BRIEF REPORT

Seeing What's Possible: Disconnected Visual Parts Are Confused for Their Potential Wholes

Chenxiao Guan and Chaz Firestone
Johns Hopkins University

Perception research traditionally investigates how *actual* states of the world are seen—how we perceive the shapes, colors, and locations that objects actually have. By contrast, everyday life provokes us to consider *possible* states of the world that have not yet (and may not ever) actually obtain. When assembling furniture or completing a jigsaw puzzle, for example, we may appreciate not only the particular shapes of individual objects but also their potential to combine into new objects with distinct shapes of their own. What is the nature of this experience? Here, we explore how visual processing extracts not only what objects *are* but also what they *could become*. In 7 experiments inspired by the puzzle game Tetris, subjects responded to a particular target within a stream of distracting "tetrominoes"; surprisingly, subjects false-alarmed more often to pairs of tetrominoes that could create their target than to pairs of tetrominoes that couldn't—essentially confusing *possible* objects for real ones. This pattern held for several types of objects and transformations, could not be explained by various forms of response bias, and persisted even when shape information was completely incidental to the task. We suggest that possible states of the world are not only contemplated in deliberate reflection but also automatically represented by more basic mechanisms of perception and attention.

Keywords: perception, shape, recognition, possibility, spatial cognition

Things could be different than they are now. The capacity to grasp this fact is a remarkable achievement of the human mind, which is capable of representing not only *actual* states of the world but also *possible* states—states that have not yet, and may not ever, actually obtain. We can consider whether it will rain tomorrow (or what might have happened if it had rained yesterday), we can weigh the potential consequences of our decisions, and we can even wonder what the world might be like with different political leaders, social structures, or physical laws. This psychological representation of possibility is a fundamental aspect of human cognition (Phillips & Knobe, 2018), and recent work has begun to elucidate its role in exceptionally diverse cognitive processes, including moral judgment (Young & Phillips, 2011), physical and probabilistic reasoning (Shtulman & Carey, 2007), theory of mind (Chernyak, Kushnir, Sullivan, & Wang, 2013), decision-making (Burns et al., 2019), linguistic expression (Knobe & Szabó, 2013), and the neural bases (De Brigard, Addis, Ford, Schacter, & Giovanello, 2013) and developmental trajectories (Buchsbaum, Bridgers, Skolnick Weisberg, & Gopnik, 2012; Shtulman, 2009) of such representations.

At the same time that the representation of possibility has been explored in these broad and wide-ranging contexts, such cases have also been constrained in at least one fundamental way: They involve only higher level *cognition* about possibilities—the capacity to reflect, consider, and reason about ways the world might be. By contrast, everyday life may provoke us not only to *think about* possible states in moments of deliberate reflection but also to *recognize* such possibilities in a more immediate, reflexive, and even visual way. For example, when assembling furniture or completing a jigsaw puzzle (see Figure 1), we may appreciate not only the particular properties that individual objects actually have in the moment but also their *potential* to combine into new objects, which may have entirely distinct properties of their own. Such cases raise an intriguing possibility: Do representations of possible states arise not only in deliberate higher level cognition but also in the course of automatic visual processing?
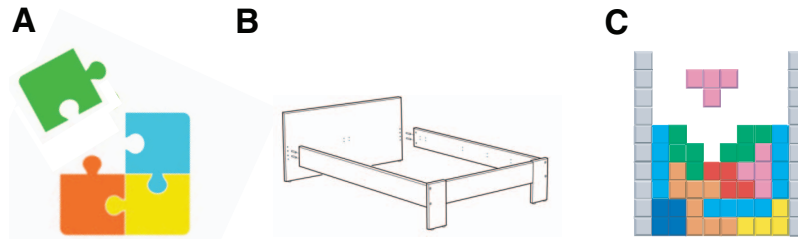
The Perception of Possibility?

On one hand, appreciating possible ways our world might *look* often feels deliberate and effortful, just as when we contemplate how our social and political environment might be different from what it is now. Indeed, the very fact that puzzle games require

Figure 1. Scenarios and stimuli that evoke impressions of possibility. When completing puzzles or assembling furniture, we may appreciate not only the visual features that individual objects actually possess in a given moment (e.g., the particular shape of a given puzzle piece) but also their potential to create new objects with their own distinct features (e.g., the shape of the composite object that multiple puzzle pieces could create). See the online article for the color version of this figure.

sustained concentration (and correlate with other measures of intelligence in adults and children; Hawes, Moss, Caswell, & Poliszczuk, 2015; Jirout & Newcombe, 2015) seems to implicate just this sort of explicit higher level reasoning (or perhaps effortful attentional routines; Michal, Uttal, Shah, & Franconeri, 2016). Even basic processes of mental rotation, which presumably play a role in such representations, typically require sustained effort over a period of several seconds when they are initiated consciously and deliberately (Shepard & Metzler, 1971).

On the other hand, the scenarios described above can also involve flashes of insight whereby one seems to *see,* all-at-once, the possible products of discrete objects. Indeed, players of puzzle games such as Tetris often report "seeing" the structures that would result from various combinations (see Figure 2A) even before the pieces physically fall into place. Moreover, this may even occur involuntarily and effortlessly, including when one is not actively seeking such a result. For example, even without any action intention or goal on the part of the observer (cf. the representation of affordances; Bekkering & Neggers, 2002), the images in Figures 1 and 2A seem to shout their possibilities at observers, who may find themselves compelled to represent the composite objects that such parts can make even when such possibilities have little or no relevance to the observer's goals or purposes.

