AGENT PREFERENCE IN DEVELOPMENT: THE ROLE OF ANIMACY AND EVENT STRUCTURE

CEREN BARIŞ

ÖZYEĞİN UNIVERSITY

DECEMBER, 2024

AGENT PREFERENCE IN DEVELOPMENT: THE ROLE OF ANIMACY AND EVENT STRUCTURE

by
Ceren Barış

Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Arts

in
Psychology

Advisor: Assoc. Prof. Ercenur Ünal

Graduate School of Social Sciences
Özyeğin University
İstanbul

December, 2024

AGENT PREFERENCE IN DEVELOPMENT: THE ROLE OF ANIMACY AND EVENT STRUCTURE

Approved by:

Assoc. Prof. Ercenur Ünal, Advisor

Department of Psychology

Özyeğin University

Prof. Ali İzzet Tekcan

Department of Psychology

Özyeğin University

Prof. Ayşe Aylin Küntay Copty

Department of Psychology

Koç University

Approval Date: December 26, 2024

DECLARATION OF ORIGINALITY

I hereby declare that I am the sole author of this thesis and that this is the true copy of my thesis, including the final revisions, approved by my thesis committee. All data and information have been obtained, produced, and presented in accordance with the rules of research ethics and principles of academic honesty. As required by these rules, to the best of my knowledge I have acknowledged ideas, thoughts, and any copyrighted material in accordance with the standard referencing rules. I certify that any part of this thesis has not been submitted for a degree or diploma in another educational institution.

Ceren Barış

ABSTRACT

According to linguistic theory, thematic roles in language are hierarchically organized in terms of their salience and this hierarchy is rooted in their counterparts as event participants in cognition. Here, we seek developmental evidence for this proposal in Turkish-speaking 3- to 5-year-olds. We examine whether children show a preference for Agents over Patients in two-participant causative events, such as a woman tickling a man. To examine whether Patient animacy has an effect on this preference, we prepared images depicting events in which the Patient was either a human or an inanimate object. Modified versions of these depictions were created by applying a color change to either the Agent or the Patient. Our stimuli also included back-to-back depictions that disrupted the event structure while keeping other visual features constant. In an eye-tracked change detection task, participants viewed causative event depictions flickering between original and modified images, with color changes applied to either the Agent or the Patient. We measured change detection accuracy and proportion of fixations to the changing role. Our findings revealed an asymmetry between Agents and Patients only in face-to-face events with an inanimate Patient: participants were less likely to detect changes in and they looked more towards the Patient compared to the Agent. Looking more toward an entity might indicate increased cognitive effort to detect the change because the entity is less salient. No differences were observed in face-to-face events with an animate Patient or back-to-back events. These findings suggest an interplay of event roles and animacy in Agent preference and support the generalizability of the phenomenon.

Keywords: Event cognition, agent preference, animacy, eye tracking

ÖZET

Dilbilim teorileri, dildeki tematik rollerin hiyerarşik olarak düzenlendiğini ve bunun bilişteki olay katılımcılarına dayandığını öne sürer. Bu çalışmada, bu görüşü gelişimsel olarak incelemek için Türkçe konuşan 3 ila 5 yaşındaki çocuklar incelendi. Çocukların, bir kadının bir adamı gıdıklaması gibi iki katılımcılı nedensel olaylarda, eylemden etkilenen katılımcıdan ziyade eylemi gerçekleştiren katılımcıya yönelik bir tercih gösterip göstermedikleri incelendi. Bu tercihin, eylemden etkilenen katılımcının canlı veya cansız olmasından etkilenip etkilenmediğini araştırmak amacıyla, eylemden etkilenen katılımcının insan olduğu ve cansız bir nesne olduğu olay görselleri hazırlandı. Görsellerin eylemi gerçekleştiren ya da eylemden etkilenen katılımcıya renk değişimi uygulanarak değiştirilmiş tasvirleri oluşturuldu. Görsel özellikler sabit tutularak olay bütünlüğünü ortadan kaldırmak için aynı olayların katılımcıların sırt sırta oldukları tasvirleri kullanıldı. Deney katılımcılarına göz hareketlerinin kaydedildiği bir değişim algılama görevi kapsamında orijinal ve renk değişimi olan olay tasvirlerinin art arda yanıp sönmesi izletildi. Katılımcıların değişim algılama doğruluğu ve değişen olay katılımcılarına yönelik bakış oranları ölçüldü. Bulgular, olaydan etkilenen katılımcının cansız bir nesne olduğu yüz yüze olaylarda, deney katılımcılarının, eylemden etkilenen katılımcıdaki değişiklikleri daha az doğru tespit ettiklerini ve eylemden etkilenen katılımcıya daha fazla baktıklarını gösterdi. Eylemden etkilenen katılımcının canlı olduğu yüz yüze veya sırt sırta olaylarda ise herhangi bir fark gözlemlenmedi. Bulgular, olay rollerinin ve eylemden etkilenen katılımcının canlılığının eylemi gerçekleştiren katılımcı tercihindeki etkileşimini önermekte ve tercihin genelleştirilebilirliğini desteklemektedir.

