UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Visual perception supports 4-place event representations: A case study of TRADING

Permalink
https://escholarship.org/uc/item/8gf225m1

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Khlystova, Ekaterina
Williams, Alexander
Lidz, Jeffrey
et al.

Publication Date
2024

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Visual perception supports 4-place event representations: A case study of TRADING

Ekaterina A. Khlystova (ekhlystova@ucla.edu)
Department of Linguistics, UCLA
Los Angeles, CA 90095

Alexander Williams (alxndrw@umd.edu)
Departments of Linguistics & Philosophy, U. of Maryland
College Park, MD 20742

Jeffrey Lidz (jlidz@umd.edu)
Department of Linguistics, U. of Maryland
College Park, MD 20742

Laurel Perkins (perkinsl@ucla.edu)
Department of Linguistics, UCLA
Los Angeles, CA 90095

Abstract

Events of social exchange, such as givings and tradings, are uniquely prevalent in human societies and cognitively privileged even at early stages of development. Such events may be represented as having 3 or even 4 participants. To do so in visual working memory would be at the limit of the system, which throughout development can track only 3 to 4 items. Using a case study of trading, we ask (i) whether adults can track all four participants in a trading scene, and (ii) whether they do so by chunking the scene into two giving events, each with 3 participants, to avoid placing the visual working memory system at its limit. We find that adults represent this scene under a 4-participant concept, and do not view the trade as two sequential giving events. We discuss further implications for event perception and verb learning in development.

Keywords: event perception; participant relations; visual working memory; giving events; trading events

Introduction

Events of social exchange, such as givings and tradings, are interesting from both a cognitive and anthropological perspective. Humans may be unique among primates in the prevalance of givings and tradings (Brosnan & De Waal, 2002; De Waal, 1989; Enloe, 2003; Feistner & McGrew, 1989; Gurven, 2004). Giving events are perceived with special status early in human development, with infants as young as 12 months distinguishing givings from takings in terms of the roles of the actors (Schöppner, Sodian, & Pauen, 2006), the minimal number of participants expected (Tatone, Geraci, & Csibra, 2015), and their status as object-mediated social interactions (Tatone et al., 2015; Tatone, 2017). These findings raise questions about how these rich event concepts are yielded by human perceptual systems over the course of development.

In adults, recognition of event relations occurs automatically and rapidly in visual perception, indicating that the human perceptual system is tuned to extract event relations from brief visual exposure (Hafri, Papafragou, & Trueswell, 2013; Hafri, Trueswell, & Strickland, 2018). This rapid extraction is likely facilitated by the object file system, a visual working memory system that maintains attentional indices on individuals (Feigenson, Carey, & Hauser, 2002; Feigenson & Carey, 2003; Simon, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999; Yu, Li, Zhu, Tian, & Lau, 2023). This system has a limit of 3 items in infants and 3 to 4 items in adults (Sperling, 1960; Trick & Pylyshyn, 1993; Scholl & Pylyshyn, 1999; Cowan, 2001; Halberda, Simons, & Wetherhold, 2006; Feigenson et al., 2002; Feigenson & Halberda, 2004, 2008), but its capacity can be increased by grouping items into hierarchically-organized "chunks" (Miller, 1956; Chase & Simon, 1973; Gobet et al., 2001; Moher, Tuerk, & Feigenson, 2012; Feigenson & Halberda, 2004; Feigenson & Carey, 2005).

Plausibly, events of giving and trading are often represented as having a number of participants (3 or 4) that may tax the capacities of visual working memory, both in adults and early in development, unless additional strategies are deployed. Here, we focus on a case study of trading, a canonical 4-place event, in order to ask whether the visual working memory system can readily yield an event percept with 4 participants. We then ask whether the limits of this system may be circumvented by a perceptual strategy that chunks a trading scene into two sequential events: namely, two givings, with only 3 participants each. We find that (i) adults view a trading scene under a 4-participant representation, but (ii) they view this scene under a single event percept rather than as two sequential giving events. These findings invite further questions about the representation of these events in development, with implications for how young children acquire verbs of giving and trading.

