Article

PERCEPTION

Perception
2020, Vol. 49(9) 905–925

DOI: 10.1177/0301006620946532
journals.sagepub.com/home/pec
SAGE

Visual and Haptic Perception of Affordances of Feelies

Catherine Dowell and Alen Hajnal
School of Psychology, University of Southern Mississippi, United States

Wim Pouw
Donders Institute for Brain, Cognition and Behaviour, Radboud University, the Netherlands; Max Planck Institute for Psycholinguistics, the Netherlands

Jeffrey B. Wagman
Department of Psychology, Illinois State University, United States

Abstract

Most objects have well-defined affordances. Investigating perception of affordances of objects that were not created for a specific purpose would provide insight into how affordances are perceived. In addition, comparison of perception of affordances for such objects across different exploratory modalities (visual vs. haptic) would offer a strong test of the lawfulness of information about affordances (i.e., the invariance of such information over transformation). Along these lines, "feelies"— objects created by Gibson with no obvious function and unlike any common object—could shed light on the processes underlying affordance perception. This study showed that when observers reported potential uses for feelies, modality significantly influenced what kind of affordances were perceived. Specifically, visual exploration resulted in more noun labels (e.g., "toy") than haptic exploration which resulted in more verb labels (i.e., "throw"). These results suggested that overlapping, but distinct classes of action possibilities are perceivable using vision and haptics. Semantic network analyses revealed that visual exploration resulted in object-oriented responses focused on object identification, whereas haptic exploration resulted in action-oriented responses. Cluster analyses confirmed these results. Affordance labels produced in the visual condition were more consistent, used fewer descriptors, were less diverse, but more novel than in the haptic condition.

Corresponding author:
Catherine Dowell, University of Southern Mississippi, 118 College Drive #5025, Hattiesburg, MS 39406, United States.
Email: catherine.dowell@usm.edu

Keywords
affordance, haptic, visual, feelies, cluster analysis

Date Received: 11 February 2020; accepted: 10 July 2020

Gibson (1962, 1963) described perception as the active process of obtaining information from the environment. The detection of information yields the discovery of the meaning of contexts, objects, or events. Affordances—opportunities for behavior given the fit between properties of the environment and capabilities of the organism—are activity-specific meanings (Turvey, 2019). For example, for a piece of fruit on a tree to be retrievable, it must be small enough for the person to grab, and the person must be tall enough and be able to reach far enough to grasp the fruit. Without these multifaceted conditions being met, the fruit does not afford retrieving. Affordances, according to Gibson (1979), are perceived through exploratory behaviors of multiple perceptual systems. For example, whether an aperture can be passed through can be perceived by viewing that aperture (Warren & Whang, 1987) or by listening to sounds projected through that aperture (Gordon & Rosenblum, 2004) and whether an inclined surface can be stood on can be perceived by viewing that surface or by probing that surface with a limb or a hand-held object (Hajnal et al., 2016; Malek & Wagman, 2008). Such comparable performance across perceptual systems is to be expected given the hypothesis that patterns in ambient energy arrays are lawfully related to the fit between animal and environment. Such patterns are invariant across transformations (including the identity of structured energy media; Gibson, 2002). That is, not only is the structure in a given patterned energy array lawfully related to the animal-environment fit that created it, but such structure in a given energy array is lawfully related to structure in other energy arrays.

Research has also shown that many affordances of objects (including throwableness, hammer-with-ableness, and poke-with-ableness) can be haptically perceived—especially when the person is able to explore the object by hefting, wielding, or otherwise manipulating it (Bingham et al., 1989; Carello & Turvey, 2000; Hajnal et al., 2020; Harrison et al., 2011; Lederman & Klatzky, 1987; Wagman & Carello, 2001). Many of these same affordances of objects (as well as many others, including catchableness, stretchableness, scoop-with-ness, floatability, and stand-on-ableness) can be visually perceived—especially when the person is able to explore the object by moving their head, body, or eyes with respect to the object (Hajnal et al., 2020; Michaels & Oudejans, 1992; Ye et al., 2009; Yu et al., 2010). Importantly, however, the objects used in most of these studies have been human-made objects. Most human-made objects are created with the intention to serve specific purposes (Dennett, 1987) and therefore usually have salient designed affordances (Rachwani et al., 2020). However, natural objects (e.g., rocks, twigs, branches, etc.) have no designed affordances (salient or otherwise). How are affordances of these types of objects perceived by vision, by touch, or otherwise? It is not clear how one would go about systematically studying perception of affordances of such objects. One solution to this problem would be to use objects that are human-made but that were not designed with any specific function in mind and with which perceivers are completely unfamiliar.

Objects fitting these exact specifications were developed by Gibson (1962) and are known as feelies. All feelies are uniquely distinctive in shape but share the common characteristics of having six protuberances, a convex underside, and being approximately equal in size and mass (see Figure 1). Feelies were described by Gibson as “unlike any familiar object, or anything with a name” (Caviness, 1962, p. 3), making them ideal stimuli to investigate affordance perception while avoiding the issue of designed affordances.

Whereas past research using feelies has focused on perceptual discrimination tasks (Davidson et al., 1974; Goodnow, 1971; Fairhurst et al., 2018; Experiments 2 and 3 of Norman et al., 2012), the goal of this study was to focus on perception of affordances. The few studies that compared visual and haptic perception of tool use have discovered not only that information obtained visually tends to be more salient than information obtained haptically, but also that functional properties of objects are accessible by both modalities (Michaels et al., 2007). Importantly, the perception of affordances is multimodal (Mantel et al., 2015; Streit, Shockley, & Riley, 2007; Streit, Shockley, Riley, et al., 2007) and suggests that each perceptual modality samples information from a global energy array (Stoffregen et al., 2017). The multimodal nature of perception of unfamiliar novel objects is also

Figure 1. Three-Dimensional-Printed Versions of the Original 10 Feelies.
Left column (top to bottom): feelies 2, 4, 6, 8, and 10. Right column (top to bottom): feelies 1, 3, 5, 7, and 9. All feelies were painted a homogenous dark gray and had a smooth, slightly rubber-like surface texture.

complemented by neurophysiological evidence for activation of visual cortical areas during haptic manipulation (James et al., 2002). Given these findings and the ambiguity and novelty of the objects, it was predicted that in this study multiple affordances will be identified for each object. To the extent that functional properties of objects can be detected by different modalities sampling from the global array, it was also predicted that affordances identified haptically and visually will be organized into significantly distinct but overlapping categories. Specifically, we hypothesized that observers' perception of feelies' possible uses will differ in terms of the total number of unique uses identified (assessed with word frequency analysis), the semantic space of the feelies' uses discovered through each modality (semantic analyses), and the type of linguistic categories that may emerge from the responses (word class analysis).

Method

Participants

There was a total sample of 32 students from the University of Southern Mississippi. Participants were all recruited using the university's SONA participant pool and received credit for participation that could be placed toward classes in which they were currently enrolled. Data from 3 participants were excluded: 2 for incomplete data resulting from technical issues and 1 for failure to follow instructions, leaving 29 viable participants (22 females; 7 males). Individuals were required to have normal or corrected-to-normal vision. Prior to the start of the experiment, participants were asked to complete a survey to assess handedness (Coren, 1993). The results of the survey showed that 28 students were right-handed, 3 were left-handed, and 1 was ambidextrous.