Anahtar Kelimeler: Olay bilişi, eylemi gerçekleştiren tercihi, canlılık, göz izleme

ACKNOWLEDGMENTS

I would like to begin by expressing my deepest gratitude to Dr. Ercenur Ünal, an exceptional mentor whose guidance has profoundly shaped my academic and scientific training. Her unwavering support, encouragement, and insight since the day we began working together have been invaluable, leaving a lasting impact that I will carry throughout my career. I am also sincerely grateful to Prof. Dr. Aylin Küntay and Prof. Dr. Ali Tekcan for taking part in my committee and providing useful feedback that enhanced my work.

This research would not have been possible without the financial support of TÜBİTAK, the Scientific and Technological Research Council of Turkey, through the 2210-A Master's Fellowship. The study itself was supported by a TÜBİTAK 1001 grant (grant number 121K259), awarded to Dr. Ercenur Ünal. I am thankful for their contributions to my academic journey.

I extend my heartfelt thanks to the members of the Language, Cognition, and Development Lab for their support throughout this project. Special thanks to Dilara for her creative work in stimulus preparation, and to Deniz, Elif, Hazan, and Lena for their invaluable assistance with data collection. I am particularly grateful to Bahar and Bilge for their unwavering support, their patience in answering my questions, and for being the best lab mates I could hope for.

I would like to thank every child who participated in this study for bringing joy to this project as none of this would be possible without them.

I am forever thankful for my incredible friends Ata, Elif, Emre, Süeda, Zülal, and Zeynep. Their encouragement and motivating words were a source of strength during the challenging times of preparing this thesis. Their insights and unwavering support have been invaluable, and they remain my chosen family for life.

Finally, I owe my deepest gratitude to my family for their unwavering love and encouragement. To my mom, who sacrificed so much and provided me with everything I needed to pursue my education—I cannot thank you enough. To my dad and Demet Abla, thank you for believing in my goals and dreams. To Boran and Yankı, thank you for lifting my spirits every day and for being the best brothers anyone could ask for.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

We perceive the world in a continuous manner but divide our experience into meaningful units with spatial and temporal boundaries called ‘events’ (Zacks & Tversky, 2001). Events are made up of a “set of participants, activities, and outcomes that are bound together by causal interrelatedness” (Elman, 2009, p. 572). In linguistics, event participants are studied under *thematic roles* (Dowty, 1991; Grimshaw, 1981; Jackendoff, 1983; 1990; Pinker, 1989; Rissman & Majid, 2019) and these roles are argued to have a hierarchical organization in language (*Thematic Hierarchy*; Grimshaw, 1981; Pinker, 1989). Within this view, the causer of an action (the Agent) is more salient than the role affected by the action (the Patient; Baker, 1997; Jackendoff, 1990). Since thematic roles are assumed to be based on event participants (Baker, 1997; Dowty, 1991; Grimshaw, 1981; Jackendoff, 1990; Pinker, 1989), it is possible that this hierarchy is reflected in cognition as well and recent studies with adults support this parallel (Cohn & Paczynski, 2013; Isasi-Isasmendi et al., 2023; Ünal et al., 2024). However, very few studies have explored Agent preference in early childhood (see Brocard et al., 2024; Ünal et al., 2021). Furthermore, research suggests the possible role of animacy in this asymmetry of the saliency between Agents and Patients (New et al., 2007; Thorpe et al., 2001; see also Scholl & Gao, 2013). Here, we investigate Agent preference in early childhood and the role of the animacy of the Patient. We begin by discussing the linguistic evidence on thematic roles, then move onto empirical evidence about the perceptual salience of event participants.