The Current Experiments: Automatic Representation of Visual Possibility

Here, we aimed to capture this experience in an especially direct way. We asked whether the experience of possible objects is so powerful that subjects would positively indicate their physical presence even when they do not yet exist, and in a way that would intrude on performance in an otherwise straightforward task. We instructed subjects to respond, under time pressure, to a particular target appearing within a stream of distractors. Sometimes, the distractors were object pairs that could create the subject's target in combination; otherwise, the distractors could not create the subject's target but shared other low-level properties. We asked whether subjects might confuse possible objects for real ones and reflexively indicate the target's presence even when it was completely absent from their environment—suggesting that the mind represented the distracting objects in terms of the possible objects they could create.

Experiment 1: Confusing Possible Objects for Actual Objects

Does the mind represent disconnected parts in terms of the possible objects they could create? Experiment 1 investigated this question by asking whether, under pressure to maintain vigilance and respond quickly, subjects might confuse possible objects for real ones.

Method

All data and materials for the experiments reported here (including preregistrations of Experiments 2, 3A, 3B, 4, 5, and 6) are available at https://osf.io/cq89g/.

Participants. For this study, 300 subjects were recruited through Amazon Mechanical Turk (for discussion of this pool's reliability, see Crump, McDonnell, & Gureckis, 2013). This sample size was chosen simply because it seemed very large; in all studies that follow, we used (and preregistered) this same sample size.

Stimuli. Stimuli consisted of custom-designed "Tetris" pieces appearing either in pairs or alone (see Figure 2A). The pieces were composed of gray squares whose shading implied a raised, textured surface. The target image was a 5 × 5 block of these squares (i.e., a square composed of smaller squares). To create the distractor images, we began with six "families"; for each family, a contiguous four-square piece was removed from the larger block and placed one square's height above the block, as if floating above it. A differently shaped piece (corresponding to the tetrominoes of Tetris) was removed for each of the six families, and these six images served as the stimuli for what we refer to as the *potential* condition, because the removed piece could "create" the complete target image in combination with the lower block. As shown in Figure 2A, the image presented to subjects included both of these spatially disconnected pieces shown at one time (cf. Zhou, Zhang, Ding, Shui, & Shen, 2016, who explore attention to objects that are integrated over sequential presentations).

For the *no potential* condition, we simply shuffled the floating pieces between families, so that the floating piece and the block could not combine to create the target image but remained matched on a wide array of low-level properties, including luminance, contrast, texture, color, average height, width, and spatial frequency; the only difference between the *potential* and *no potential* groups was thus the *relationship* between the floating tetromino
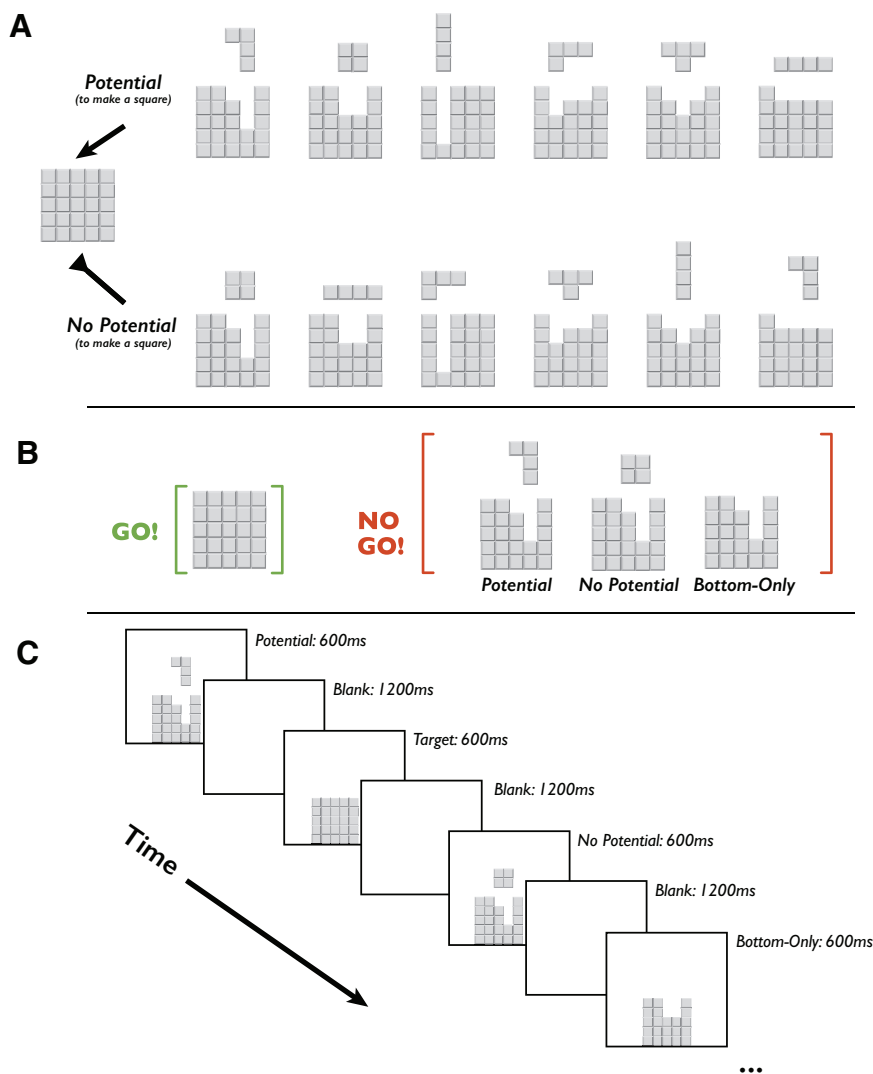
Figure 2. Stimuli and task for the present studies. Panel A: Some pairs of objects have the potential to easily combine into other objects, whereas some pairs of objects do not. The object pairs shown here vary in their ability to combine into a square and served as the stimuli in many of the present experiments. Panel B: In Experiment 1, subjects were given a simple Go/No-Go task in which they were told to press a key whenever they saw a complete square but never at any other time. Panel C: The experimental session consisted simply of a stream of images separated by 1,200-ms blanks. See the online article for the color version of this figure.