Present Study

We conducted four experiments to investigate high-adicity event perception through the case study of tradings. In Experiment 1, we investigated the conceptual structure under which adults view a scene depicting a trade, and find that they explicitly represent all four event participants. We then asked whether they do so by chunking the scene into two giving events, with three participants each. Experiments 2 through 4 answer this question by pitting our trading scene against a control stimulus which may also plausibly be viewed under a 4-place concept. We find that adults also represent all four participants in this control stimulus, but through a different event structure: this control scene is perceived as two separate sequential events, but the trading scene is not.

Experiment 1

Our first experiment diagnosed the adicity of adults' event percepts through a similarity-rating task that adapted a method inspired by Gordon (2003) and introduced in He (2015), Perkins, Knowlton, Williams, and Lidz (under review), and Wellwood, He, Lidz, and Williams (2015). These previous studies asked whether participants would view what was plausibly a change in participant structure – for instance, from an event seen as having 3 participants to one seen as having 2 – as more noteworthy than a change in another physical property of the event, such as manner of motion. All else being held equal, if people

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

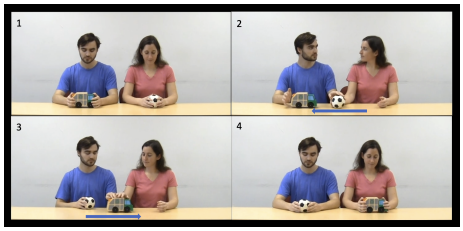

Figure 1: Illustration of 'trading'

treat a change in putative participant structure as more noteworthy than a change to another event property, this may be taken as evidence about the conceptual structure under which they viewed the event. Perkins et al. (under review) tested infants on videos of a girl picking up a truck, with a boy sitting idly by, compared to the girl taking the truck from the boy. They found that infants viewed this change as more noteworthy than a change to the manner in which the truck was moved. Ensuring that no other perceptual differences could explain this pattern of results, this suggests that infants viewed this participant change as an important conceptual difference: infants viewed the 'taking' scene, but not the 'picking-up' scene, under an event concept in which the girl, truck, and boy were all explicitly represented as participants.

Here, we adopted this logic to test adults' representations of a scene in which a boy gives a girl a truck, and the girl gives the boy a ball (Figure 1). We asked adults to compare this video to videos in which one of the actors or items was no longer a participant, or the manner of motion was changed. Under one hypothesis, this stimulus scene might be viewed as a TRADING in which all 4 participants are explicitly represented in the conceptual structure (1):

(1) **TRADING**-AGENT1-AGENT2-ITEM1-ITEM2

If one of the items is removed from the action – for example, the ball is passed back and forth between the actors and the truck remains unmoved – this might now be viewed as a 3-participant PASSING (2). Similarly, if one of the actors looks off to the side while the second actor swaps the two items, this might be viewed as a 3-participant SWAPPING (3):

(2) **PASSING**-AGENT1-AGENT2-ITEM

(3) **SWAPPING**-AGENT-ITEM1-ITEM2

We contrast these changes in participant structure to a change in manner of motion. If the ball and truck are moved not by sliding them on the table, but by lifting them off of the table, this might be seen as a TRADING with a different manner, but not a different number of participants (4).

(4) **LIFTINGTRADE**-AGENT1-AGENT2-ITEM1-ITEM2

If people view the trading but not the passing or swapping scenes under a 4-participant concept, then changing from (1) to (2) or (3) will involve a change in conceptual structure. All else equal, we predict that these 'participant changes' will therefore be viewed as more noteworthy than a manner change, which does not change the adicity of the conceptual representation. In particular, we would expect changes to the hypothesized event participants to be rated as less similar to the original trading scene than a change to the manner of motion.

Methods

Participants 24 adults (12 female; ages 18-63) were recruited via Prolific. Participants were from the United States or the United Kingdom and were paid $6 for participating.

