobule	107	6.17	-56	-38	48
R Middle Frontal Gyrus	154	5.89	48	40	16
L IFG (p. Triangularis)	187	5.50	-44	36	12
L Inferior Temporal Gyrus	148	5.17	-58	-56	0
Parametric Model					
Increasing LOA (Q-and-A average)					
L Superior Frontal Gyrus	1319	8.95	-18	48	44
L Superior Medial Gyrus	-	8.50	-8	60	22
L Rectal Gyrus	-	6.44	-2	40	-20
L Middle Temporal Gyrus	192	7.34	-60	-10	-14
No Label	147	6.82	22	24	16
No Label	99	6.65	30	-76	14
L Angular Gyrus	157	5.92	-42	-60	30
L Inferior Occipital Gyrus	180	5.17	-38	-84	0
L Fusiform Gyrus	-	4.68	-34	-66	-12
L Middle Occipital Gyrus	-	4.62	-22	-94	8
L PCC	130	5.07	-6	-54	28
No Label	184	4.75	26	-88	6
Decreasing LOA (Q-and-A average)					
R Middle Frontal Gyrus	1122	6.93	46	38	18
R Middle Frontal Gyrus	-	5.51	36	14	56
R IFG (p. Triangularis)	-	4.44	58	22	18
L IFG (p. Triangularis)	394	6.41	-44	36	12
L Inferior Parietal Lobule	181	6.08	-54	-40	46
R SupraMarginal Gyrus	-	5.84	54	-44	38
R Inferior Parietal Lobule	-	5.46	44	-58	54
R Superior Medial Gyrus	261	5.82	4	30	46
L IFG (p. Opercularis)	153	5.79	-40	6	34
L Inferior Temporal Gyrus	183	5.26	-58	-56	0
L Angular Gyrus	-	4.91	-32	-62	40

Note. All analyses were conducted with a cluster-forming threshold of $p < .001$, followed by cluster-correction at a family-wise error rate of .05. Montreal Neurological Institute (MNI) coordinates. Listed are local maxima for all peaks separated by a minimum of 20 mm. The extent is listed only for the maximum in each cluster; the remaining peaks are marked with a dash. Regions were automatically labeled using the SPM Anatomy Toolbox (Version 2.0; Eickhoff et al. 2005).

Table S4
Peaks from clusters observed in whole-brain analyses for the two models examining the prepotent (Question) LOA at each trial and the within-trial shfit in LOA (from Question to Answer).

Analysis Name			MNI		
Region Label	Extent	*t*	x	y	z
Increasing Prepotent LOA					
L Superior Medial Gyrus	914	8.92	-8	60	24
L Superior Frontal Gyrus	914	8.59	-18	48	44
L Middle Temporal Gyrus	207	7.39	-60	-10	-14
L Middle Temporal Gyrus	168	5.82	-46	-60	28
L Rectal Gyrus	221	5.70	-2	40	-20
L PCC	117	5.00	-6	-54	28
Decreasing Prepotent LOA					
L IFG (p. Triangularis)	380	6.66	-44	36	12
R SupraMarginal Gyrus	552	6.43	48	-42	44
R Middle Frontal Gyrus	1069	6.27	36	14	56
R Middle Frontal Gyrus	1069	6.19	46	38	20
R IFG (p. Orbitalis)	1069	5.74	46	36	0
R Superior Medial Gyrus	233	5.83	4	32	46
L IFG (p. Triangularis)	153	5.82	-42	6	32
L Inferior Temporal Gyrus	209	5.54	-58	-56	0
L Inferior Parietal Lobule	544	5.45	-56	-38	48
L Angular Gyrus	544	5.37	-30	-62	40
Positive Signed LOA Shifts					
L Middle Temporal Gyrus	282	5.81	-60	-6	-20
L Superior Medial Gyrus	128	5.36	-6	60	20

Note. All analyses were conducted with a cluster-forming threshold of p < .001, followed by cluster-correction at a family-wise error rate of .05. Montreal Neurological Institute (MNI) coordinates. Listed are local maxima for all peaks separated by a minimum of 20 mm. The extent is listed only for the maximum in each cluster; the remaining peaks are marked with a dash. Regions were automatically labeled using the SPM Anatomy Toolbox (Version 2.0; Eickhoff et al. 2005).