Materials and Apparatus

Experimental Stimuli. The 10 original feelies (Gibson, 1962) were three-dimensionally scanned and were then three-dimensional (3D) printed to be used as stimuli (Phillips & Egan, 2016; Norman et al., 2012). The objects were printed via a Tevo 3D Tarantula 3D printer and are composed of polylactic acid with 70% of the interior filled in and a printing increment of 0.2 mm. The original feelies are shown in Figure 1. All feelies have a slightly different shape, but the topological configuration is identical. This means that all objects have a smooth surface curvature all throughout and six identifiable apexes of varying proportions. Table 1 provides information about the physical characteristics of the feelies used in the experiment.

Materials. The participant and experimenter sat at opposing sides of a table, separated by a black felt curtain in the middle (see Figure 2). The black cloth occluding screen was used to prevent the participant from seeing the object being explored in the haptic condition and to prevent him or her from seeing more than one object at a time in the visual condition. The screen also occluded the researcher's face, preventing participants from attempting to gauge reactions to their responses. Participants in the haptic condition wore thin nitrile gloves to minimize any potential minute textural differences from 3D printing between the set of objects.

Design

This study used a mixed design with Modality (visual and haptic) as a between-subjects independent variable with two levels and Object (10 feelies) as a within-subjects independent

Table 1. The Mass, Volume, and Density of Each Feelie.

Object number	Mass (kg)	Volume (mL)	Density (kg/cm^3)
1	0.09986	150	0.00066
2	0.09804	150	0.000653
3	0.07126	125	0.000570
4	0.07126	150	0.000475
5	0.08896	125	0.000712
6	0.07126	100	0.000712
7	0.06491	150	0.000433
8	0.07217	150	0.000481
9	0.05401	100	0.000540
10	0.08125	125	0.000650
Mean	0.07729	132.5	0.000589
SD	0.01543	20.58	0.000103

Note. SD = standard deviation.

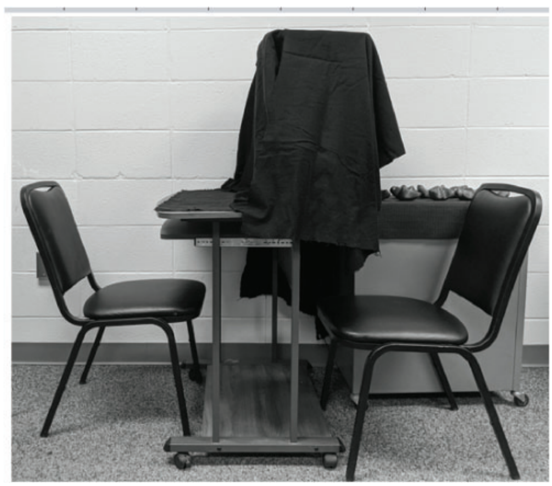

Figure 2. Experimental Setup With a Black Occluding Screen Between Participant and Experimenter Seated at Opposite Sides of a Table.

variable. Participants in the visual condition viewed the objects but were unable to touch them. Participants in the haptic condition touched the objects but were unable to see them. Each participant explored each object 3 times over a series of three blocks for a total of 30 trials. Each block included all objects presented in a random order, controlling for any order effects. All participants were randomly assigned to a modality condition—14 were assigned to the haptic condition and 15 were assigned to the visual condition.

Procedure

After reading and signing the informed consent document, participants filled out The Lateral Preference Survey (Coren, 1993). Participants in the haptic condition were asked to put on nitrile gloves, place their hands behind the occluding screen, and confirm that they could not see anything beyond the screen. Participants in the visual condition were told that they were free to move their head and torso as much as they would like, so long as they remained seated. They were also reminded to only look at each object and not touch or move the object in any way.

On a particular trial, participants were presented one of the 10 objects and were permitted to either observe or handle the object (depending on condition) for 1 minute before it was taken away. In the vision condition, participants observed each feelie in a static, randomly oriented position under standard indoor lighting conditions. Participants were told not to touch the object during visual exploration (see Figure 3) but were allowed to move their head and torso. In the haptic condition, the participant was permitted to use both hands to move and manipulate the object, so long as it stayed behind the screen and in the area of the table surface (e.g., tapping it against the surface of the table, rolling it around on the table, spinning it on the table, and tossing it between hands, see Figure 4). Participants in both conditions were asked to name any and all possible uses of that object. They were allowed to start verbally describing or naming possible uses of the objects at any time before the 1-minute exploration time expired and were able to continue listing possible uses even after the 1-minute exploration time was over. They indicated when they were finished by verbally replying “done”. There was no time limit on responses. All responses were audio recorded and transcribed by researchers.

Data Analysis

Text Analysis. Once all responses were transcribed, the possible uses for each object were analyzed for frequency by condition and block. Responses were grouped by object, by modality condition, and by blocks of trials. Table 2 presents an example of how the responses were organized for Object 1 in the visual condition. For example, from the table, one can discern that the word “decoration” was used 4 times to describe a potential use for Object 1 in Block 1 of the visual condition. It is important to note that these responses may have come from different participants. All 10 objects were analyzed according to this procedure.

The total number of uses listed for a particular object in a given condition was obtained by summing the number of unique words (denoting potential uses listed) within each block. Shared responses between blocks for each object were responses that were listed at least once

Figure 3. The Experimental Setup During the Visual Condition, With a Participant (on the Left) Viewing One Object at a Time in Front of the Occluding Screen.

in all three blocks. Responses that were listed in only two of the three blocks were not included. Shared responses were summed by counting the number of responses that appeared across all three blocks. For example, if the word "decoration" appeared in the list for Block 1, Block 2, and Block 3 for Object 1 in the visual condition, it would be considered a shared response for that object. The total number of unique responses (listed words) for each condition was calculated by summing the number of listed words for each of the 10 objects. To compare shared responses between visual and haptic conditions, a proportion of shared to total responses was calculated for each condition and each object. Within each block, the number of an object's shared responses was divided by the total number of unique responses for that object in that block of a given modality condition to obtain a proportion of shared responses for each object in each block. Proportion of shared responses provided information about perceived use within a particular modality. We also calculated how many responses were shared across visual and haptic modalities. This was done by comparing all the responses to each object in the visual and haptic modality. While the shared responses across blocks provide information about the consistency with which participants perceive objects *within a single modality*, the number of shared responses across two modalities could provide information about consistency with which different participants responded to the objects *in different modalities*. This could be a more powerful test of the invariance of perception, given the between-subjects nature of the modality manipulation.

Response Time Analysis. Response times were recorded for each trial. The beginning and end times of trials were obtained in two steps: Initially, the experimenter delineated the start and end of each trial by manually clicking a mechanical pen during the experiment. Begin-time clicks were initiated at the moment when the object was placed in front of the participant (visual condition) or when the participant made initial tactile contact with the object (haptic condition). End-time clicks occurred following the verbal response of "done" given by

Figure 4. The Experimental Setup During the Haptic Condition.
The participant (on the left) places his hands behind the black occluding screen and handles one object at a time.