Stimuli We developed a series of video stimuli that manipulate the possible participant structure under which the scenes could be represented. All of the videos contained the same four potential participants: a girl, a boy, a ball, and a truck, all visible throughout the event. The videos always begin with the girl holding the ball and the boy holding the truck, looking down at the respective item, unless otherwise specified. In 'trading' videos, the two actors exchange the two items, one after the other, after making brief eye contact (Figure 1). In 'item change' videos, the event proceeds exactly as in a trading video, except that one of the items is no longer moved. The actors pass either the truck or the ball back and forth, with the second item present but unmoved. In 'person change' videos, the event proceeds exactly as in a trading video except that the items are exchanged by only one of the actors; the second actor looks off to the side and does not participate. Together, the person change and item change videos comprised the participant change manipulations. In 'manner change' videos, the event proceeds exactly as in a trading video, but with the actors lifting the items rather than sliding them across the table.

Three tokens of each event type and manipulation were recorded, with the timing of each motion identical across event types in order to match their perceptual properties. This was achieved by using pre-recorded audio cues and a metronome during filming. Each token was exactly 10 seconds.

Test trials were created by pairing two videos together with one second of black screen between them. Each trial was 21 seconds. Each of the three tokens per event type was paired with all three tokens of the relevant manipulation (for example, Trade Token 1 was paired with Ball-Subtraction Token 1, 2, and 3, etc.). This resulted in 18 total trials for each type of video pair. Video pairs were matched for order of movement, with either the ball moving first in both videos or the truck moving first, but never a mixture of the two. Baseline control stimuli were developed by pairing two 'trading' tokens together; however, no two identical tokens were paired together. This led to a total of 12 control trials. Overall, 120 trials were created for Experiment 1. These trials were broken into two lists, with 66 trials per list: half of the experimental trials (54), and all 12 control trials, counterbalanced for order of presentation and direction of movement.

Procedure The experiment was conducted online through LabVanced (Finger, Goeke, Diekamp, Standvoß, & König, 2017). Participants were told that a video editor had lost

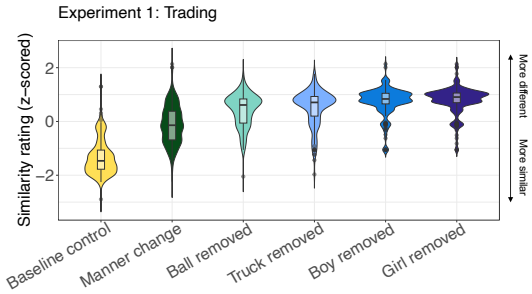

Figure 2: Similarity ratings from Experiment 1

footage for a film, and that they needed to judge how likely the film director would be to notice the change between the first video (the "lost footage") and the second video (a "substitute take"). Videos played automatically and could not be paused or replayed. All trials consisted of a video pair, followed by a 7-point Likert scale with the prompt "How likely is the director to notice the change?", with 1 being "Very unlikely" and 7 being "Very likely." Participants were also asked to rate their confidence on a 4-point scale. Prior to the test trials, participants were given two practice trials, one with a manner change and one without. No feedback was given.

Participants were randomly assigned to one of the two lists, with condition manipulated within participants. Condition type was pseudo-randomized by trial such that no condition was seen more than twice in a row. Test trials were counterbalanced across participants for both the order of token-type presentation (trade-change vs. change-trade) and the order of item movement (ball-first vs. truck-first).

Results

Any test trials whose similarity rating or time to complete the rating task was more than 2 SD from the mean were excluded (108 outliers excluded for a total of 1450 trials analyzed). The confidence ratings revealed no significant effects and thus will not be discussed further. For purposes of visualization, similarity ratings were z-scored by participant and plotted for each condition (Figure 2). Lower z-scores indicate a greater degree of similarity, while higher scores indicate a lower degree of similarity.

We fit a linear mixed effects model with similarity rating as the dependent measure and condition as a fixed effect. The maximal model that converged had a random intercept for subject. Model comparisons confirmed a significant effect of condition ($\chi^2(5) = 1421.8$, $p < 0.0001$). Pairwise comparisons revealed a significant difference between manner change and all participant change conditions ($t(1422) = 12.863$, $p < 0.0001$ for truck subtraction; $t(1422) = 11.513$, $p < 0.0001$ for ball subtraction; $t(1423) = 17.994$, $p < 0.0001$ for boy subtraction; $t(1422) = 17.24$, $p < 0.0001$ for girl subtraction).