1.1 Thematic Roles in Language and the Thematic Hierarchy

In linguistic theory, thematic roles capture the relationship between a verb and its constituents, such as noun or prepositional phrases (Davis, 2011; Levin, 2014). In simple causative events, the *Agent* is the doer of the action and, therefore, the causer of the event, while the *Patient* is the one directly affected by the action. For example, in the sentence 'The man pushed the woman,' the Agent is encoded in the subject noun phrase, 'the man,' and reflects who performed the pushing action. The direct object noun phrase, 'the woman' is the Patient and reflects who is directly affected by the action. As in the example above, the Agent is usually encoded in the subject noun phrase and the Patient is encoded in the direct object noun phrase (Baker, 1997; Jackendoff, 1990). More complex events can have additional roles, such as the *Goal* (the endpoint of the action), the *Source* (the beginning of the action), or the *Instrument* (the means of performing the action). In the sentence 'The girl rode on her bicycle from her house to the school'; the Goal is 'her school,' the Source is 'her house,' and the Instrument is 'her bicycle'.

In one view of the Thematic Hierarchy, Agents are suggested to hold the highest prominence among event roles, followed by Patients, Goals, Sources, and lastly, Instruments (Baker, 1997; Grimshaw, 1981; Jackendoff, 1990; Pinker, 1989; see Rissman & Majid, 2019 for a review) and roles higher in the hierarchy contribute more to the event structure (Levin & Rappaport Hovav, 2005). Gärdenfors (2014) specifically notes that Agents and Patients are the key roles of a causative event since they are the two elements of causality: The causer and the caused. This formulation of the Thematic Hierarchy is largely based on how different roles behave in language. For instance, Agents and Patients tend to be mentioned more frequently (Ünal et al., 2021; 2024) and encoded in more central syntactic positions like Subject and Object, respectively (Grosz et al., 1995). Furthermore, languages are overall more likely to have a word order where the subject is mentioned before the object when both are mentioned in a sentence (Greenberg, 1963;

Haspelmath & Siegmund, 2006). Among languages that have a dominant word order, Subject-Object-Verb (SOV) and SVO are the most common orders, followed by VSO (Dryer, 2013). In all these word orders, the subject precedes the object. The Thematic Hierarchy is further supported by sign language/homesign and sentence comprehension and prediction studies (For sign language/homesign see Flaherty, 2014; Sandler et al., 2005; Padden et al., 2009; Goldin-Meadow & Mylander, 1998; Rissman & Goldin-Meadow, 2017; for sentence comprehension and prediction see Altmann, 1999; Boland et al., 1995; Carlson & Tanenhaus, 1989; Kamide et al., 2003; Kim & Osterhout, 2005; Mauner & Koenig, 2000; Trueswell et al., 1994; Frankland & Greene, 2015).

1.2 Event Participants in Cognition

Even though the Thematic Hierarchy was formulated to explain linguistic representation of event roles, it has been assumed to have roots in cognition (Baker, 1997; Dowty, 1991; Grimshaw, 1981; Pinker, 1989; see also Rissman & Majid, 2019; Strickland, 2017). According to Jackendoff (1990), thematic roles are "relational notions defined structurally over conceptual structure" (p. 47) and "every putative thematic role assignment must be justified on the grounds of its place in conceptual structure" (p. 50).

Recently, empirical evidence for this proposal has been growing. One piece of evidence comes from studies showing that event participants can be rapidly extracted from brief viewings of events. In one study, participants could reliably answer questions about the event and event participants (Agent and Patient) after viewing the events for 37 or 73 ms (Hafri et al., 2013). Another study asked participants to locate an individual based on their gender or shirt color (Hafri et al., 2018). They found that participants showed a role switch cost, meaning they were slower to locate the target if the target's event role was different from the previous trial. Based on these findings, the authors noted that the extraction of abstract event participants is rapid and spontaneous. In a more recent study using a similar paradigm, the stimuli did not depict actual events, but Agent-Patient roles were assigned based on body posture only (Agents had more dynamic, and Patients had less dynamic postures; Vettori et al., 2023). Regardless, the results once again showed a role switch cost. These findings point to event participants, particularly Agent and Patient, playing a significant role in the conceptual structure of as they are subject to spontaneous and fast extraction in human cognition.

Another study investigated the rapid extraction of event participants as well (Dobel et al., 2007). They asked participants to describe the event or to name the actors, objects, and action. They uncovered similar findings: Participants could identify the Agent and Patient in 200, 250 and 300 ms. Furthermore, their results showed that Agents were consistently identified more correctly than Patients. This suggests the higher relative salience of Agents over Patients. This Agent preference in human cognition has been investigated further. Cohn and Paczynski (2013) conducted a series of experiments to investigate the idea that Agents initiate the building of event structure. They used sequential images that depicted visual narratives. In one experiment, they investigated the predictive power of Agents and Patients in preparatory states. They found that Agents elicited a greater level of prediction about the immediate event in the next panel than Patients. In the following experiment, their participants spent more time looking at Agents than Patients. Lastly, seeing Agents before actions reduced the viewing time of the completed actions compared to seeing Patients before actions. The authors noted that these results support the idea that Agents play a central role in understanding and predicting event structure.

