Affordances and Intentionality: Reply to Roberts

Michael L. Anderson and Anthony Chemero

Franklin & Marshall College

In this essay we respond to some criticisms of the guidance theory of representation offered by Tom Roberts. We argue that although Roberts' criticisms miss their mark, he raises the important issue of the relationship between affordances and the action-oriented representations proposed by the guidance theory. Affordances play a prominent role in the anti-representationalist accounts offered by theorists of embodied cognition and ecological psychology, and the guidance theory is motivated in part by a desire to respond to the critiques of representationalism offered in such accounts, without giving up entirely on the idea that representations are an important part of the cognitive economy of many animals. Thus, explorations of whether and how such accounts can in fact be related and reconciled potentially offer to shed some light on this ongoing controversy. Although the current essay hardly settles the larger debate, it does suggest that there may be more possibility for agreement than is often supposed.

Keywords: action, intentionality, representation, affordances

We would like to begin by thanking Dr. Roberts for his critical engagement with the guidance theory, and Ray Russ and the editors at *Journal of Mind and Behavior* for arranging this exchange. Although we will argue below that Roberts' criticisms miss their mark, the question of the relationship between affordances and the intentional states treated by the guidance theory is indeed an important one. We appreciate this opportunity to offer our thoughts on the best way to approach the issue.

We'll not spend much time here explaining the guidance theory — just enough to make cogent the responses to follow. The reader is directed to Anderson and Rosenberg (2008) and to the first few pages of Roberts' article (2009, this issue) for more detailed and expansive accounts. Briefly, then, the guidance theory says that a given token T represents an entity E for a subject

S if and only if T is standardly used to provide guidance to S when taking action with respect to E. More briefly: T represents E just in case it is used to guide actions targeting E. Such representations are in error on this account when an action taken with respect to E would fail because of some feature of the action-guiding token. To stick with an example developed in Anderson and Rosenberg (2008) and discussed by Roberts, a particular string in a file on a computer hard drive represents a particular printer just in case that string is standardly used by the computer to guide actions it takes with respect to that printer — for instance, causing the printer to spool a certain document. But suppose that the last four characters of that string are such that under some circumstances the computer commands the printer to staple the document. However, the printer has no stapler. Colloquially, we might say that the string as written represents the printer as having a stapler, but it does not. Note, however, that for the guidance theory this error should not be analyzed in terms of an inaccurate picture or description, but is instead a matter of guiding actions that fail because of this guidance. What it *is* for a representation to be in error is for it to be such that it is used to guide actions that will thereby fail.

Roberts' Critique of the Guidance Theory

That's the guidance theory in a nutshell; Roberts objects to the theory on three grounds. He argues first that states can mis-represent even when action succeeds, second that this account of intentionality does not allow one to specify the target Es with sufficient precision and third that the guidance theory can attribute implausible mental content to animals in certain cases. We'll take each of these concerns in turn.

Error and Successful Action

To make the case that representational error cannot just be a matter of action failure, Roberts cites the famous Titchener circles illusion in the context of the two visual systems hypothesis (Milner and Goodale, 1995). In the Titchener circles illusion (reproduced below for those unfamiliar with it) the center circle in the left grouping appears larger than the center circle in the right grouping, even though they are the same size.

In a clever set of experiments using poker chips as circles, Aglioti, DeSouza, and Goodale (1995) asked participants to pick up one of the center chips, saying: "if you think the two disks are the same size, pick up the one on the left; if you think they are different in size, pick up the one on the right" (p. 681). The choice patterns showed that participants were subject to the illusion, as expected. However, measurements of hand pre-formation and grip aperture suggested that participants' motor movements were guided by the actual, and not
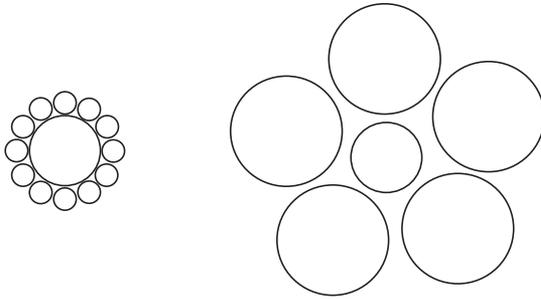
Figure 1: The Titchener circles illusion.

the apparent size of the targets. The authors interpreted this as further evidence that conscious perceptual judgments and online motor guidance are in fact handled by different specialized visual processing streams.