Discussion

We find that changes to all four hypothesized participants of our trading scene were rated as more noteworthy than a change to the manner of motion. This result is predicted under the hypothesis that adults perceived the participant-change conditions as involving not merely a change in the physical event properties, but also a change in conceptual structure: namely, a change between a 4-place and a 3-place concept. All else equal, this suggests that adults perceived our trading scene under a concept with all four participants explicitly represented, as in (1).

Adults have a reported visual working memory limit of 3 or 4 (Sperling, 1960; Trick & Pylyshyn, 1993; Scholl & Pylyshyn, 1999; Cowan, 2001; Halberda et al., 2006). Thus, encoding and tracking all 4 participant relations in our trading scene places the visual working memory system at its limit. To alleviate demands on this system, it is possible that people may be representing the scene under two separate event concepts, each with fewer participant relations. This could be interpreted as the implementation of a chunking technique (Miller, 1956; Chase & Simon, 1973; Gobet et al., 2001; Feigenson & Halberda, 2004; Feigenson & Carey, 2005; Moher et al., 2012). For example, rather than being viewed under a single TRADING event concept, our trading scene might be viewed as two sequential GIVINGs :

(5) a. **GIVING1**-AGENT-RECIPIENT-ITEM
b. **GIVING2**-AGENT-RECIPIENT-ITEM

This type of conceptual structure has parallels to previous semantic analyses of TRADE as a two-part predicate composed of a primary GIVING and a symmetrical GIVING (Jackendoff, 1992). We tested this possibility in Experiments 2-4 by comparing the representation of our 'taking' scene to the representation of a different scene – a 'giving-then-disposing' – which is likely to be viewed as two sequential events.

Experiment 2

We tested perception of another potentially 4-participant event: one actor gives an item to another actor, who then disposes of his or her original item by sliding it to the side (Figure 3).

This event type was chosen for two reasons. First, this scene differs minimally from a trading in that there are two actors exchanging two items. As such, it serves as a comparison to determine whether adults are *only* able to track four participants in trading scenes, or if this ability holds across other related scene types. Second, it is plausible that this scene will be viewed under two sequential event concepts – an actor giving an item to the other is followed by the second actor disposing of his or her original item, as in (6). This allows it to serve as a useful comparison to 'trading,' to investigate whether a given scene is viewed under one event percept or two.

(6) a. **GIVING**-AGENT-RECIPIENT-ITEM
b. **DISPOSING**-AGENT-ITEM

In Experiment 2, we first ask whether adults track all four participants in our giving-then-disposing scene. This sets the stage for asking whether these participants are perceived in relation to a single event, or in relation to two sequential events.

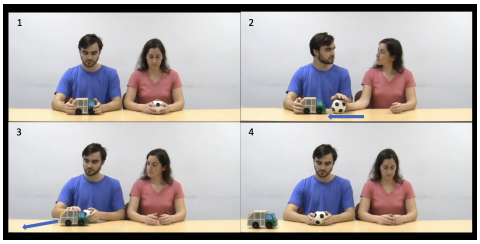

Figure 3: Illustration of 'giving-then-disposing'

We investigate this second question in Experiments 3-4.

Method

Participants 24 adults (12 female; ages 24-69) were recruited via Prolific. Participants were from the United States or the United Kingdom and were paid $6 for participating.

Stimuli To test the number of participants represented in our 'giving-then-disposing' scene, we performed the same manipulation as in Experiment 1, comparing similarity ratings for videos in which a participant was removed and videos in which the manner of motion was changed. The 'item change' videos consisted of one actor giving an item to the other actor, who took possession of that item before sliding it to the side. The second actor's original item remained present but unmoved. The 'manner change' videos consisted of the actors giving and disposing of their items by lifting them off the table.

In Experiment 1, we saw that the changes to the actors' participation were noted very strongly by subjects. Thus, we did not include 'person change' videos in this experiment, under the assumption that the actors would continue to be robustly perceived as event participants. However, to avoid ceiling effects for our crucial manipulations, we included a ceiling control condition: a giving-then-disposing token matched with another token of the same scene which was temporally reversed (i.e., played backwards). Backwards-motion tokens were never paired with their forward counterpart, resulting in 12 total ceiling control trials. Pilot testing confirmed that the temporal reversals were not perceived as unnatural. 12 baseline control trials were created as in Experiment 1.