and the block beneath it, which either could or could not create the target image by translation.

Finally, we also included the negative parts remaining after the tetromino-shaped pieces were removed, without the floating pieces on top; we refer to those images as *bottom only* images. These 19 images (six potential, six no potential, six bottom only, and one target) served as the stimulus set.

Due to the nature of online experiments, we cannot specify here the exact size, color, or brightness (etc.) of the images, because we could not know each subject's particular viewing conditions. However, any distortions introduced by a given subject's viewing distance or monitor settings would have been equated across all stimuli and conditions.

Procedure. Subjects completed a simple Go/No-Go task in which they were instructed to press a key (spacebar) whenever the target image appeared onscreen and never for any other type of image (see Figure 2B). To ensure that the instructions were clear, we included a preview of each trial type, with explicit instructions for what to do when that image appeared (e.g., *Press the space-bar!; DO NOT press the spacebar!*).

During the experiment, images appeared in a continuous stream (see Figure 2C), with each image remaining on-screen for 600 ms and then disappearing for 1,200 ms, after which the next image appeared for another 600 ms (this delay was long enough to prevent the experience of apparent motion); responses were accepted for this entire 1,800-ms duration. These parameters were

chosen so that the task would be easy but require constant vigilance; even though we expected performance to be high, lapses in vigilance might reveal the operation of implicit or impulsive processes that would otherwise be successfully inhibited.

There were 84 trials total: 24 *target* trials, 24 *potential* trials (four repetitions of each *potential* image), 24 *no potential* trials (four repetitions of each *no potential* image), and 12 *bottom only* trials (two repetitions of each *bottom only* image). The trial order was randomized for each subject, with the constraint that half of the *target* trials were directly preceded by a *potential* trial, and half of the *target* trials were directly preceded by a *no potential* trial.

Subjects received feedback throughout the experiment: Each correct target identification caused the border of the frame in which the images appeared to briefly flash green (for 200 ms); incorrectly responding that a target was present caused the border to flash red.

Importantly, "possibility" and related notions were never mentioned and indeed the potential for distractors to combine into the target was completely irrelevant and incidental to the actual task subjects were given—which was simply to press the spacebar for a complete square and never for anything else.

Finally, we excluded subjects based on two criteria. First, any subject who failed to contribute a complete data set was excluded from further analysis. Second, any subject who failed to perform significantly above chance (i.e., above 75%) on any of the key conditions (potential, no potential, and target) was also excluded; this criterion prevented inclusion of subjects who misunderstood the task or were otherwise distracted, and it was applied identically to all of the experiments reported here. We also preregistered these criteria, as well as an analysis plan, in every subsequent experiment.

Readers can experience the task for themselves at http://perceptionresearch.org/tetris.

Results and Discussion

As expected, task performance was high: Even without exclusions, accuracy (i.e., the proportion of trials with correct detections or correct rejections) was 93.4%. However, after excluding 17 subjects (6%) who had incomplete data and 38 subjects (13%) who failed to perform above 75% accuracy on all the key conditions, accuracy was now 98.1%. (No analysis here depended in any way on these exclusions; i.e., all effects reported here remained statistically significant in the same direction\even without excluding these subjects).

However, subjects did occasionally false-alarm to nontargets. Remarkably, this happened more frequently in the *potential* condition (2.52%) than in the *no potential* condition (1.09%), $t(244) = 5.33$, $p < .001$, $d = .37$ (see Figure 3A).[1] In other words, subjects were more likely to mistakenly respond that the image in front of them was a complete square when the image comprised two pieces that could create a complete square versus two pieces that could not. This initial result suggests that, at least for a moment, subjects represented the disconnected parts in terms of the complete object they could create, such that they mistakenly responded that a merely potential object was physically present on the display.