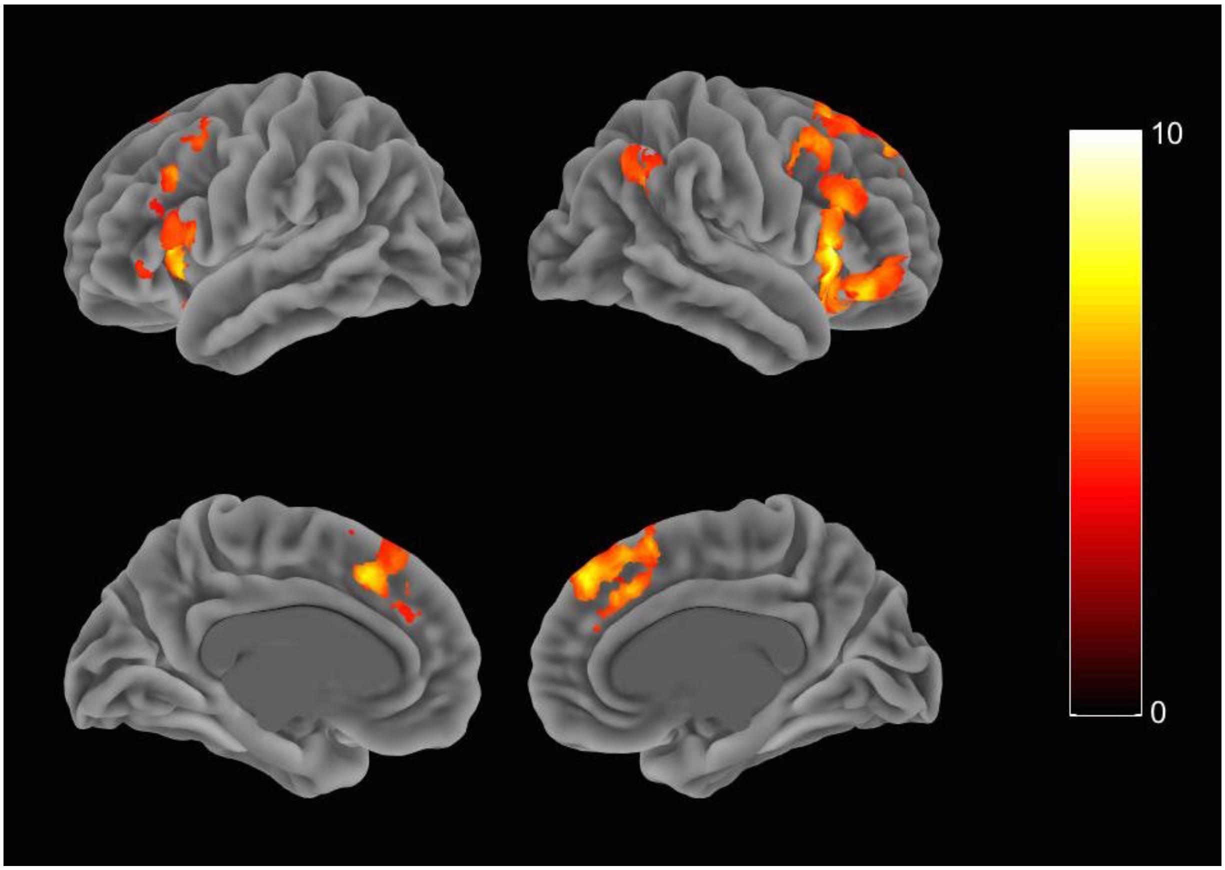
Table S5
Regions associated with increasing response time (RT).

			MNI Coordinates		
Region Label	Extent	*t*	x	y	z
R Superior Medial Gyrus	7245	10.50	4	30	50
R Insula Lobe	-	9.73	28	22	4
R IFG (p. Orbitalis)	-	7.39	46	36	-12
R Middle Frontal Gyrus	-	7.04	44	14	46
L Superior Medial Gyrus	-	6.49	-2	44	26
R Middle Frontal Gyrus	-	4.95	22	54	32
R Middle Frontal Gyrus	-	3.77	36	58	6
L IFG (p. Orbitalis)	1816	7.79	-34	20	-2
L IFG (p. Triangularis)	-	6.65	-46	24	34
R Supramarginal Gyrus	731	6.10	48	-46	40
R Middle Temporal Gyrus	182	5.37	62	-38	-2

Note. All analyses were conducted with a cluster-forming threshold of $p < .001$, followed by cluster-correction at a family-wise error rate of .05. Montreal Neurological Institute (MNI) coordinates. Listed are local maxima for all peaks separated by a minimum of 20 mm. The extent is listed only for the maximum in each cluster; the remaining peaks are marked with a dash. Regions were automatically labeled using the SPM Anatomy Toolbox (Version 2.0; Eickhoff et al. 2005).

Figure S1

Surface rendering of regions associated with longer response time (RT). Regions were identified in a whole-brain search conducted with a cluster-forming threshold of $p < .001$, followed by cluster-correction at a family-wise error rate of .05. Rendering created using the SurfPlot Tool (http://mrtools.mgh.harvard.edu/index.php/SurfPlot). Left hemisphere is left.



Appendix A. Instructions used for normative data collection.

Normative Responses to Why/How Questions

Scale: Yes, No

Instructions for Why-Questions: People often justify their actions with reasons. You will judge the acceptability of reasons for actions. For each judgment, you will see a unique pairing of a reason with an action. For example: "Q: Why ride a bike? A: Exercise." In this example, your task would be to judge whether or not the phrase "exercise" describes a commonly accepted reason for the action described by "ride a bike". If you think it is a commonly accepted reason, then answer 'Yes'. If you think it isn't a commonly accepted reason, then answer 'No'.