Recording and trial creation was carried out as described in Experiment 1. Each participant saw a total of 60 trials: 36 experimental, 12 ceiling controls, and 12 baseline controls.

Procedure The procedure was identical to Experiment 1.

Results

The z-scored similarity rations are shown in Figure 4. A linear mixed effects model analysis with similarity rating as the dependent measure, condition as fixed effect, and a random intercept for subject revealed a significant effect of condition ($\chi^2(4) = 1241.2$, $p < 0.001$). Pairwise comparisons confirmed a significant difference between item change and manner change conditions ($t(1229) = 5.755$, $p < 0.0001$ for ball subtraction, and $t(1229) = 6.153$, $p < 0.0001$ for truck subtraction).

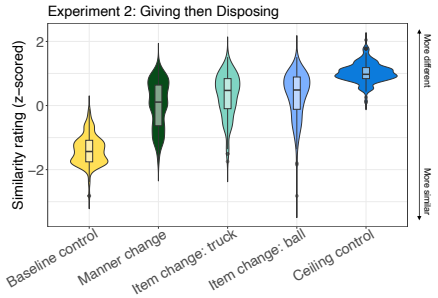

Figure 4: Similarity ratings from Experiment 2

Discussion

The findings in Experiment 2 suggest that adults represent all possible participants in our giving-then-disposing scene. Just as for trading, adults viewed a change in the hypothesized participant structure as more noteworthy than a change in manner of motion, suggesting that the truck and ball filled privileged participant relations in their conceptual representation.

Participants' success in both Experiments 1 and 2 sets us up to test whether they are using a chunking strategy: namely, whether they view the trading scene as two sequential 'giving' events, and view the giving-then-disposing scene as a 'giving' event that prompts a 'disposing' event. We test this possibility in Experiments 3 and 4 by adopting manipulations from the causal perception literature (Leslie, 1982, 1984). In Experiment 3, we introduce a manipulation to disrupt the sequence of the two hypothesized events, by reversing their relative order. In Experiment 4, we introduce a manipulation to disrupt the coherence of a hypothesized single event percept, by inserting a pause in the middle.

Experiment 3

In Experiment 3, we manipulated the order in which the potential sub-events occur in our trading and giving-then-disposing scenes. If each scene type is viewed under two sequential event concepts, then changing the relative order of the actions reverses this sequence, which should result in a noteworthy difference for the perceiver. But if the scene is viewed under a single event concept, then all else being held equal, changing the order of the actions might be less disruptive.

By hypothesis, we expect giving-then-disposing to be viewed under two event concepts: a GIVING followed by a DISPOSING, as in (6). A change in the relative order of the movements, to a DISPOSING followed by a GIVING, should therefore be a noteworthy difference, due to the change in sequential position of the two events. If the trading scene is also viewed under two event concepts – a GIVING followed by a second GIVING, as in (5) – then we would expect a reversal in the order of the two GIVINGs to be similarly noteworthy. This predicts that there should be no interaction of scene type by condition. But if the trading scene is viewed under only one event concept, then we would expect a change to the relative order of movement to be less noteworthy. This predicts an interaction of scene type by condition.

Methods

Participants 48 participants (24 female; ages 21-68) were recruited via Prolific. Participants were from the United States and the United Kingdom and were paid $6 for participating.

Stimuli The novel manipulation in Experiment 3 was an order change. For trading scenes, this was achieved by pairing a trading video in which the girl first gives her ball to the boy, with a trading video in which the boy first gives his truck to the girl. In giving-then-disposing scenes, the order change was achieved by pairing the giving-disposing scene described in Experiment 2 with another type of scene developed for this experiment. In this new scene type, one actor first "disposes" of his or her item by sliding it off to the side, after which the second actor "gives" his or her item to the first actor.

Baseline control and manner change conditions for each scene type were identical to those in Experiments 1 and 2, respectively. To prevent ceiling effects, a person change condition was included for trading scenes (as in Experiment 1), and the backwards-motion condition was included for giving-then-disposing (as in Experiment 2). Although the critical manipulation is the order change, the control and manner change conditions were included to keep the experimental setup as similar as possible to Experiments 1 and 2.