To further support this interpretation, we also examined performance on correct detection trials—that is, trials in which subjects correctly answered that the target image was present. As noted earlier, half the *target* trials were preceded by a *potential* trial, and half were preceded by a *no potential* trial. Intriguingly, subjects were faster to identify the presence of a target when the previous trial was an image that could create the target versus when it could not: 456 ms versus 465 ms, $t(244) = 3.72$, $p < .001$. This effect also remained when excluding response times longer than 600 ms (i.e., after the target disappeared from the display): 435 ms vs. 441 ms, $t(242) = 3.70$, $p < .001$, $d = .24$. This result is further consistent with our interpretation that subjects represented the discrete objects in terms of the new composite object they could create, as here that representation appeared to prime the complete object, boosting recognition of it on subsequent trials.

Finally, we observed the false-alarm effect not only over the entire session but also when considering just the final half of trials, with a higher false-alarm rate for potential (1.1%) versus no potential (0.4%), $t(244) = 2.85$, $p = .005$. By this point, subjects had already received extensive feedback about their responses, so there can be no question that they understood which trials required responses and which did not; still, they were apparently unable to inhibit their representations of potential and so were more likely to mistakenly indicate the target's presence for trials that could create it than for trials that could not. Indeed, this is what we mean when we describe these results as reflecting "automatic" processing: a process that the mind engages in even when it is irrelevant to one's active goals and intentions—and indeed even when it would behoove subjects to simply ignore this property altogether.

Experiments 2–4: Generalization

How general and reliable is this representation of possibility? And could our results be explained by other factors? Experiments 2–4 extended these results to rotations (Experiment 2), translations (Experiments 3A and 3B), and other shape classes (Experiment 4), establishing this phenomenon's reliability, generalizing it to new contexts, and ruling out several alternative explanations.

Method

Experiments 2–4 were identical to Experiment 1 except as noted here. Four new groups of 300 subjects participated in the four experiments (1,200 subjects total). We also preregistered the hypotheses, sample sizes, exclusion criteria, and analysis plans for all four experiments.

In Experiment 2, subjects saw the same images as in Experiment 1, except that the entire stimulus was rotated 90° counterclockwise. One possible interpretation of Experiment 1 is that subjects represented the floating tetromino as a "falling" object (due to

---

[1] If these false-alarm rates seem small or rare, note that they are perhaps artificially "deflated" by our exclusion criteria, which required that subjects performed well on all trial types (and so leave out precisely those subjects who committed many false alarms). If those subjects are not excluded, the false-alarms rates become 7.80% for the *potential* condition and 3.34% for the *no potential* condition—and remain statistically significant, $t(291) = 4.91$, $p < .001$. However, consistent with our preregistered analysis plans, we continue to report the postexclusion data throughout this article.
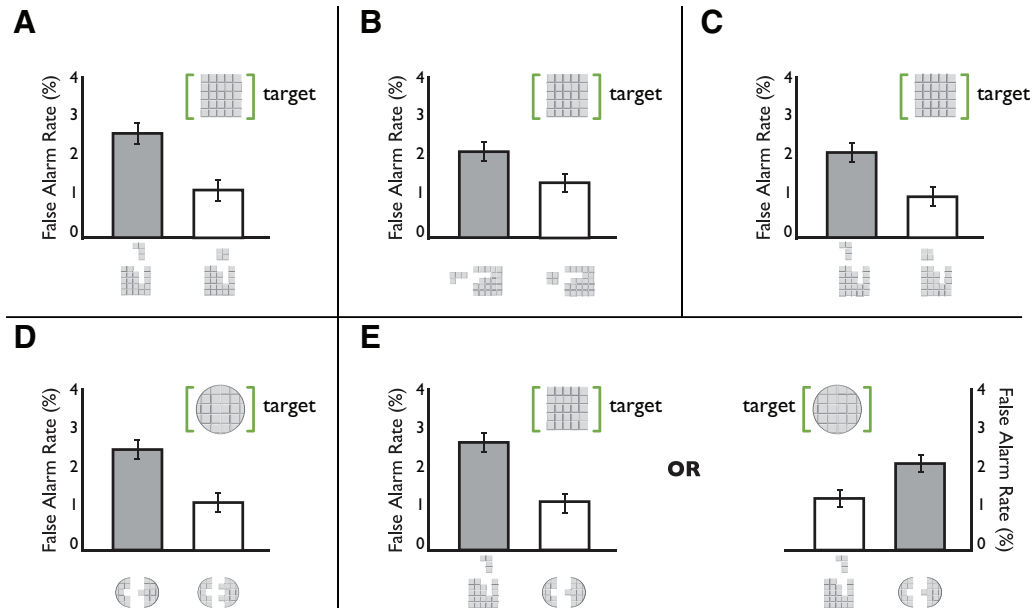
Figure 3. Results from Experiments 1–5, showing automatic representation of objects in terms of their potential (only Experiment 3A is not shown here). Panel A: In Experiment 1, the subjects' target was a complete square; subjects false-alarmed more often to pairs of objects that could create their target in combination (potential) than to pairs of objects that could not (no potential). Panel B: Experiment 2 replicated Experiment 1 but with all stimuli rotated counterclockwise by 90°; the same pattern of results emerged. Panel C: Experiments 3A and 3B included shifted stimuli in which the floating tetromino was misaligned with the block beneath it; again, the same pattern of results emerged (this graph shows data from Experiment 3B). Panel D: Experiment 4 generalized previous effects to circles. Panel E: In Experiment 5, subjects were randomly assigned to have either a square target or a circle target; otherwise, they saw the same distractors. Subjects continued to false-alarm to the distractors that could create their target than to those that couldn't, even when this property was not determined by the distractors themselves but instead by the relationship between the distractors and the target. See the online article for the color version of this figure.

gravity), which is known to bias representations of location (Hubbard & Bharucha, 1988) via representational momentum (Freyd, 1983); this could have produced a representation of a complete square, but only because of the accidental alignment between the tetromino and the larger block. However, here a "gravitational" influence on the component pieces would not produce a complete square (see Figure 3B), allowing us to rule out this possibility.