Table 2. Example of Text Analysis Based on Responses From All Participants in the Visual Condition to Object 1 Across Three Blocks of Trials.

Block 1		Block 2		Block 3	
Response	Frequency of occurrence	Response	Frequency of occurrence	Response	Frequency of occurrence
Decoration	4	**Toy**	4	**Decoration**	4
Throwing	4	**Toss**	3	**Toy**	4
Paperweight	3	**Decoration**	3	**Paperweight**	3
Toy	2	**Throwing**	3	Model	3
Massager	1	Hold	3	**Stress**	2
Weapon	1	**Playing**	3	**Throwing**	2
Toss	1	**Stress**	2	**Art**	2
Stress	1	**Paperweight**	2	Ball	1
Look	1	**Squeezing**	2	Spinning	1
Gripping	1	Ball	1	Souvenir	1
Playing	1	Spinning	1	Tool	1
Standing	1	Turning	1	**Toss**	1
Art	1	Rolling	1	Pulling	1
Relief	1	Crafting	1	Holding	1
Pet	1	Bouncing	1	Bouncing	1
Squeezing	1	**Art**	1	**Playing**	1
Doorstop	1	**Relief**	1	Mold	1
Model	1			**Relief**	1
				Hammer	1
				Squeezing	1
				Paint	1

Note. The words in boldface are responses that are shared across all three blocks (e.g., paperweight), suggesting consistency in perception of the object's potential uses. In this example, there are 10 shared responses across blocks. The responses are rank ordered by the frequency of occurrence. Filler words and noncontent words were not included in this analysis.

participants.[1] The duration of verbal responses was determined using EUDICO Linguistic Annotator (ELAN) (version 5.2, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands), a professional audio transcription and annotation software. Audio recordings of each participant were loaded into ELAN, where an experimenter then manually selected portions of the recording that were marked as verbal responses (content words referring to object use) and indicated the duration of the uttered word on the software timeline.

Natural Language Processing. Spacy (https://spacy.io/; e.g., Srinivasa-Desikan, 2018) and a `R` wrapper `SpacyR` (Benoit et al., 2018) were used for all Natural Language Processing (NLP) procedures. For an overview of the English corpus models, see https://spacy.io/models. For the part of speech tagging, we used a faster but smaller corpus database ("en_core_sm")[2] and for the word2vec, we used the largest corpus possible ("en_core_lg"). The R-script for the word class analysis was also used to lemmatize[3] the potential uses mentioned, which means that inflected forms of a term (e.g., throwing, thrown) are reduced to a single base term ("throw"). These lemmatized tokens would then be submitted for word2vec computations (this script can be found on the open science framework (OSF), https://osf.io/xgwub/). The python-based word2vec procedure can be found at https://osf.io/qyj5n/, and the R-script taking as input word2vec comparisons so as to plot semantic spaces and compute clusters can be found at https://osf.io/yaq2h/.

Results

Response times and responses (both shared responses and total number of unique responses) were analyzed. The number of shared responses for each object was analyzed using a between-subjects *t* test to compare the amount of overlap in listed words between conditions. The number of shared responses listed for each object over blocks of trials did not differ in the visual ($M = 10.9$, standard deviation [SD] $= 1.45$) and haptic condition ($M = 9.6$, $SD = 2.55$), $t(9) = -1.494$, $p = .169$. Remarkably, the number of shared responses *across* modalities ($M = 20.1$, $SD = 4.82$) was higher than *within* the visual condition, $t(9) = 5.84$, $p < .001$, and the haptic condition, $t(9) = 6.59$, $p < .001$, respectively. The cases for the *t* tests were the 10 feelies, not participants, thus the degrees of freedom equaled 9.

We conducted a 2(Modality) × 3 (Block) mixed-design analysis of variance on the total number of responses used to describe each object. Participants provided significantly more responses for each object in the haptic condition ($M = 25.1$, $SD = 0.61$) than in the visual condition ($M = 22.3$, $SD = 0.57$), $F(1, 9) = 11.478$, $p < .008$, $\eta_p^2 = .561$. There was also a significant Modality × Block interaction, indicating that the difference in the number of responses between conditions varied by block, $F(2, 18) = 8.786$, $p < .002$, $\eta_p^2 = .494$ (see Figure 5). Specifically, participants provided significantly more responses in the second block of the haptic condition compared with the second block of the visual condition. There was no significant main effect of Block, $F(2, 18) = 2.902$, $p = .081$.

Although participants listed more potential uses per object in the haptic condition, there was a significantly higher proportion of shared responses per object across blocks in the visual condition ($M = 0.49$, $SD = 0.01$) than in the haptic condition ($M = 0.39$, $SD = 0.02$), $F(1, 9) = 9.611$, $p < .013$, $\eta_p^2 = .516$. There was also a significant Modality × Block interaction, indicating that the difference in the proportion of shared responses between conditions varied by block, $F(2, 18) = 7.752$, $p < .004$, $\eta_p^2 = .463$ (see Figure 6). Participants provided a higher proportion of shared responses in the second block of the visual condition compared with the haptic condition than in other blocks. There was no significant main effect of Block, $F(2, 18) = 2.675$, $p = .096$.

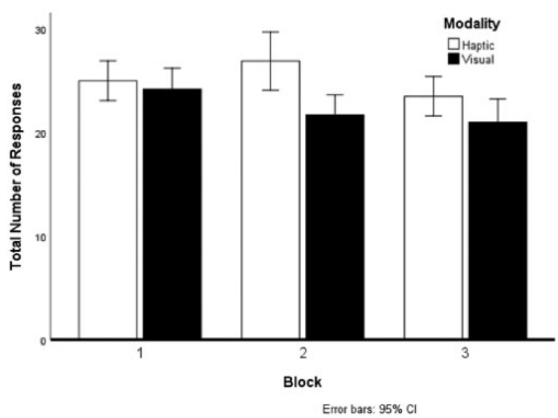

Figure 5. Total Number of Responses as a Function of Modality Across Blocks of Trials.
Error bars represent 95% confidence intervals.
CI = confidence interval.

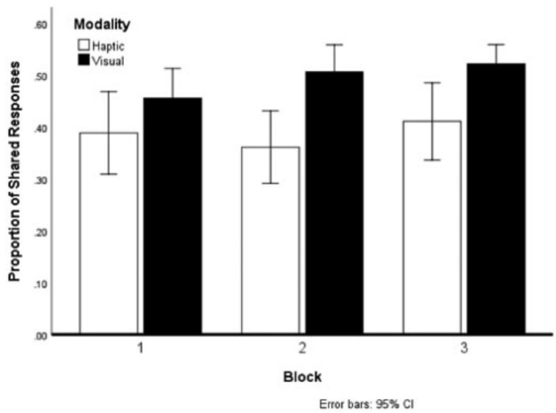

Figure 6. Proportion of Shared Responses as a Function of Modality Across Blocks of Trials.
Error bars represent 95% confidence intervals.
CI = confidence interval.