Procedure The procedure was the same as the previous experiments, with the addition of scene type as a between-subjects factor. Half of the participants were assigned to the 'trading' condition and half to the 'giving-then-disposing' condition. The experiment consisted of 60 test trials in the giving-then-disposing condition and 66 test trials in the trading condition: 12 baseline controls, 18 manner changes, 18 order changes, and 18 or 12 ceiling controls for trading and giving-then-disposing, respectively[1].

Results

The z-scored similarity ratings for Experiment 3 are shown in Figure 5. We fit a linear mixed effects model with similarity rating as the dependent measure, fixed effects of condition and scene type, and a random intercept for subject. Model comparisons revealed significant main effects of both condition ($\chi^2(3) = 1707.6$, $p < 0.0001$) and scene type ($\chi^2(1) = 5.287$, $p = 0.0215$), and importantly, a significant interaction of condition by scene type ($\chi^2(3) = 86.736$, $p < 0.0001$). As predicted, order changes were viewed as significantly less noteworthy in the trading scene than in the giving-then-disposing scene ($t(2730) = -8.593$, $p < 0.0001$). As a benchmark comparison, the order change was also viewed as significantly less noteworthy than the manner change for trading ($t(2704) = -11.430$, $p < 0.0001$), but just as noteworthy as the manner change for giving-then-disposing ($t(2704) = -0.363$, $p = 0.9836$).

[1]This difference arises from the fact that we did not pair backwards-motions tokens of giving-then-disposing scenes with their forward counterparts, resulting in only 12 possible pairs.

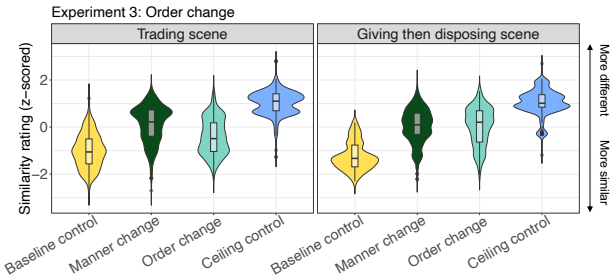

Figure 5: Similarity ratings from Experiment 3

Discussion

The findings of Experiment 3 suggest that our trading and giving-then-disposing scenes were viewed under different types of event representations. A change to the order of movement was rated as significantly more noteworthy for giving-then-disposing than for trading. For giving-then-disposing, order changes were viewed as just as noteworthy as another physical change to the event (the manner of motion). For trading, order changes were rated as significantly less noteworthy than changes to other event properties. It appears that the giving-then-disposing percept was disrupted by a reversal to the order of motion, but the trading percept was not. This is consistent with the hypothesis that the giving-then-disposing scene is perceived under a two-event structure (a GIVING followed by a DISPOSING), whereas the trading scene is perceived as a single coherent TRADING event, and not as two sequential GIVINGs.

Experiment 4

Experiment 4 aimed to marshall further support for our interpretation of Experiment 3. Here, we asked whether each event percept would withstand a disruption to its timing. For both the trading and giving-then-disposing scenes, we manipulated the timing with which the possible sequential events occurred by inserting a pause at the hypothesized event boundary, after the first item was moved.

If a scene is being viewed under a single event concept, the insertion of a pause should be viewed as noteworthy. It should break the coherence of the single event percept, causing the scene to be viewed as two sequential events instead. If, on the other hand, the scene is viewed under two event concepts initially, then inserting a pause will not disrupt the event percept as substantially, and thus should not be as noticeable a change. On the hypothesis that our giving-then-disposing scene is viewed as two sequential events, but our trading scene is viewed as one event, then we again predict a condition by scene type interaction. In this case, we expect the interaction to go in the opposite direction as for Experiment 3: the crucial timing manipulation should be viewed as more noteworthy for 'trading' than for 'giving-then-disposing.'