In Experiment 3A, the floating tetromino was horizontally displaced by one square's width. Here, the floating tetromino would have to move both horizontally and vertically to combine with the larger block. This also altered the negative space between the floating tetromino and the block below, ensuring that this factor did not explain previous results. In Experiment 3B, the floating tetromino was horizontally displaced by two squares' width, to even further disrupt the gestalt alignment of negative space between the pieces (see Figure 3C).[2]

In Experiment 4, we expanded our stimuli beyond rectilinear objects to circular "puzzle piece" images. We extracted a circle from the same 5 × 5 grid of squares from Experiment 1 and divided the circle into halves that either could or could not combine to create the whole (see Figure 3D). These objects differed from those used earlier not only with respect to their shape but also perhaps with respect to their familiarity, because these sorts of objects do not appear in any popular puzzle games that we know of (unlike the Tetris-inspired pieces used earlier). Here, there were

96 trials: 24 *potential* trials, 24 *no potential* trials, 24 *target* trials, and 24 *piece* trials (in which only the left or right half of the puzzle pieces from the other trials was shown). Across all four experiments, we predicted that subjects would false-alarm more in the *potential* conditions than in the *no potential* conditions, even when the combined object that the two pieces could form required a more convoluted path and even for other classes of shapes.

Results

All four experiments replicated the pattern of results from Experiment 1.

In Experiment 2, with shapes turned on their sides (so that there was no "gravitational" influence), there were more false alarms in the *potential* condition than in the *no potential* condition (2.08% vs. 1.22%), $t(243) = 3.48$, $p < .001$, $d = .24$ (see Figure 3B).

In Experiment 3A, with the floating tetromino shifted, the same pattern emerged: 2.38% versus 1.07%, $t(242) = 5.86$, $p < .001$, $d = .41$. This also occurred in Experiment 3B, where the floating tetromino was shifted even further: 2.09% versus .90%, $t(218) = 4.38$, $p < .001$, $d = .34$ (see Figure 3C).

---

[2] We thank an anonymous reviewer for proposing this variation of Experiment 3A.

In Experiment 4, with circle-based puzzle pieces instead of square tetrominoes, the same pattern emerged again: 2.43% versus 1.04%, $t(244) = 5.51$, $p < .001$, $d = .38$ (see Figure 3D).

These results establish the robustness and reliability of the visual system's representation of possibility: The present results could not be explained by a simple gravitational influence and generalized beyond one particular spatial arrangement or shape type.

Experiment 5: Isolating Possibility

We've described these results in terms of potential objects that discrete entities could create. But perhaps our results are explained merely by "geometric compatibility" in the *potential* conditions. In other words, perhaps simply *noticing that two objects can fit together* is exciting on its own—causing false alarms—without invoking the object they would subsequently *create* as a result. Experiment 5 tested this possibility by holding geometric compatibility constant, thereby isolating potential per se.

Method

In this experiment, 300 new subjects were randomly assigned to a square target (as in Experiment 1) or a circular target (as in Experiment 4). All subjects saw the same distractors: four *potential* images from Experiment 1, four *potential* images from Experiment 4, and a sample of *bottom only* trials from Experiment 1 and *piece* trials from Experiment 4. There were 120 trials total.

The key feature of this experiment's design was that even though all subjects saw the same distractors, some were potential squares and some were potential circles. For subjects with a *square* target, the potential-square distractors were the *potential* trials and the potential-circle distractors were the *no potential* trials (see Figure 3E). For subjects with a *circle* target, this was reversed. This equated all conditions for the degree of geometric compatibility of the shapes in the *potential* and *no potential* trials to specifically isolate whether the images could create the subject's target, over and above a bias to respond simply when two objects could neatly fit together.

Results

Again, there were more false alarms in the *potential* condition than the *no potential* condition (2.34% vs. 1.08%), $t(234) = 5.35$, $p < .001$, $d = .39$ (see Figure 2E), even when fit was matched across conditions. This suggests that, beyond any effect of noticing that two objects can fit together, discrete objects are truly represented in terms of what they can *create*.

Experiment 6: Seeing What's Possible When What's Possible Is Irrelevant

The previous results suggest that representations of possibility arise naturally and spontaneously, in that the mind computes possible objects even when the task does not require such computations. However, across all of these experiments, subjects were required to attend to the shape of the objects (because their target was defined by its shape), and so the task may still have seemed to invite a broad consideration of shape properties. These results alone still suggest a strong degree of automaticity, because these various tasks never required the computation of possible shapes,

and yet such possibilities still drove responses. Nevertheless, we can ask just *how* automatic such representations can be: Are possible shapes computed not only in tasks that do not require attention to possible shape but also in tasks that do not even require attention to shape at all?