Response times were numerically longer in the haptic condition ($M = 57.14$, $SD = 25.07$) than in the visual condition ($M = 45.74$, $SD = 25.94$), but this difference was not significant, $F(1, 27) = 2.113$, $p = .158$. Response time decreased over blocks of trials, $F(2, 54) = 13.13$, $p < .001$, $\eta_p^2 = .33$.

Semantic Content Data Analysis

In addition to duration and quantity, responses were also analyzed for semantic content. Wordclouds displaying the most frequently used words in each modality were created, with more frequently used words being larger than less frequent words (see Figure 7). Inspection of these wordclouds suggests that the haptic corpus contained more concrete, action-based words such as "throw" and "hold," while the visual corpus had more abstract, object-based words like "decoration," "toy," and "stress." Similarly, the 20 most frequent words in each condition can also be seen in Figure 8, which shows the proportion that each response occurred relative to the total number of affordance responses given in that condition.

Natural Language Processing

To further investigate semantic content of the responses, we conducted several NLP procedures which allowed us to estimate word class differences (i.e., Part of Speech Tags). We also conducted an analysis of the conceptual affordance space using "word2vec" so as to see how the lemmatized responses were semantically related (as predicted by a neural network trained on word associations of a large English corpus).

Word Class Analysis

For producing word classes for each response, we entered individual words in the way they appeared in their original phrases (i.e., without lemmatization), and we then let Spacy predict the word class (we used the "en_core_web_sm" corpus for this procedure). It should be noted that while currently relevant Verb vs. Noun word classes are straightforward (e.g., "throwing", "decorating"), in some cases the word class is ambiguous when presented

Figure 7. Wordclouds (Created Using the Voyant Tools Software) for the Visual (Top Panel) and the Haptic Modality Condition (Bottom Panel).
The most frequent words in the visual corpus were decoration (109), paperweight (86), toy (71), throw (63), and stress (59). the most frequent words in the haptic corpus were throw (141), hold (89), paperweight (55), toy (53), and decoration (36). the word frequencies are listed in parentheses.

without its accompanying sentence, for example, "glue" and "massage "can serve as a noun or a verb. In these ambiguous cases, our NLP procedure based its word class prediction on biases in the corpus it was trained on.

Possible use responses were tagged as either nouns or verbs. Based on these categorizations, the visual condition was found to contain more nouns (58% vs. 41% in the haptic condition) and fewer verbs (42% vs. 56%) than in the haptic condition. In addition, a hierarchical mixed-effects logistic regression (Gelman & Hill, 2007; Jaeger et al., 2011) was conducted to investigate the relationship between modality, objects, and Part of Speech (verb or noun) as an outcome variable. Modality was a significant predictor of verbal responses ($B = 0.716$, standard error $[SE] = 0.300$, $p = .020$, odds ratio $= 2.046$), suggesting that participants uttered more nouns than verbs in the visual condition. These results further support the pattern of differences in semantic content between conditions seen in the wordclouds.

Semantic Similarity Combined With Graph

We further conducted a similarity analysis of all the lemmatized responses that were produced by the participants. Using Spacy, we compared all unique responses with one another

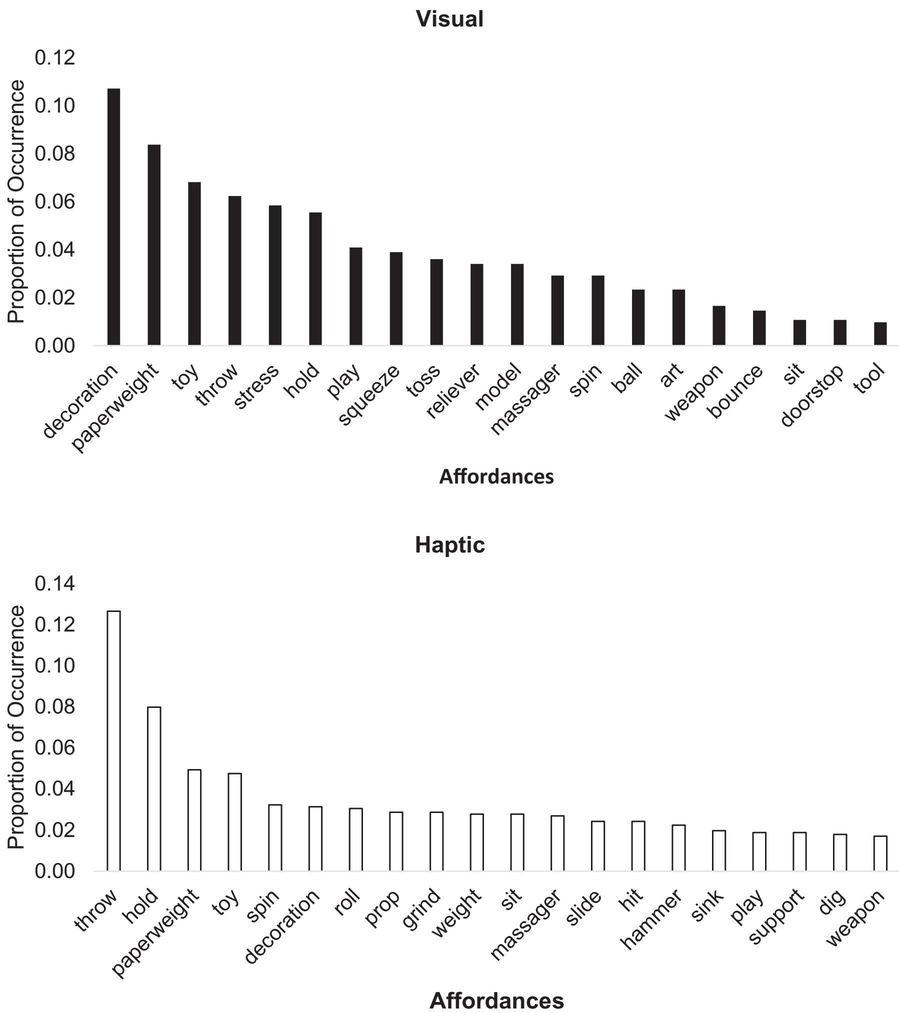

Figure 8. The 20 Most Common Responses in Both Visual (Black) and Haptic (White) Conditions. Response frequencies are shown as proportions of the total number of times each response was given relative to the total number of responses given in each condition.

by producing a semantic similarity score for each compared pair of words (https://spacy.io/usage/vectors-similarity). Semantic similarity scores ranged from 0 (*not similar*) to 1 (*similar*) and are roughly based on the degree to which both lemmas appeared in the corpus in similar contexts. To be precise, word2vec procedures make use of neural networks which are trained such that they map out probability distributions of a token's (e.g., sentence or word) co-occurrence with other tokens they typically co-occur with in a corpus (i.e., the token's context). The model can then represent words or sentences as vectors in a multidimensional semantic space, with each vector being a representation of their meaning (as defined by the typical context they pop up in). Distance (i.e., angular disparity) between vectors (say "dog" vs. "cat") is a measure of how often responses occur under similar contexts (dog and cat may often co-occur with tokens such as "pet" and "owner"), and the degree to which they do so is a proxy for their semantic relatedness.