This, of course, is just the point that rescues the guidance theory from Roberts' first criticism. Roberts claims that the participants' success in picking up the chip even when their representation is demonstrably in error shows that error and action failure can be dissociated. But, in fact in this experiment participants took *two* actions, guided by two different mental representations.[1] First, they chose which disk to target, and then they grasped the targeted disk. The first representation — the one that guided the choice — was indeed in error, and resulted in an action that failed in its intent. The second representation — the one that guided the grasping — was not in error, and the action succeeded. Admittedly, this point would be clearer had the instructions been: "choose the right disk if they *are* different in size" or, better yet "choose the disk that is larger," but we trust the point goes through despite the potential confusion caused by the word "think" in the instructions.[2] Participants asked to choose the larger disk would fail, for they would not end up with the larger disk (see Anderson and Rosenberg, 2008, p. 78), even though they would successfully end up holding the disk they chose.

In fact there are a number of illusions of this sort, for instance the size-weight illusion in which observers judge the larger of two objects of equal mass to be lighter (Amazeen and Turvey, 1996; Shockley, Carello, and Turvey, 2004)

---

[1]According to guidance theory, that is. It is worth noting that the present authors do not agree on whether this task involves mental representation. In fact, Chemero doubts that representations are required to explain any of the illusions described in this section. He also doubts that frogs represent flies, whether as flies or as black specks.

[2]Indeed, insofar as the instructions as written are asking participants to choose based on characteristics of the *appearance*, participants make the *correct* judgment. One *does* appear larger than the other. So, there is no misrepresentation here.

and various grade and distance illusions, in which observers judge hills steeper and distances further when wearing a heavy backpack than they do when unencumbered (Proffitt, 2006; Proffitt, Stefanucci, Banton, and Epstein, 2003). In these cases, subjects appear to be in error, but in ways that either do not affect, or may even enhance the success of motor actions taken with respect to the targets — for instance by correctly priming the motor system for the amount of effort required to manipulate or navigate the entities in question. The guidance theory approaches all these cases in the same two steps. First, by looking for evidence that motor actions and explicit perceptual judgments might be guided by different representations, as is not infrequently the case. Second — when the evidence seems to point to a unified representation guiding both sorts of actions — by distinguishing between actions that will succeed and those that will fail as a result of the use of the representation in guidance. To support the guidance theory, one must only identify, for any representation that is in error, the set of actions that would fail as the result of its use. It is no counterexample to this account to find a set of actions that will succeed given use of the representation in question — even representations that are seriously in error may successfully guide *some* actions in some circumstances. What is wanted as a counterexample to the guidance theory treatment of error is a representation that is in error even though *all* of the actions it is used to support would succeed. That would be a very interesting case for discussion, but Roberts' example is not of this sort.

Content Indeterminacy

The second objection Roberts raises against the guidance theory is of content indeterminacy. Insofar as there are many ways to describe a targeted entity, the content of the intentional state might be ambiguous; one could be representing some entity E *qua* fly, or *qua* black speck. Conversely, as he points out, the same action might be taken with respect to different targets. Thus, there appears to be the potential for bi-directional ambiguity in the guidance theory.

For the answer to this objection, we need to return to the definition of the intentional connection as given in Anderson and Rosenberg (2008), and repeated above: T represents E just in case it is standardly used to guide actions with respect to E. From the standpoint of the guidance theory, it doesn't matter if there could be multiple ways to *describe* E; the guidance theory is not providing a descriptive theory of content, but an account of the fundamental basis for the intentional connection. Whether or not content seems potentially ambiguous from the third-person perspective — since any such attribution is constrained only by standards of rational interpretation, which leave room for competing equally justified theories about another's mental state — from the first-person perspective of the guidance theory content is entirely determinate: it is the

entity with respect to which a given action (or set of actions) is taken; it is the focus of the animal's effort to change or control the world. Whether it is better to describe the state of the animal as a fly-experiencing state, or a black-dot-experiencing state when it is representing E with T will depend on the set of actions supported by T in a range of circumstances. That the right answer to *this* question could remain scientifically underdetermined is not sufficient reason to suppose that there is any actual ambiguity about what E (or which Es) is/are picked out by T.