Isasi-Isasmendi and colleagues (2023) also investigated Agent preference in cognition. Their participants were speakers of Basque and Spanish. These languages were chosen because Basque overtly marks Agents and leaves Patients unmarked, while Spanish uses unmarked nominative case for both Agents and Patients. Participants were briefly exposed (300 ms) to two-participant, Agent-acting-on-Patient events. After a mask image, participants were either asked to perform an event description or a probe recognition task ('Was this the person in the image?'). The results showed higher specificity in event description task, greater accuracy, and faster response times in probe recognition for Agents than Patients. Furthermore, participants' first fixations were mostly to Agents rather than Patients or Actions. These results were true for speakers of both languages and there were no differences between languages. While these results support a cognitive preference for Agents, they also contribute to the generalizability of the phenomenon to syntactically diverse languages.

Another study of interest is by Ünal and colleagues (2024). They investigated asymmetries in event roles salience in language and cognition in relatively complex causative motion events that included Agents, Patients, Goals, and Instruments. In a linguistic description task, they found that English speakers mentioned Agents more than Patients, Patients more than Goals, and Goals more than Instruments. In a visual search task (e.g., 'Look at the one performing the action as quickly as possible'), participants fixated faster on Agents compared to Patients (and Patients faster than Goals and Instruments). In a change blindness task, they found faster change detection for Patients than for Goals and for Goals than Instruments. Interestingly, changes in Patients were detected faster than changes in Agents. Researchers argue that this was because the whole object changed color for Patients (as well as Goals and Instruments), however only part of the Agents (their clothes) changed color.

In summary, research reveals that Agents are consistently prioritized over other event participants. Studies on Agent preference highlight Agents' predictive power (Cohn & Paczynski, 2013) and attentional salience (Isasi-Isasmendi et al., 2023; Ünal et al., 2024), thus we can conclude that they play a central role in event perception. Conclusions are strengthened due to similar findings across languages (Isasi-Isasmendi et al., 2023), tasks (Ünal et al., 2024), and methods (Cohn & Paczynski, 2013; Isasi-Isasmendi et al., 2023; Ünal et al., 2024).

1.2.1 Agent Preference in Non-Human Animals

If Agents are the key participants in understanding events and Agent preference is potentially a foundational cognitive bias in how we perceive causative events, it might have evolved earlier than humans. Recent studies on Agent preference in non-human primates are in favor of this view. In a study by Brocard and colleagues (2024), great apes [western chimpanzees (Pan troglodytes verus), western lowland gorillas (Gorilla gorilla gorilla), and Sumatran orangutans (Pongo abelii)], human adults and children (2-12 years of age) watched videos of two participant Agent-Patient events on a touchscreen. The events included animate or inanimate Agent and Patient combinations, so there were four conditions (animate acting on animate, animate acting on inanimate, inanimate acting on animate, and inanimate acting on inanimate). The last frames remained on the screen after the videos and participants made a choice to touch either the Agent or the Patient. Results showed a robust Agent preference among all the species in the animate acting on inanimate condition. Further, Agent preference increased with age among humans in this condition. In the animate acting on animate condition, only adult humans and orangutans showed an Agent preference. In the inanimate acting on animate condition, only human children and gorillas preferred Agents over Patients. Lastly, no Agent preference was observed in the inanimate acting on inanimate condition. Authors argue that the animate acting on animate events may be more complex than animate acting on inanimate events or participants may expect reciprocity in animate acting on animate events, which may have led to the reduced Agent preference in animate acting on animate condition. This study provides important insights into our understanding of the evolution and working mechanisms of Agent preference. To our knowledge, there is only one more study on Agent preference among apes. In this study, researchers investigated Agent preference in chasing events among Guinea baboons (Papio papio; Meewis et al., 2024). They used a change detection paradigm where the participants had to indicate in which participant the color change occurred. Their results showed that the Guinea baboons were faster to detect the change to the Agents than the change to the Patients. Even though it requires further investigation, these studies suggest that Agent preference is not specific to humans and may have evolved much earlier. This supports the idea that Agent preference forms an essential part of cognitive event processing.