When the algorithm performs a comparison for each response word pair, it produces a similarity matrix **M** where each cell M_{ij} contains a comparison with a unique response i (e.g., "paperweight") and a unique response j ("display"). Of note, when $i = j$ (so when "cat" and "cat" are compared) then $M_{ij} = 1$ as two identical responses have a similarity of 1. If we invert the similarity matrix ($1-\mathbf{M}$), we get a *distance matrix* **M'** where each comparison now gives a semantic distance of 0 when they are perfectly similar and 1 when they are fully dissimilar. Consequently, the diagonal of **M'** is now always zero, and we have a distance matrix. Distance matrices are very useful as they can be represented in approximative way using multidimensional scaling[4] as a fully connected undirected weighted network where each node/vertex (response i) is connected with each other node (response j). The length of the connections/edges represents how semantically *dissimilar* the responses are (thus node distance represents semantic dissimilarity). See Figure 9 which shows the semantic network. We used `R` package `igraph` for network plotting.

One of the graph measures that we derived from this is the distance of the response relative to all other responses as a network centrality measure indicating semantic novelty for higher score. The measure is computed by taking the average of distances for each response relative to the other responses.

It is possible that responses named in the haptic versus visual condition are more likely to occupy more and less central positions in the network, which would indicate different affordance perception. Indeed, a linear mixed-effects model including Modality and Object as predictors found a significant effect of Modality ($B = 0.02$, $SE = 0.01$, $p = .04$, effect size[5] $d = 0.25$) on semantic novelty, specifically that visual perception generated more novel responses than haptic perception.

Cluster Analysis

The aforementioned distance matrix was also used for a hierarchical cluster analysis (with a "ward.D2" algorithm implemented by `R` package `factoextra`, Kassambara & Mundt, 2017). The hierarchical cluster analysis uses an iterative process where each closest pair in similarity space is connected in trees. Then, those trees are compared with other trees in terms of closeness in similarity space, until trees reach a single root. The two largest and most distant tree branches were used as categories in subsequent logistic regression to test whether the two modalities matched this clustering. The tree "dendogram" and the network representations of cutting the tree at the last two branches at the top of the cluster hierarchy are shown in Figure 10. With each response now being clustered as belonging to Tree 1 or 2 (depending on the cluster analysis), we can assess whether visual or haptic condition was more likely to produce responses in one semantic cluster over another. The cluster analysis was based on the semantic similarity distance matrix in order to determine whether perceivers use different categories of labels in the visual and haptic condition. This postulation was supported by the finding of a significant effect of Modality ($B = 0.92$, $SE = 0.38$, $p = .023$, odds ratio $= 2.509$) on cluster Categories 1 and 2 based on a hierarchical mixed-effects logistic regression with Modality and Object as predictors. This result indicated that the smaller of the two clusters was connected to the visual modality and contained a larger proportion of nouns than the larger cluster. Thus, the results of cluster analysis corroborated our preceding statistical analyses.

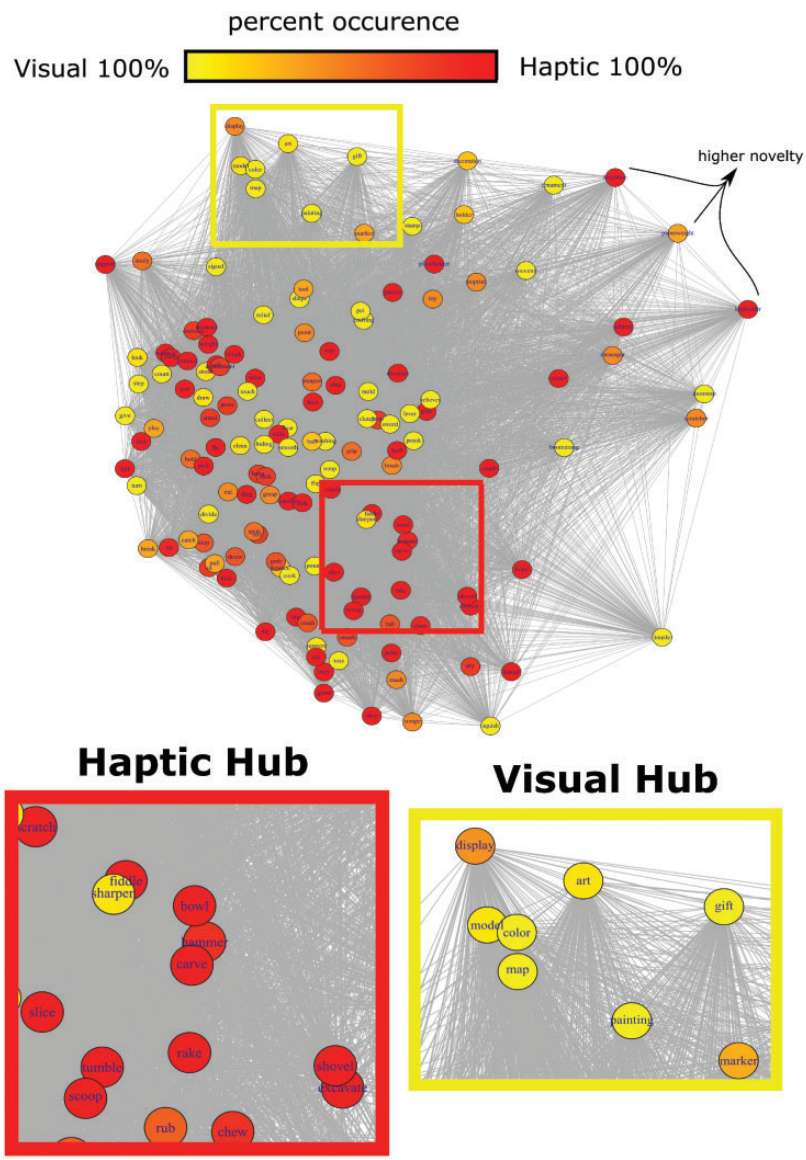

Figure 9. The Semantic Network Is Shown Where Each Connection Length Between Nodes (Responses) Indicates the Degree of Semantic Dissimilarity Between Those Nodes.
Response nodes occupying the same region indicate that they were estimated as more semantically similar by the word2vec procedures. We further colored the nodes for the relative frequency with which they were named in the haptic (more red) versus visual (more yellow) condition. Some small hubs emerged which primarily occurred in one versus the other condition, however a more formal cluster analysis was performed to test whether visual and haptic condition clustered together reliably. Some nodes with high semantic novelty are indicated in the top right corner of the network, but any node that is more peripheral in network has higher semantic novelty scores.
Note. Please refer to the online version of the article to view the figure in colour.