Experiment 6 tested this question by asking whether possible and actual objects facilitate judgments about each other even when those judgments are about the object's color rather than its shape. In other words, we asked whether possible objects are represented not only when possibility is irrelevant to the task but also when shape itself is irrelevant, because no judgment the subjects make is ever about shape. If possibility plays a role even under these circumstances, this would imply an even stronger degree of automaticity for the computation of possible objects.

Method

This experiment also involved 300 new subjects. All subjects saw the same set of stimuli, which here included both the square-based and circle-based objects from Experiment 5. As in previous experiments, there were *potential* and *no potential* conditions, here determined by the relationship between two adjacent trials (for reference, we refer to these trials as Trial T and Trial T − 1). In the *potential* condition, the two images shown on Trials T and T – 1 were possible and actual versions of each other: for example, a square and a pair of objects that could create a square. In the *no potential* condition, the two images on Trials T and T – 1 were not versions of each other at all: for example, a square and a pair of objects that could create a circle (the same permutations also appeared for circles and possible circles, and circles and possible squares).

However, unlike in previous experiments, subjects were never asked to make any judgments about the shapes of the objects on the screen. Instead, subjects made judgments about the color of the stimuli: All stimuli could be rendered either in blue or in yellow, and the subject's task was simply to indicate the color of the object on the screen on every trial (by pressing either *B* or *Y;* see Figure 4A). There were 128 trials total, including every possible two-way permutation of the shapes, colors, instances, and so forth. We predicted that color judgments for a given object (on Trial T) would be facilitated by first seeing some possible or actual version of that object previously (on Trial T – 1) and that this would manifest as faster response times for possible-actual matches than for possible-actual mismatches.[3] We also preregistered the hypotheses, sample sizes, exclusion criteria, and analysis plans for this experiment.

Results

Response times for the *potential* condition were faster than for the *no potential* condition (469 ms vs. 476 ms), $t(216) = 5.62$, $p < .001$, $d = .38$ (see Figure 4B). In other words, possible and actual objects were represented as being similar enough that they facilitated judgments of completely incidental properties (here, color).

---

[3] We thank an anonymous reviewer for a detailed and helpful proposal of this experimental design.
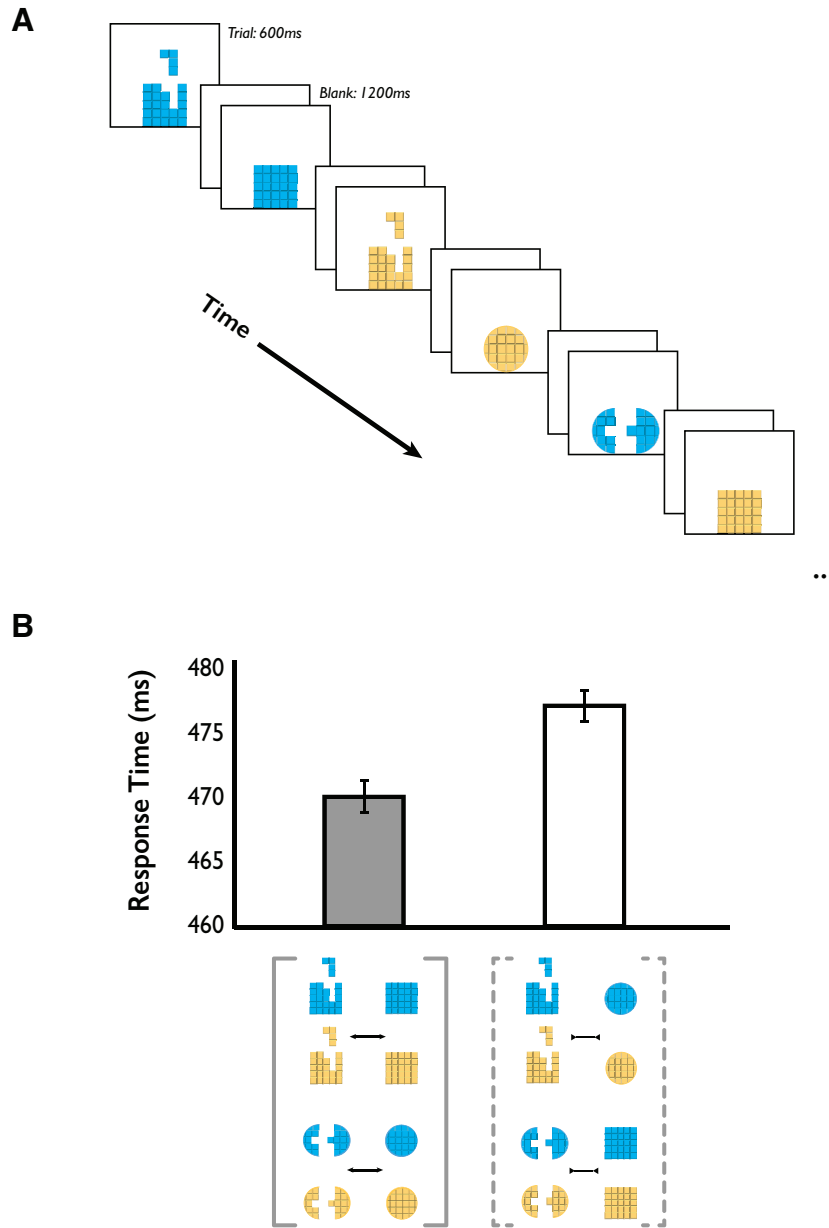
Figure 4. Design and results of Experiment 6. Panel A: A stream of images that included both square- and circle-based objects was presented. Sometimes, adjacent trials involved possible and actual versions of the same object (*potential*); at other times, adjacent trials did not involve possible and actual versions of the same object (*no potential*). Crucially, all stimuli could be rendered in blue or yellow, and the subject's task was simply to indicate the color of the object with a key press, with no shape judgment required at all. Panel B: Subjects were faster to indicate the color of an object for *potential* trials (gray bar in Panel B) compared to *no potential* trials (white bar). Error bars indicate the standard error of the difference between condition means. Note that, even though the x-axis shows color matches, our analysis collapsed over color matches and color mismatches (such that, e.g., a yellow square preceded by a blue possible square was still considered a potential trial). See the online article for the color version of this figure.