Instructions for How-Questions: The actions we perform typically have many parts that collectively describe how to perform the action. For example, consider the action of "brushing one's teeth". Parts of this action commonly include "grab a toothbrush", "apply toothpaste", and "rinse one's mouth". Below you will see a series of questions about how to perform a variety of different actions. Each question is followed by an answer. For example: "Q: How to brush one's teeth? A: Apply toothpaste." In this example, your task would be to judge whether or not the phrase "apply toothpaste" describes a commonly accepted part of the action described by "brush one's teeth". If you think it is a commonly accepted part, then answer 'Yes'. If you think it isn't a commonly accepted part, then answer 'No'.

Abstract/Concrete

Scale: 1 = Completely Abstract, 4 = Equally Concrete and Abstract, 7 = Completely Concrete

Below are a number of phrases that can be used to describe aspects of human action. You will rate how concrete versus abstract each phrase is on a scale from 1 to 7. To make sure you understand what we mean by “concrete” and "abstract", please keep the following dictionary definitions in mind when making your ratings:

- Concrete: existing in a material or physical form; real or solid; not abstract
- Abstract: existing in thought or as an idea but not having a physical or concrete existence

If you think a phrase is completely concrete, rate it as a 7 (completely concrete). If instead you think it is completely abstract, rate it as a 1 (completely abstract). If you think it is as concrete as it is abstract, rate it as a 4 (equally concrete and abstract). The other numbers represent intermediate points on the scale. Please use the entire range of the scale when making your ratings.

Non-imageable/Imageable

Scale: 1 = Very difficult to image, 3 = Fairly difficult to image, 5 = Fairly easy to image, 7 = Very easy to image

Below are a number of phrases that could be used to describe a person's action. For each phrase, use the scale provided to rate how easily it evokes a mental image of what it would be like to perform the action. For example, consider the phrase "brushing one's teeth". This phrase might evoke a visual image of a toothbrush, the taste of toothpaste in your mouth, and/or the feeling of gripping a toothbrush in your hands and moving it back and forth in your mouth. If for a given phrase images like these arise quickly and easily, regardless of what sense they refer to, then rate it a 7 (very easy to image). If instead the phrase evokes no mental image or only does so with great difficulty, then rate it a 1 (very difficult to image).

Even if you've never performed the action described by the phrase, it could still evoke mental images. For instance, many people have never performed the action described by "fire a gun", but the phrase still likely evokes mental images in these people, for instance, an image of what a gun looks like or what a gunshot sounds like.

Broad/Specific

Scale: 1 = Very Broad, 3 = Fairly Broad, 5 = Fairly Specific, 7 = Very Specific

Below are a number of phrases that could be used to describe a person's action. Some of these phrases are very specific. Take for example the phrase "Pick up a pencil". This phrase is very specific in that there are only a few ways in which a person might pick up a pencil. In other words, the phrase "Pick up a pencil" leaves little to the imagination. In contrast, consider a phrase like "Express creativity". This phrase is very broad in that there are a large number of ways in which a person might express themselves. In other words, the phrase "Express creativity" leaves a lot to the imagination.

Use the scale provided to rate how broad vs. specific each phrase is. Please consider the entire range of the scale when making your ratings.

Mind/Body

Scale: 1 = Only Mind, 3 = Mostly Mind, 5 = Mostly Body, 7 = Only Body

Actions involve the actors' mind and the actor's body. Because of this, actions can be described in ways that highlight either the state of mind of the actor or the specific bodily actions that the actor is using to perform the action. Below are phrases that could be be used to describe a person's action. You will rate the extent to which each phrase highlights the role of the actor's mind versus their body in performing the action.

Please consider the entire range of the scale when making your ratings. In other words, do not assume that the phrase must refer to either the mind or the body. Instead, try to be thoughtful about the subtle differences among the descriptions that would suggest more or less involvement of the mind relative to the body and vice versa.

NOTE: If you don't believe that the mind is actually different than the body or if you simply don't fullly know what I mean by 'mind' and 'body', then feel free to assume that by 'mind' I mean the actor's brain and by 'body' I mean everything other than the actor's brain.

Positive/Negative (Emotional/Neutral)

Scale: 1 = Positive, 4 = Neutral, 7 = Negative

Below are a number of brief English phrases. Read each phrase and consider the extent to which it evokes feelings that are positive vs. negative. Please consider the entire range of the scale when making your ratings.

Supplementary References

Eickhoff, S, Stephan, K, Mohlberg, H, Grefkes, C, Fink, G, Amunts, K, Zilles, K. 2005. A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25:1325–1335.

Spunt, R, Adolphs, R. 2014. Validating the why/how contrast for functional mri studies of theory of mind. Neuroimage 99:301–311.