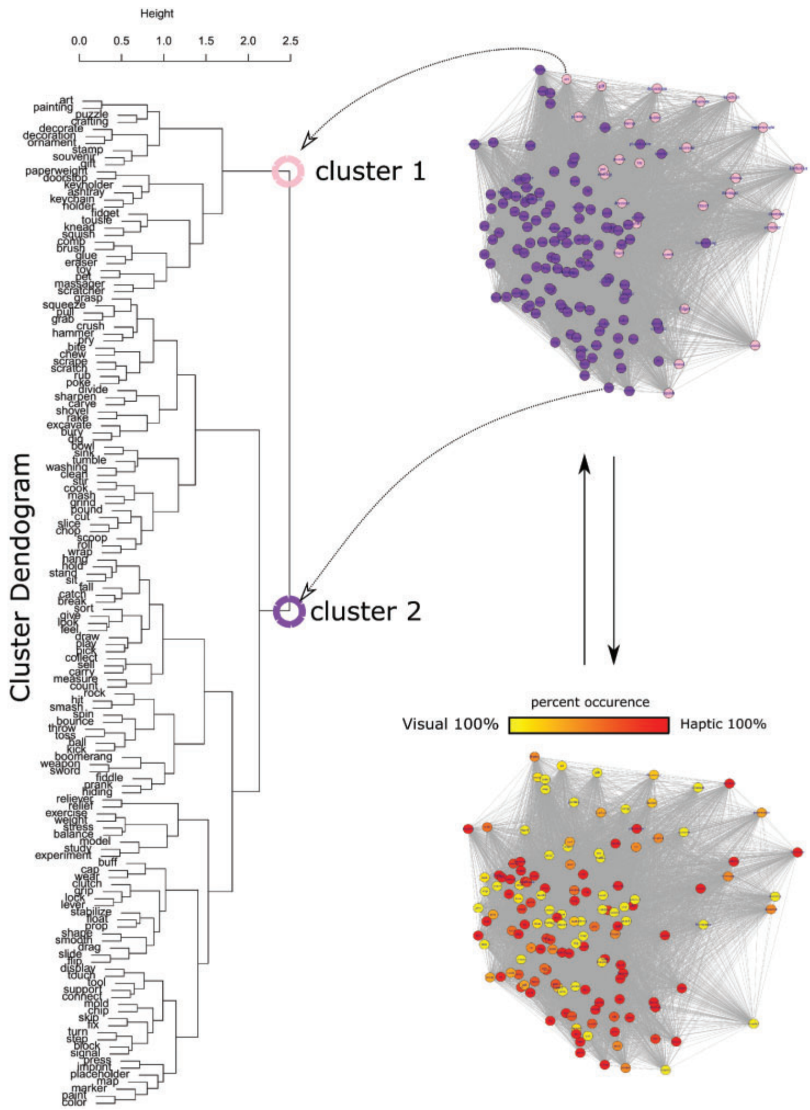

Figure 10. On the Left, the Cluster Dendogram Is Shown.
The "height" axis indicates how much dissimilarity there is at that level of the tree comparisons. For example, the top two responses "art" and "painting" are very close in similarity space (height score is almost 0), while "puzzle" and "crafting" are more dissimilar (with height score of 1). We forcefully cut the tree at the two roots, so that we have two clusters indicated in purple and pink in the network. We do not know yet whether these identified clusters in the meaning space reflect differences in modality (visual vs. haptic), or may be clustering for other reasons. Therefore, we need to assess whether the two clusters map onto the degree of mentions in the visual (yellow shadings) and haptic modality (red shadings). Noncoincidentally those estimated cluster assignments covaried with modality assignments, for example, purple clustered affordances are more often mentioned in the haptic condition (are more often red shaded) and pink clustered affordances are more often mentioned in the visual condition (yellow shaded). Thus, semantic structure of affordance responses in the visual versus haptic condition seemed different, as it is more likely that a condition occupies a certain cluster as defined by hierarchical cluster analysis.
Note. Please refer to the online version of the article to view the figure in colour.

Discussion

The aim of this study was to investigate perceived affordances of unfamiliar human-made objects without designed functions when such objects were explored haptically and visually. The use of both visual and haptic exploration allowed us to investigate how affordance perception is influenced by information in different energetic media obtained by different perceptual systems. Gibson (1962) argued that certain properties of objects (e.g., slant, curvature, edges and corners), and later that properties of objects and events in general (Gibson, 1966, 1979) lawfully structure energy media such that this structure is informative about such properties. Moreover, such properties structure different energy media in analogous ways, allowing for potentially equivalent perception of such properties across modes of exploration.

We first analyzed the duration and quantity of responses. There were no differences in the amount of overlap of responses for each object over trials in the visual and haptic conditions or in the response time between conditions. Participants provided significantly more uses for each object in the haptic condition and a higher proportion of shared responses in the visual condition, but both of these effects depended on block, suggesting that these differences were not global but rather context dependent (Table 2; Figures 5 and 6).

We also analyzed the word class distributions and the semantic content of the responses. Together, the set of analyses showed that participants used more concrete and action-based descriptions (verbs) in the haptic condition and more abstract and object-based words (nouns) in the vision condition, and participants generated more novel responses in the vision condition than in the haptic condition (Figures 7 and 8). Furthermore, the affordances more often mentioned in the visual or haptic condition tended to occupy different regions in the semantic space, as indicated by our cluster analysis. Overall, the results support the hypothesis that multiple affordances would be identified for each object by each modality and that affordances identified by means of each modality would be organized into significantly distinct but overlapping categories.

The results of the experiment reported here showed that although there is substantial overlap, participants list different possible uses in the visual and haptic modality conditions for "feelies" — human-made objects that were not designed with any specific function. When coupled with previous research by Norman et al. (2012) showing individual differences among participants in whether vision or haptics enabled better shape discrimination of feelies, our findings could be interpreted as evidence against the amodal nature of information, undercutting the notion that the same affordance can be perceived through different transformations of the requisite energy arrays.

Despite this possibility, there are several reasons why the differences between modalities found in this study may not necessarily provide evidence against invariance of perceptual information. First, the present task was not well-defined. In everyday life, perception and actualization of affordances are always nested in the context of both goals and task constraints (see Wagman, Caputo, et al., 2016a, 2016b). Perception of affordances of a puddle of water, for example, is constrained by the (or a) superseding task goal (e.g., cross from one side to the other vs. play in the puddle) together with task constraints (e.g., small vs. large puddle). Importantly, previous research has shown that perception of affordances is influenced by both such factors. For example, perception of whether an object can be reached depends on both why and how the reaching task will be performed (Wagman, Cialdella & Stoffregen, 2019). Moreover, research has also shown that perception of affordances for a given behavior more closely reflects the action capabilities for that behavior when that affordance is nested within the context of a superseding goal than when it is not (Doyon et al., 2015; Heft, 1993; Wagman, Bai, et al., 2016). There were no such superseding goals in the present experiment. In fact, the task was quite abstract. That is, rather than perceiving affordances of an object given a goal and task constraints, the participants were essentially tasked with perceiving the goals and task constraints that might apply for a given an object. The abstractness of the responses generated by participants (e.g., "display" and "gift") especially in the visual condition (e.g., "reliever" and "model") likely reflects the abstractness of the task. In this sense, the lesson of this study seems to be that perception of affordances of objects, like perception of affordances of surface layout, is less well-constrained in the absence of such goals and task constraints (Doyon et al., 2015; Heft, 1993; Wagman, Bai, et al., 2016). Intention constrains perception, and perception has an intentional character (Turvey, 2019). Perception without intention potentially devolves into sensation that is based on psychophysical properties.