Given this, it should also be clear that the fact that one can take the same (or a relevantly similar) action with respect to different entities does not imply any confusion in the assignment of content. Content is not determined by some set of focus-neutral action descriptions — like hit, taste, or run — such that all tokens supporting the same set of actions have the same content and pick out the same entity. Not everything that one can taste is thereby represented as equivalent. Different tokens can have overlapping sets of supported actions but take different foci, and thus represent different entities. Similarly, there can be tokens with non-overlapping sets of supported actions that nevertheless take the same focus, and thus represent the same thing for different purposes or in different circumstances. This arrangement does not introduce ambiguity, but is a sign of the cognitive richness and flexibility of complex creatures.

Implausible Content Ascriptions

Finally, Roberts raises a version of Pietroski's (1992) argument that some of the content ascriptions dictated by consumer-semantic theories might be implausible given relevant facts about the creatures involved. The objection goes like this: suppose a given animal evolved a sensitivity to a particular frequency of red light, because regions reflecting this light are free of this creature's only natural predator. According to standard teleo-semantic accounts the proper function of this evolved sensitivity is to enable predator-avoidance, and the content of the red-experiencing state must therefore be something like: "no predators here!" But by hypothesis, these animals are not capable of representing predators. They are capable of detecting that certain shade of red and have a preference for regions that have it. So there appears to be a contradiction between the content ascribed by the theory, and the actual representational capacities of the creature.

Although the guidance theory does not rely on any concept of proper function to determine content, Roberts thinks that it may be vulnerable to a similar objection. For suppose further that many different entities in the environment have the color in question — rocks, trees, prairie voles — and so when the animal sees the color and uses it to locate (guide its locomotion toward) safe regions, it is in fact guiding its action with respect to many different entities. If so, Roberts

argues, the guidance theory dictates that the animals represent rocks, trees, and prairie voles; but these animals are capable of no such feat.

The trouble with this argument is that it assumes that the animal's actions are taken with respect to whatever happens to be spatio-temporally contiguous with the target of the action, as if its actions in fact target the largest independent thing containing the action's focus, rather than the focus itself — in this case, that particular color red. But consider: if a face-sensitive animal is confronted with the face-vase illusion, that circumstance does not magically transform its face-targeting actions into face-and-vase targeting actions. However it may look to those of us capable of discriminating faces and vases, faces are and remain the foci of the actions guided by its relevant face representations, even when those faces are arrayed to portray vases. The same is true of our red-seeking animal here — however it might initially look to us, the animal guides its actions with respect to redness, as careful analysis of its behavior would show. And insofar as its actual behavior *does* show this, as Roberts stipulates it will, then this means that the animal does not — cannot — take actions with respect to rocks, trees, and prairie voles, and does not guide its actions with respect to these things or any of their tree-y, rocky, or vole-y features. Instead, its action is guided with respect to — and it therefore represents — that red. The example thus yields no contradiction between the animal's capacities and the guidance theory's ascription of content.

Affordances and the Guidance Theory

For the reasons outlined above, we do not think that Roberts' arguments against the guidance theory hit their mark, and so we do not believe that it needs to be modified along the lines that Roberts proposes. Nevertheless, the question of the relation between affordances and the intentional states analyzed in the guidance theory is indeed important and interesting, and we would like to take this opportunity to lay out what we take to be the right approach to the issue. To do that, we need first to get clear on what affordances are.