Methods

Participants 48 participants (24 female; ages 20-66) were recruited via Prolific. Participants were from the United States

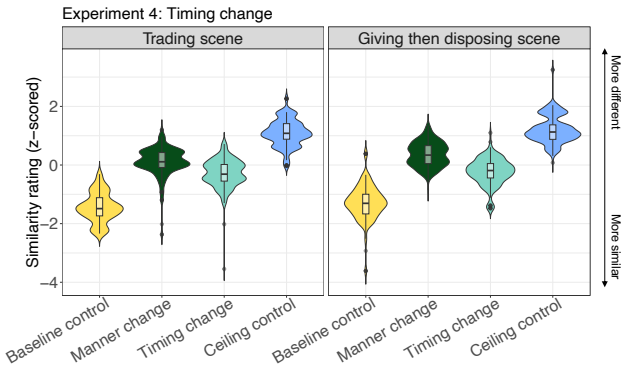

Figure 6: Similarity ratings from Experiment 4

or the United Kingdom and were paid $6 for participating.

Stimuli The video stimuli for Experiment 4 were identical to those for Experiment 3, except that we replaced the order change stimuli with 'timing change' stimuli. These timing change stimuli were created by recording new tokens of the trading and giving-then-disposing scenes, now with a 1-second pause after the first actor gives their item to the second actor. As before, an audio recording with cues for each movement alongside a metronome beat was used to ensure the pause was equally long for all tokens and that movement occurred at the same time points throughout the scene. Token pairs were created as in the previous experiments.

Procedure The procedure was identical to Experiment 3.

Results

The z-scored similarity ratings for Experiment 4 are shown in Figure 6. We fit a linear mixed effects model with with similarity rating as the dependent measure, fixed effects of condition and scene type, and a random intercept for subject. Model comparisons revealed significant main effects of both condition ($\chi^2(3) = 3041.87$, $p < 0.0001$) and scene type ($\chi^2(1) = 46.34$, $p < 0.0001$), and importantly, a significant interaction of condition by scene type ($\chi^2(3) = 2.79$, $p = 0.0325$). As predicted, timing changes were viewed as more noteworthy for the trading scenes than for giving-then-disposing ($t(1751) = 2.140$, $p = 0.0325$). Although timing changes were viewed as less noteworthy than manner changes for both scene types ($t(1735) = -10.619$, $p < 0.0001$ for trading; $t(1736) = -13.222$, $p < 0.0001$ for giving-then-disposing), timing changes were viewed as more similar to manner changes for trading, and less similar to manner changes for giving-then-disposing.

Discussion

Consistent with the results of Experiment 3, we again found a predicted condition by scene type interaction, confirming a difference in how our trading and giving-then-disposing scenes are represented. A pause inserted into the trading scenes was seen as more noteworthy than the same length of pause inserted in same position in the giving-then-disposing scenes. This suggests that the giving-then-disposing percept was not disrupted by the insertion of a pause, as predicted under the hypothesis that this scene was already viewed as two sequential events. However, the trading percept was disrupted by the insertion of a pause, suggesting that it was not originally viewed as two separate events.

General Discussion

This paper examined the conceptual representation of trading as a case study of high-adicity event perception. We found that adults view a scene of trading under a 4-participant structure, in which both traders and both traded items are explicitly represented. Moreover, by comparing against another closely related, plausibly four-participant event, we found converging evidence that the trading scene was viewed under one TRADING event concept, rather than as two sequential GIVINGs. These findings are interesting in light of reported constraints on visual perception: in order to represent all four participants in relation to a single event, the visual working memory system may be operating at its reported limit of 4 (Sperling, 1960; Trick & Pylyshyn, 1993; Scholl & Pylyshyn, 1999; Cowan, 2001; Halberda et al., 2006). While it is possible that our task did not sufficiently tax adults' visual working memory, participants reported that this task was challenging. It was therefore likely informative about perceptual limits. Our findings also suggest that these limits are not being circumvented by chunking the trading scene into sequential events, each with fewer participants. Instead, adult visual perception appears capable of yielding a 4-place event percept without this particular type of internal structure.

This finding has potential linguistic implications for semantic analyses of verbs of TRADING that treat this predicate as composed of two causally-related GIVINGs (Jackendoff, 1992). Our findings suggest that, in nonlinguistic visual perception, our trading scene was not viewed as two sequential giving events. However, this does not preclude the possibility that this scene may be represented with other types of internal structure. For instance, instead of chunking the scene into two sequential events, people may instead chunk the event participants into groups that bear similar relations to their events: for instance, two traders, and two things traded. We leave this possibility as a question for future work.