Second, as described earlier, Gibson (1966, 1979) proposed that a given object or event structures energy arrays such that this structure is informative about that object or event. Moreover, structure in one energy array is lawfully related to structure in another energy array. Consequently, structure in any one of these energy arrays may be sufficient to provide information about a given affordance. Importantly, however, this does not necessarily mean that a person will necessarily have equivalent abilities to perceive a given affordance (or affordances of a given object) by means of different perceptual modalities. A lawful relationship between structure in one energy array and structure in a different energy array guarantees that the patterns exhibited in each array are *analogous* but does not guarantee that such patterns are *identical* (Wagman & Abney, 2012). Therefore, perception of affordances (or affordances of a given object) ought to be analogous (though not necessarily identical) across perceptual modalities.

Moreover, practice (especially practice that includes feedback) leads to changes in what variable is used to perceive a given property (i.e., attunement) and how it is used to do so (i.e., calibration). Differences in how well-practiced a person is in perceiving a given affordance (or affordances of a given object) by means of different perceptual modalities may lead to differences in how well attuned and calibrated they are to structure in each of those energy arrays that is informative about such affordances. Therefore, the differences across modes of exploration observed in the experiment reported here may have resulted from the differences in how readily analogous structure in different energy arrays was detected by participants (and not necessarily from the absence of such analogous structure across energy arrays). Relatedly, the fact that block and modality interactively influenced both the number of reported affordances for each object and the proportion of shared reported affordances suggests that perceived affordances are changing with practice performing the task, perhaps due to changes in attunement or calibration. One possibility for future research is to use a transfer of recalibration paradigm to investigate the degree to which attunement or calibration to (or exploration of) structure in a given energy array transfers from one modality to another (Stephen & Hajnal, 2011; Wagman & Abney, 2012).

Third, in both the visual and haptic exploration conditions, the perceiver actively explored the object by engaging in exploratory behaviors. However, in the visual condition, the objects themselves remained stationary. This is in contrast to the haptic condition in which the objects themselves were manipulated by the perceiver. The stationarity of the objects in the visual condition may have hampered perception of certain dynamic properties of the objects such as mass, mass distribution, and moveableness. Dynamic properties of a given object can (and do) structure the optic array in ways that provide information about such properties. However, this generally requires that the object(s) themselves move (or are moved). Motions are lawfully related to the underlying dynamics that bring about such motion. Therefore, the kinematic patterns created by moving objects are (or can be) informative about the (properties of the) objects themselves, a principle known as the kinematic specification of dynamics (Runeson, 1977; Runeson & Frykholm, 1981, 1983; Runeson et al., 2000; Streit, Shockley, & Riley, 2007; Streit, Shockley, Riley, et al., 2007).

The fact that objects were stationary in the vision condition means such patterns were unavailable to participants in this condition. The fact that objects were occluded in the haptic condition means that such kinematic patterns were also unavailable to this condition. However, participants in the haptic condition had access to the dynamics of object motion as they hefted and wielded each object. Therefore, participants in the vision condition may have been at a relative disadvantage in the availability of information about dynamic object properties. This may be one reason why participants in this condition generated more nouns and fewer verbs than in the haptic condition. Despite this difference, there were more shared responses across modalities than within the visual modality. This suggests that the opportunity to create transformations that reveal invariant stimulation patterns may be more important than whether those transformations bring about dynamic changes, kinematic changes, both, or neither. One possibility for future research is to develop conditions that are more conducive to the kinematic specification of dynamic properties.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Catherine Dowell https://orcid.org/0000-0003-4501-9197

Notes

1. Upon discovering that occasionally participants either forgot to respond that they were finished and needed to be prompted or that there was a considerable lapse between the actual final response and the response of "done," it was decided that the end of responding (after the final response for a trial was given) would also be marked during audio transcription and used as the end times for trials in lieu of pen clicks.
2. Note for the part of speech (POS) tagging it did not matter if we used large or small corpus.
3. Lemmatization is the process of converting words into their root format that one would typically find in a dictionary entry. For example, the root format of "displaying," and "displayed" is "display".
4. The distance matrix is only fully representable in an N-dimensional space, but with visualizing the network one can reduce the dimensionality by approximating a two-dimensional topology through multidimensional scaling.
5. Effect size calculations were based on the formula analogous to Cohen's *d* (Westfall et al., 2014) recommended for mixed-effects models.

References

Benoit, K., Matsuo, A., & Benoit, M. K. (2018). R Package:'spacyr.' https://mran.microsoft.com/snapshot/2018-03-30/web/packages/spacyr/spacyr.pdf

Bingham, G. P., Schmidt, R. C., & Rosenblum, L. D. (1989). Hefting for a maximum distance throw: A smart perceptual mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 507–528. https://doi.org/10.1037/0096-1523.15.3.507

Carello, C. & Turvey, M. T. (2000) Rotational dynamics and dynamic touch. In: M. Heller (ed.), *Touch, representation and blindness*, pp. 26–65. Oxford University Press.

Caviness, J. A. (1962). *The equivalence of visual and tactual stimulation for solid shape perception* [master's thesis]. Cornell University.

Coren, S. (1993). The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: Norms for young adults. *Bulletin of the Psychonomic Society*, *31*, 1–3. https://doi.org/10.3758/bf03334122

Davidson, P. W., Abbott, S., & Gershenfeld, J. (1974). Influence of exploration time on haptic and visual matching of complex shape. *Perception & Psychophysics*, *15*, 539–543. https://doi.org/10.3758/bf03199299

Dennett, D. C. (1987). *The intentional stance*. The MIT Press.

Doyon, J., Hajnal, A., Wagman, J., McGathy, M., Clark, J., & Palatinus, Z. (2015). Are we overthinking it? Haptic perception of geographic slant is accurate when embedded within a secondary task. *Journal of Vision*, *15*, 980–980. https://doi.org/10.1167/15.12.980

Fairhurst, M. T., Travers, E., Hayward, V., & Deroy, O. (2018). Confidence is higher in touch than in vision in cases of perceptual ambiguity. *Scientific Reports*, *8*, 9. https://doi.org/10.1038/s41598-018-34052-z

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, *69*, 477–491. https://doi.org/10.1037/h0046962

Gibson, J. J. (1963). The useful dimensions of sensitivity. *American Psychologist*, *18*, 1–15. https://doi.org/10.1037/h0046033

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

Gibson, J. J. (1979). *The ecological approach to visual perception*. Psychology Press.

Gibson, J. J. (2002). A theory of direct visual perception. In A. Noë & E. Thompson (Eds.), *Vision and mind: Selected readings in the philosophy of perception* (pp. 77–91). MIT Press.