What Affordances Are

Affordances are opportunities for behavior by particular animals. Gibson (1966) argued that affordances are the primary entities that are perceived, and perceiving affordances is perceiving the meaningful world. Importantly for current purposes, affordances are not merely entities in the environment, and they are also not projections of meaning by animals onto a merely physical environment. Affordances are features of animal–environment systems, and exist in such systems only in virtue of animals that have the appropriate abilities to perceive and take advantage of them. There is some controversy within the ecological psychology

community about how exactly to understand affordances. Turvey (1992) suggests that affordances are dispositional properties of the environment, where the actualizing conditions include the abilities of animals. In contrast, Chemero (2003, 2009) suggests that affordances are relations between the abilities of animals and environmental features. Although we favor the relational view, and will continue our discussion in terms of it, it should be noted that for present purposes, the disposition and relational views are not different from one another (indeed, Chemero and Turvey [2007] argue that the views are nearly identical). Whichever side in this insider argument is correct, everyone within the ecological psychology community agrees that (1) affordances exist only within coupled animal–environment systems and (2) when animals perceive affordances, they perceive something about both themselves and their environment.

So, here, we will take affordances as relations between features of the environmental situations and the abilities of animals. An affordance is a function or mapping from abilities (e.g., the ability to read) and situational features (e.g., the presence of text, appropriate lighting conditions, and so on).

$$\text{Affordance} = \text{< Ability, Situational Features >}$$

Note that the very same environment will have different affordances for different animals. Different animals, with different abilities, may have physically co-located but nonetheless non-overlapping niches. For example, a human and a bacterium may share a physical location (as when a bacterium is inside a human), but their niches will not overlap. Gibson argues that this is the way to make sense of the mutuality of animals and environments. An animal's abilities imply an ecological niche. Conversely, an ecological niche implies an animal.

When an animal perceives an affordance for a particular action, it perceives a relation between an ability that it has and an environmental situation in which that ability can be exercised. This relation is not static. The affordances that are available to a particular animal change quickly, and the change is caused by prior affordances and action upon those affordances. Acting changes the situation, which creates new affordances for the animal.

This becomes clearer in an example. Consider a healthy adult human approaching a building. The environmental situation in which the person's standing body is facing a staircase, with riser heights that are less than .87 the person's leg length (Warren, 1984), means that the person's stair climbing abilities can be put into action; that is, the affordance *climbability* is present and perceivable. When the person perceives and acts on the affordance repeatedly, climbing the stairs, her position with respect to the building's entrance door has changed. This change in the situation has made other of the person's abilities applicable and further affordances available. Thus the person engages her ability to reach, which alters the situation by placing her hand in contact with the door handle,

which makes her ability to grasp applicable and the affordances for grasping available and perceivable. Grasping changes the situation so that abilities to turn become applicable, and turning affordances are available to the person. Subsequent actions lead to further changes in the situation, further changes to which abilities are applicable, and further changes to which affordances are available (*pushability*, affordances for aperture passing, and on and on).

Of course, the particulars of this flow would be different for different animals. A healthy adult human, a human infant, a dog, and a bacterium in front of this same building would all exhibit a very different flow of situation-ability relations and the actions that cause changes to them.

With this basic description of affordances in place, we can turn to the relationship between affordances and the guidance theory. We will see that the relationship is very different from what Roberts imagines.

Affordances and Guidance

Allow us first to say what the relationship between affordances and the representing tokens of the guidance theory *is not*. It is not the case that guidance theory tokens *are* affordances, nor is it the case that guidance theory tokens *represent* affordances. Affordances are perceived by organisms in circumstances, and to say that organisms represent these perceptions adds an unnecessary layer of cognitive machinery between organisms and the world. Moreover, we should remember that not all animals have — or need — representations, and those that *do* have them don't necessarily need them for every action they take. Finally — and perhaps this should go without saying — although the guidance theory says that representations are such in virtue of their role in guiding action (all representations are action-guiders), it does not claim that everything that guides action is therefore a representation. All this is simply a way of reminding the reader that the guidance theory offers an analysis of the nature and function of intentional states; it is not a comprehensive account of the sensorimotor and cognitive control systems of animals.

This being said, there are at least four ways in which guidance theory tokens are related to affordances. We'll briefly outline each first, then explain them more fully in the context of an extended example.

1. Guidance theory tokens can act in concert with perceived affordances in guiding an animal's actions.
2. Affordances are among the relationships between an animal and its environment that help determine the focus of its actions.
3. Guidance theory tokens can partly consist in traces left by perceiving affordances. Among other things, this can allow for action-guidance in the absence of the represented entity, and thus in the absence of the

affordance. Note this is not a case of representing the affordance; one represents an entity with the aid of prior experiences of the affordance.