General Discussion

Perception research has traditionally concerned how *actual* states of the world are seen: how we perceive the shapes, colors, and locations that objects actually have (or had in the past; see Chen & Scholl, 2016) and how such objects relate to one another (Hafri, Papafragou, & Trueswell, 2013; Linkenauger et al., 2015;

Strickland & Scholl, 2015; Zhou et al., 2016). By contrast, here we have suggested that visual processing can also represent merely *possible* states. When searching for a particular target among distractors, subjects mistook the distractors for their target when the distractors could create it—suggesting that the mind processed such objects in terms of their *potential*.

Seeing What Isn't (Yet) There

The notion that an object might be represented as being physically present even when the visual system has direct access to only the parts that could *become* that object goes beyond extant accounts of object representation and recognition. For example, many theories of object recognition include a role for those objects' parts (e.g., Biederman, 1987; Marr & Nishihara, 1978); however, even these theories assume that such parts are already *connected* or otherwise in their appropriate locations (cf. Bertamini, Friedenberg, & Kubovy, 1997). By contrast, here we suggest that disconnected parts of objects can activate representations of their complete wholes even when those parts are not currently arranged in the form of those wholes and instead could only possibly become them at some later time.

Moreover, the possibilities explored here differ from mere extrapolation based on motion (De Freitas, Myers, & Nobre, 2016) or semantic meaning (Hsu, Taylor, & Pratt, 2015) and perhaps even representational momentum (Freyd & Finke, 1984). Representational momentum is the process by which objects and events are (mis)represented as physically displaced in some implied direction. For example, a rotating rectangle that suddenly disappears will be misremembered as having been farther along its rotational trajectory than it really was, because subjects extrapolate its future position. Could the possibilities explored in the present work be explained by this sort of mechanism?

First, we note that all stimuli in our studies were completely static throughout the experiment; thus, any extrapolation by subjects must not have been based on any actual motion of the objects (because no such motion existed in the first place). However, representational momentum can also arise for implied motion, rather than actual motion, as when it occurs for static stimuli that would be subject to the force of gravity (Freyd, 1983). Because we ruled out a purely gravitational bias in Experiments 2 and 4, our results might thus reveal that possibilities themselves can serve as implied directions for objects—as if object parts are somehow magnetically attracted to their complete states—which would be a substantive and previously unknown driver of this core cognitive process.[4]

But second, no study of representational momentum has, to our knowledge, shown that this process can create new object representations in the way we explore here. In other words, it's one thing for an object to be represented as physically displaced from its true location (e.g., a rotating square being represented as more rotated than it really was, or a tetromino being represented as located closer to its bottom counterpart than it really is), but it's quite another for this process to then result in a brand new shape representation involving its combination with another object (e.g., the complete square that would result from the tetromino's displacement). In that case, our work would go beyond anything previously known about representational momentum, suggesting that it can operate not only over object locations but also over object identities themselves.

The possibilities we explore here also seem distinct from those associated with "affordances" (e.g., graspability; Bekkering & Neggers, 2002). In our experiments, the stimuli were simply static shapes without any particular high-level associations beyond the subject's goal to search for a particular target (in Experiments 1–5), and subjects had no intention or ability to act upon them.

And in Experiment 6, nothing about their completed shape was even relevant to the task. Indeed, whereas affordances capture relations between the observer and the world, the cases explored here relate the world to itself (regardless of who is there to interact with it), by capturing how objects in the world might interact and combine to look different in the future.

Looking Forward

How general is this phenomenon? Our work here explored different orientations and configurations of familiar Tetris pieces and also extended the phenomenon to other stimuli beyond this familiar environment (e.g., the circles in Experiments 4–6). Still, future work could ask whether these results extend even further, perhaps to irregular shapes that may be more complex than squares and circles, to shapes with different topological properties (e.g., internal holes), to pieces that are even further separated than in the present experiments, and so forth.

Conclusion

More generally, the results here suggest representations of possibility play a role not only in higher level cognition (Phillips & Knobe, 2018) but also in perception itself: We can see not only how the world is but also how it could be.