Goodnow, J. J. (1971). Eye and hand: Differential memory and its effect on matching. *Neuropsychologia*, *9*, 89–95. https://doi.org/10.1016/0028-3932(71)90066-2

Gordon, M. S., & Rosenblum, L. D. (2004). Perception of sound-obstructing surfaces using body-scaled judgments. *Ecological Psychology*, *16*, 87–113. https://doi.org/10.1207/s15326969eco1602_1

Hajnal, A., Olavarria, C. X., Surber, T., Clark, J. D., & Doyon, J. K. (2020). Comparison of two psychophysical methods across visual and haptic perception of stand-on-ability. *Psychological Research, 84*, 602–610. https://doi.org/10.1007/s00426-018-1076-6

Hajnal, A., Wagman, J. B., Doyon, J. K., & Clark, J. D. (2016). Perception of stand-on-ability: Do geographical slants feel steeper than they look? *Perception*, *45*, 768–786. https://doi.org/10.1177/0301006616638597

Harrison, S. J., Hajnal, A., Lopresti-Goodman, S., Isenhower, R. W., & Kinsella-Shaw, J. M. (2011). Perceiving action-relevant properties of tools through dynamic touch: Effects of mass distribution, exploration style, and intention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 193–206. https://doi.org/10.1037/a0020407

Heft, H. (1993). A methodological note on overestimates of reaching distance: Distinguishing between perceptual and analytical judgments. *Ecological Psychology*, *5*, 255–271. https://doi.org/10.1207/s15326969eco0503_3

Jaeger, T. F., Graff, P., Croft, W., & Pontillo, D. (2011). Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguistic Typology*, *15*, 281–319. https://doi.org/10.1515/lity.2011.021

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, *40*, 1706–1714. https://doi.org/10.1016/s0028-3932(02)00017-9

Kassambara, A., & Mundt, F. (2017). *Factoextra: extract and visualize the results of multivariate data analyses*. R package version 1.0.7. https://CRAN.R-project.org/package = factoextra

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, *19*, 342–368. https://doi.org/10.1016/0010-0285(87)90008-9

Malek, E. A., & Wagman, J. B. (2008). Kinetic potential influences visual and remote haptic perception of affordances for standing on an inclined surface. *The Quarterly Journal of Experimental Psychology*, *61*, 1813–1826. https://doi.org/10.1080/17470210701712978

Mantel, B., Stoffregen, T. A., Campbell, A., & Bardy, B. G. (2015). Exploratory movement generates higher-order information that is sufficient for accurate perception of scaled egocentric distance. *PLoS One*, *10*, 1–26. https://doi.org/10.1371/journal.pone.0120025

Michaels, C. F., & Oudejans, R. R. D. (1992). The optics and actions of catching fly balls: Zeroing out optical acceleration. *Ecological Psychology*, *4*, 199–222. https://doi.org/10.1207/s15326969eco0404_1

Michaels, C. F., Weier, Z., & Harrison, S. J. (2007). Using vision and dynamic touch to perceive the affordances of tools. *Perception*, *36*, 750–772. https://doi.org/10.1068/p5593

Norman, J. F., Phillips, F., Holmin, J. S., Norman, H. F., Beers, A. M., Boswell, A. M., Cheeseman, J. R., Stethen, A. G., & Ronning, C. (2012). Solid shape discrimination from vision and haptics: Natural objects (*Capsicum annuum*) and Gibson's "feelies. *Experimental Brain Research*, *222*, 321–332. https://doi.org/10.1007/s00221-012-3220-7

Phillips, F., & Egan, E. (2016). *Gibson feelies* (Version 1.1) [3D Object Files]. http://www.skidmore.edu/~flip

Rachwani, J., Tamis-LeMonda, C. S., Lockman, J. J., Karasik, L. B., & Adolph, K. E. (2020). Learning the designed actions of everyday objects. *Journal of Experimental Psychology: General*, *149*, 67–78. https://doi.org/10.1037/xge0000631

Runeson, S. (1977). On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology*, *18*, 172–179. https://doi.org/10.1111/j.1467-9450.1977.tb00274.x

Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 733–740. https://doi.org/10.1037/0096-1523.7.4.733

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*, 585–615. https://doi.org/10.1037/0096-3445.112.4.585

Runeson, S., Juslin, P., & Olsson, H. (2000). Visual perception of dynamic properties: Cue heuristics versus direct-perceptual competence. *Psychological Review*, *107*, 525–555. https://doi.org/10.1037/0033-295x.107.3.525

Srinivasa-Desikan, B. (2018). *Natural language processing and computational linguistics: A practical guide to text analysis with Python, Gensim, spaCy, and Keras*. Packt Publishing Ltd.

Stephen, D. G., & Hajnal, A. (2011). Transfer of calibration between hand and foot: Functional equivalence and fractal fluctuations. *Attention, Perception, & Psychophysics*, *73*, 1302–1328. https://doi.org/10.3758/s13414-011-0142-6

Stoffregen, T. A., Mantel, B., & Bardy, B. G. (2017). The senses considered as one perceptual system. *Ecological Psychology*, *29*, 165–197. https://doi.org/10.1080/10407413.2017.1331116

Streit, M., Shockley, K., & Riley, M. A. (2007). Rotational inertia and multimodal heaviness perception. *Psychonomic Bulletin & Review*, *14*, 1001–1006. https://doi.org/10.3758/bf03194135

Streit, M., Shockley, K., Riley, M. A., & Morris, A. W. (2007). Rotational kinematics influence multimodal perception of heaviness. *Psychonomic Bulletin & Review*, *14*, 363–367. https://doi.org/10.3758/bf03194078

Turvey, M. T. (2019). *Lectures on perception: An ecological perspective*. Routledge.

Wagman, J. B., & Abney, D. H. (2012). Transfer of recalibration from audition to touch: Modality independence as a special case of anatomical independence. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 589–602. https://doi.org/10.1037/a0025427

Wagman, J. B., Bai, J., & Smith, P. J. K. (2016). Nesting in perception of affordances for stepping and leaping. *Attention, Perception, & Psychophysics*, *78*, 1771–1780. https://doi.org/10.3758/s13414-016-1141-4

Wagman, J. B., Caputo, S. E., & Stoffregen, T. A. (2016a). Hierarchical nesting of affordances in a tool use task. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1627–1642. https://doi.org/10.1037/xhp0000251

Wagman, J. B., Caputo, S. E., & Stoffregen, T. A. (2016b). Sensitivity to hierarchical relations among affordances in the assembly of asymmetric tools. *Experimental Brain Research*, *234*, 2923–2933. https://doi.org/10.1007/s00221-016-4695-4

Wagman, J. B., & Carello, C. (2001). Affordances and inertial constraints on tool use. *Ecological Psychology*, *13*, 173–195. https://doi.org/10.1207/s15326969eco1303_1

Wagman, J. B., Cialdella, V. T., & Stoffregen, T. A. (2019). Higher order affordances for reaching: Perception and performance. *Quarterly Journal of Experimental Psychology, 72*, 1200–1211. https://doi.org/10.1177/1747021818784403

Warren, W. H., Jr., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 371–383. https://doi.org/10.1037/0096-1523.13.3.371

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*, 2020–2045. https://doi.org/10.1037/xge0000014

Ye, L., Cardwell, W., & Mark, L. S. (2009). Perceiving multiple affordances for objects. *Ecological Psychology*, *21*, 185–217. https://doi.org/10.1080/10407410903058229

Yu, Y., Bardy, B. G., & Stoffregen, T. A. (2010). Influences of head and torso movement before and during affordance perception. *Journal of Motor Behavior*, *43*, 45–54. https://doi.org/10.1080/00222895.2010.533213