4. As noted above, affordances are relations between paths through a state space of possible actions (animal abilities) and a state space of possible circumstances. Paths through action-space imply paths through circumstance-space (because actions affect circumstances), as well as the reverse (because circumstances shape the set of possible actions). Guidance theory tokens are *also* associated with sets of actions they support — that would be guided by the token — given sets of circumstances in which such action might be called for.[3] Thus, one can conceptualize guidance theory tokens as occupying regions of both the action-space and the circumstance-space, corresponding to the actions they could guide in those circumstances. As an extension of (3), then, and given that sensitivity to affordances is at the same time sensitivity to the contingent dynamic relations between the action and circumstances spaces, the perception of affordances might be part of the process whereby guidance theory tokens come to be associated with their proper regions of these coupled state spaces.

5. Given (4), there may be another way to analyze what it is for guidance theory tokens to be in error. A guidance theory token might be said to be in error when it is associated with the wrong location(s) in the coupled action-circumstance state-space. That is, the paths through action-space that would be taken as a result of guidance provided by a guidance theory token are inappropriate to the circumstances of the animal, in that they result in actions that will fail.

Here again, this will be clearer given an example. Imagine someone playing tennis, watching the ball sail over the net, and preparing a backhand return.

---

[3]The class of actions a token T supports is relative to the kinds of circumstances C where the system is prepared to use the token for guidance. It consists of all the actions the system can initiate or modulate in C due to its processing of T. Let us label this class of supported actions A_{supp}.

> *Definition 10:* An action A is a member of the class of actions, A_{supp}, supported by a token T used by a subject S in circumstances C if, and only if, S in C would use T for guidance regarding the initiation or manner of execution of A.

We should think of the actions in A_{supp} as focus-neutral descriptions of an action in need of association with a focus in particular initiations. So, for example, if in some circumstances a system is prepared to use a token for guidance in running, the action *running* is the focus neutral description. If the specific initiation of this action occurs when the focus of the action is a bear, the focus-neutral action "running" is initiated as the focus-specific action "running away from a bear." Actions obtain a focus in the way discussed above, by determining the ultimate entity being monitored by the subject in support of its action. Furthermore, since subjects do not initiate actions at random, for each action in A_{supp}, there will be a (possibly very large but) finite set of circumstances capable of triggering the initiation of the action. We can call this set of triggering circumstances A_{circ}. The number of triples $<A \in A_{supp}, C \in A_{circ}, Focus>$ representing supported actions A initiated in circumstances C with focus F provides a class of counterfactual *action scenarios*, A_{scene}, in which the token T provides guidance for a subject. These are the action scenarios in which T *participates*. (Anderson and Rosenberg, 2008, p. 75)

There are undoubtedly circumstances in which this action is driven entirely via the perception of affordances and the appropriate behavioral responses they suggest — positioning the body here, the front foot there, angling the racquet just so. And there may also be circumstances in which this behavior is guided almost entirely by guidance theory tokens, perhaps especially when one is just learning the sport, trying explicitly to enact a coach's advice. Here the circumstances may afford very little, and one is actively ignoring what it *does* afford as inappropriate to the skills one is trying to acquire. But more often behavior is shaped by a combination of affordances and guidance theory tokens.[4] For instance, the ball in these circumstances may afford a variety of possible swings, and the selection among them may rely in part on information about the responsiveness of this particular brand of ball given the number of minutes it has been in play (not to mention knowledge of the past play of one's opponent, and one's overall strategy for the match).[5] Thus the affordance(s), along with a guidance theory token representing the ball, together determine the character of the resulting behavior. Note also that the particular affordances to which the player is attuned; the behaviors that result from this attunement; and the sensorimotor feedback loops into which the player thereby enters, help to fix the focus of the action — and thus determine the entity being represented by any guidance theory tokens that are being used to guide the action.