Our results have further implications for language acquisition. Acquiring a verb like *trade* requires mapping a linguistic form onto a conceptual representation of an event which falls under the TRADING category. As infants have more stringent caps on visual working memory than adults (Feigenson et al., 2002; Feigenson & Carey, 2003, 2005), does the perceptual system of a young verb learner likewise readily yield 4-place representations of trading scenes? If so, we might ask what mechanisms they deploy to circumvent their visual working memory limits. If not, this would raise a puzzle for how verbs like *trade* are acquired. Further work on the perception of tradings in development may help us understand the perceptual support for acquiring high-adicity verb meanings, with implications for why so few 4-place concepts are lexicalized as simple monomorphemic verbs crosslinguistically.

Acknowledgments

We would like to thank Tara Mease, Sterling Mullenix, Katherine Howitt, and Allison Dods for their help with stimuli creation and filming. We would also like to thank audiences in the UCLA Psycholinguistics/ Computational Linguistics Seminar and the University of Maryland Language Acquisition Lab for their feedback at various stages of this work.

References

Brosnan, S. F., & De Waal, F. B. (2002). A proximate perspective on reciprocal altruism. *Human Nature*, *13*, 129–152.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology*, *4*(1), 55–81.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, *24*(1), 87–114.

De Waal, F. B. (1989). Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution*, *18*(5), 433–459.

Enloe, J. (2003). Food sharing past and present: archaeological evidence for economic and social interactions. *Before Farming*, *2003*(1), 1–23.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295–313.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological science*, *13*(2), 150–156.

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition*, *91*(2), 173–190.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences*, *105*(29), 9926–9930.

Feistner, A. T., & McGrew, W. C. (1989). Food-sharing in primates: a critical review. *Perspectives in primate biology*, *3*(21-36).

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). Labvanced: a unified javascript framework for online studies. In *International Conference on Computational Social Science (cologne)* (pp. 1–3).

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences*, *5*(6), 236–243.

Gordon, P. (2003). The origin of argument structure in infant event representations. In *Proceedings of the 28th annual Boston University Conference on Language Development* (Vol. 1, pp. 189–198).

Gurven, M. (2004). To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, *27*(4), 543–560.

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*(3), 880.

Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, *175*, 36–52.

Halberda, J., Simons, D., & Wetherhold, J. (2006). You can never attend to more than three items at once: Gestalt grouping principles explain changes in capacity. *Manuscript submitted for publication*.

He, X. (2015). *Verb learning under guidance* (Unpublished doctoral dissertation). University of Maryland, College Park.

Jackendoff, R. S. (1992). *Semantic structures* (Vol. 18). MIT press.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173–186.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, *13*(3), 287–305.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Moher, M., Tuerk, A. S., & Feigenson, L. (2012). Seven-month-old infants chunk items in memory. *Journal of experimental child psychology*, *112*(4), 361–377.

Perkins, L., Knowlton, T., Williams, A., & Lidz, J. (under review). Thematic content, not number matching, drives syntactic bootstrapping.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive psychology*, *38*(2), 259–290.

Schöppner, B., Sodian, B., & Pauen, S. (2006). Encoding action roles in meaningful social interaction in the first year of life. *Infancy*, *9*(3), 289–311.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non-numerical" account. *Cognitive development*, *12*(3), 349–372.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied*, *74*(11), 1.

Tatone, D. (2017). *The naïve sociology of resource transfer* (Unpublished doctoral dissertation). Central European University Budapest, Hungary.

Tatone, D., Geraci, A., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition*, *137*, 47–62.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 331.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999).

What representations might underlie infant numerical knowledge? *Cognitive development*, *14*(1), 1–36.

Wellwood, A., He, A. X., Lidz, J., & Williams, A. (2015). Participant structure in event perception: Towards the acquisition of implicitly 3-place predicates. *University of Pennsylvania Working Papers in Linguistics*, *21*(1), 32.

Yu, X., Li, J., Zhu, H., Tian, X., & Lau, E. (2023). Electrophysiological hallmarks for event relations and event roles in working memory. *bioRxiv*, 2023–05.