---

[4] Moreover, if we compare those experiments that involved vertically oriented stimuli (which perhaps imply some kind of gravity-based motion) against those experiments that involved horizontally oriented stimuli (which do not), there is no significant difference in the magnitude of the false-alarm biases, $t(976) = .72$, $p = .47$, just as one might predict if gravity were not the explanation of these effects. We note, however, that this analysis—which supports our interpretation—nevertheless compares experiments that were run on different subjects and at different times and so forth, and it was also never part of our preregistered analysis plan, so we do not wish to heavily rely on it. We thank a reviewer for strongly recommending this analysis.

References

Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. *Psychological Science, 13,* 370–374. http://dx.doi.org/10.1111/j.0956-7976.2002.00466.x

Bertamini, M., Friedenberg, J. D., & Kubovy, M. (1997). Detection of symmetry and perceptual organization: The way a lock-and-key process works. *Acta Psychologica, 95,* 119–140. http://dx.doi.org/10.1016/S0001-6918(96)00038-8

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147. http://dx.doi.org/10.1037/0033-295X.94.2.115

Buchsbaum, D., Bridgers, S., Skolnick Weisberg, D., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences, 367,* 2202–2212. http://dx.doi.org/10.1098/rstb.2012.0122

Burns, P., McCormack, T., Jaroslawska, A., Fitzpatrick, Á., McGourty, J., & Caruso, E. M. (2019). The development of asymmetries in past and future thinking. *Journal of Experimental Psychology: General, 148,* 272–288. http://dx.doi.org/10.1037/xge0000464

Chen, Y.-C., & Scholl, B. J. (2016). The perception of history: Seeing causal history in static shapes induces illusory motion perception. *Psy-

chological Science, 27, 923–930. http://dx.doi.org/10.1177/0956 797616628525

Chernyak, N., Kushnir, T., Sullivan, K. M., & Wang, Q. (2013). A comparison of American and Nepalese children's concepts of freedom of choice and social constraint. *Cognitive Science, 37,* 1343–1355. http://dx.doi.org/10.1111/cogs.12046

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8,* e57410. http://dx.doi.org/10.1371/journal.pone .0057410

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia, 51,* 2401–2414. http://dx.doi.org/10.1016/j.neuropsychologia.2013.01.015

De Freitas, J., Myers, N. E., & Nobre, A. C. (2016). Tracking the changing feature of a moving object. *Journal of Vision, 16*(3), 22.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics, 33,* 575–581. http://dx .doi.org/10.3758/BF03202940

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 126–132. http://dx.doi.org/10.1037/0278-7393.10.1.126

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General, 142,* 880–905. http://dx.doi .org/10.1037/a0030045

Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. *Trends in Neuroscience and Education, 4,* 60–68. http://dx.doi.org/10.1016/j.tine.2015.05.001

Hsu, P., Taylor, J. E., & Pratt, J. (2015). Frogs jump forward: Semantic knowledge influences the perception of element motion in the Ternus display. *Perception, 44,* 779–789. http://dx.doi.org/10.1177/0301 006615596903

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics, 44,* 211–221. http://dx.doi.org/10.3758/BF03206290

Jirout, J. J., & Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from a large, representative U.S. sample. *Psychological Science, 26,* 302–310. http://dx.doi.org/10.1177/09567 97614563338

Knobe, J., & Szabó, Z. G. (2013). Modals with a taste of the deontic. *Semantics and Pragmatics, 6,* 1–42. http://dx.doi.org/10.3765/sp.6.1

Linkenauger, S. A., Wong, H. Y., Geuss, M., Stefanucci, J. K., McCulloch, K. C., Bülthoff, H. H., . . . Proffitt, D. R. (2015). The perceptual homunculus: The perception of the relative proportions of the human body. *Journal of Experimental Psychology: General, 144,* 103–113. http://dx.doi.org/10.1037/xge0000028

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London: Series B, Biological Sciences, 200,* 269–294. http://dx.doi.org/10.1098/rspb.1978.0020

Michal, A. L., Uttal, D., Shah, P., & Franconeri, S. L. (2016). Visual routines for extracting magnitude relations. *Psychonomic Bulletin & Review, 23,* 1802–1809. http://dx.doi.org/10.3758/s13423-016-1047-0

Phillips, J., & Knobe, J. (2018). The psychological representation of modality. *Mind & Language, 33,* 65–94. http://dx.doi.org/10.1111/mila .12165

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703. http://dx.doi.org/10.1126/science.171 .3972.701

Shtulman, A. (2009). The development of possibility judgment within and across domains. *Cognitive Development, 24,* 293–309. http://dx.doi.org/ 10.1016/j.cogdev.2008.12.006

Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development, 78,* 1015–1032. http://dx.doi.org/10.1111/j.1467-8624.2007 .01047.x

Strickland, B., & Scholl, B. J. (2015). Visual perception involves event-type representations: The case of containment versus occlusion. *Journal of Experimental Psychology: General, 144,* 570–580. http://dx.doi.org/ 10.1037/a0037750

Young, L., & Phillips, J. (2011). The paradox of moral focus. *Cognition, 119,* 166–178. http://dx.doi.org/10.1016/j.cognition.2011.01.004

Zhou, J., Zhang, H., Ding, X., Shui, R., & Shen, M. (2016). Object formation in visual working memory: Evidence from object-based attention. *Cognition, 154,* 95–101. http://dx.doi.org/10.1016/j.cognition .2016.04.009