  Now suppose that after the match is over, the player analyzes the flow of the match, her moment-by-moment choices, imagining ways she could improve for the next contest. Here part of what informs her representations of the circumstances in which she wishes to improve her (future) actions are traces left by the experience of the affordances while in the actual circumstance. For a ball looking just so at that part of the court in this situation it is perhaps better to hit it with *this* top-spin, rather than that, or at this angle, with this force. She need not be representing what the ball affords, but is using traces left by earlier perceptions of the affordance to help represent the ball, to guide her actions (her training, her planning, her visualization) with respect to the ball. One result of this process may be to tune the set of actions that a guidance theory token supports, or the precise circumstances in which it supports them, thus conceptually altering the position of the token in the coupled action-situation possibility space. This in turn would change the likely trajectories of future

---

[4]Note again that the relative proportion of circumstances falling into these three categories is a matter of some debate in the field, including between the two authors of this paper, with some expecting representations to play a more, and others a less prominent and frequent role in guiding behavior.

[5]The example should not be taken to imply that guidance theory tokens never guide on-line action, or only do so for novices, or when affordances are missing or weak. The focus of the example is illuminating circumstances in which affordances and guidance theory tokens can play a complementary role in shaping behavior.

behaviors triggered in those circumstances, guided by the guidance theory tokens therein engaged. This may also change the particular affordances to which the player will be attuned in the future — both because she may selectively attend to a different sub-set of the affordances already available in those circumstances, and because her training may change her abilities, and thereby change the set of affordances that are in fact perceivable.

Finally, suppose that the reason the player is so intent on reimagining the match is because that backhand went awry, causing her to lose the point and the subsequent game, set and match. That is, the action failed in its intent, due at least in part to guidance provided by a guidance theory token. That token was in error, and one way to understand what it means for it to have been in error is that it occupied the wrong place in the coupled action-circumstance possibility space, causing the player to attune to the wrong sub-set of affordances, monitor the wrong aspects of the circumstance, and select the wrong parameters for her swing.

Conclusion

In this paper we have argued that, while Roberts' particular objections to the guidance theory do not go through, he has nevertheless raised an important issue: the relationship between affordances and the action-guiding representations analyzed by the guidance theory. Affordances play a prominent role in the anti-representationalist accounts offered by theorists of embodied cognition and ecological psychology (Chemero, 2009). Theories like the guidance theory are motivated in part by a desire to respond to the critiques of representationalism offered in such accounts, without giving up entirely on the idea that representations are an important part of the cognitive economy of many animals (see Svensson and Ziemke, 2005, for a brief summary of the debate). Thus, explorations of whether and how such accounts can in fact be related and reconciled potentially offer to shed some light on this ongoing controversy. Although the current essay hardly settles the larger debate, it does suggest that there may be more possibility for agreement than is often supposed.

References

Aglioti, S., DeSouza, J., and Goodale, M. (1995). Size illusions deceive the eye but not the hand. *Current Biology, 5,* 679–685.

Amazeen, E.L., and Turvey, M.T. (1996). Weight perception and the haptic size–weight illusion are functions of the inertia tensor. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 213–232.

Anderson, M.L., and Rosenberg, G. (2008). Content and action: The guidance theory of representation. *Journal of Mind and Behavior, 29,* 55–86.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology, 15,* 181–195.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, Massachusetts: The MIT Press.

Chemero, A., and Turvey, M.T. (2007). Gibsonian affordances for roboticists. *Adaptive Behavior, 15,* 473–480.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Milner, D., and Goodale, M. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Pietroski, P. (1992). Intentionality and teleological error. *Pacific Philisophical Quarterly, 73,* 267–281.

Proffitt, D.R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1,* 110–122.

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

Roberts, T. (2009). Guidance, selection, and representation: Response to Anderson and Rosenberg. *Journal of Mind and Behavior, 30,* 291–300.

Shockley, K., Carello, C., and Turvey, M.T. (2004). Metamers in the haptic perception of heaviness and moveableness. *Perception and Psychophysics, 66,* 731–742.

Svensson, H., and Ziemke, T. (2005). Embodied representation: What are the issues? In B.G. Bara, L. Barsalou, and M. Bucciarelli (Eds.), *Proceedings of the XXVII Annual Conference of the Cognitive Science Society* (pp. 2116–2121). Stresa, Italy.

Turvey, M.T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology, 4,* 173–187.

Warren, W.H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 683–703.