1.2.2 Agent Preference in Cognitive Development

If Agent preference is evolutionarily early and a core aspect of (event) cognition, it would be expected to develop at an early age. However, studies investigating Agent preference among infants and children are limited. One line of research related to event participants in infant cognition involves infants' grasp of causality and intentionality, which are crucial for interpreting social interactions (Woodward, 2013). Studies suggest that by around 10 months of age, infants begin to recognize causality in events involving two participants. Ten-month-olds who habituated to causal events dishabituated to delayed launching and no-contact (non-causal) events and vice versa, but those who habituated to one type of non-causal event did not dishabituate to the other type of non-causal event (Oakes & Cohen, 1990). By 10 months of age, infants also attribute causal force to a specific Agent linked to an observable action (Cohen & Oakes, 1993). The sensitivity to intentionality may develop even earlier. For instance, five-month-old infants show heightened attention when an Agent's actions toward an object do not match prior patterns in simple reaching tasks, indicating that they can infer goal-directed behavior (Woodward, 1998). Infants also expect Agents to take the most efficient path to achieve a goal, showing that even young infants can infer goal-directedness or intentionality based on context (Gergely et al., 1995). Further, infants at 18 months can reproduce an intended action even if they only witnessed an adult's failed attempt to complete it, suggesting that young children can infer intent beyond observable outcomes (Meltzoff, 1995).

Earlier research also shows that infants have some understanding of Agent and Patient roles. Studies demonstrated that 15-month-old infants were dishabituated when the roles of Agent and Recipient were reversed, indicating sensitivity to the expected structure of role assignments (Golinkoff, 1975; Golinkoff & Kerr, 1978). This finding was later replicated with 7-month-old infants who saw animated human Agents and Recipients (Leslie & Keeble, 1987). A recent study investigated whether seven-month-old infants represented Agent-Patient roles using finer measurement: Pupil dilation (Papeo et al., 2024). Their stimuli consisted of two-body dyads featuring Agent-like and Patient-like postures, with two male and two female models dressed in blue and orange. Each posture was designed to appear either dynamic (Agent-like) or less dynamic (Patient-like). In the habituation phase, participants viewed twelve dyads with stable Agent-Patient roles. Following habituation, they were shown alternating trials with either the same or switched roles. The results revealed that infants exhibited significant pupil dilation during role-switch trials compared to same-role trials, indicating sensitivity to the relational structure of the dyads. The authors concluded that seven-month-old infants represent and differentiate abstract Agent and Patient roles.

Research on infants' understanding of event participants demonstrates that they can recognize causality (e.g., Cohen & Oakes, 1993), intentionality (e.g., Meltzoff, 1995), and the abstract relational structure of Agent and Patient roles (e.g., Leslie & Keeble, 1987) from a remarkably early age. These findings provide distant support for a foundational role of Agent preference in cognitive development. However, infant studies directly investigating the event participants remain limited, often utilizing the same task (habituation; Golinkoff, 1975; Golinkoff & Kerr, 1978; Leslie & Keeble, 1987; Papeo et al., 2024). Furthermore, these studies indicate that infants understand the relational structure of Agents and Patients, not that they show an Agent preference.

One study has directly investigated Thematic Hierarchy and Agent preference among 3- to 5-year-old children (Ünal et al., 2021). English- and Turkish-speaking preschoolers saw images on a laptop screen. The images depicted causative events with four participants: Agent, Patient, Instrument and Goal. The Agent was an adult human and the Patient, Instrument, and Goal were inanimate objects in all of the images. An example event is a man hitting a ball into a bucket with a golf club. Each preschooler completed picture description and change blindness tasks. They were asked to describe what happened in each picture for the picture description task. For the change blindness task, new versions of the pictures in which a participant's color changed were created. Preschoolers were instructed to 'point at or name the changing object' as quickly as possible. Results of the picture description task were in line with the predictions of the Thematic Hierarchy and showed that speakers of both languages mentioned Patients more than Goals and Goals more than Instruments. On the other hand, English speakers mentioned Agents frequently, but Turkish speakers usually omitted Agents, which is not surprising considering Turkish allows subjects to be omitted because the verb is marked for person (first/second/third person) and number (singularity/plurality of the subject) (Göksel & Kerslake, 2005). Moreover, the results of the change blindness task indicated that preschoolers were faster at detecting changes in Patients and Goals than in Instruments. In line with a similar adult study by the researchers (Ünal et al., 2024), changes in Agents were not detected as accurately as expected. As mentioned above, this is probably due to the nature of the color change (clothes of the Agents changed color, as opposed to the whole object changing color in Patients, Goals and Instruments). Even though this study provides insights into the Thematic Hierarchy and Agent preference among young children, the differences in the nature of the color changes and animacy of the Agent/Patient introduce limitations to this study.

1.3 Unresolved Issues in Agent Preference

As discussed so far, Agent preference has been somewhat extensively studied. However, this work leaves open several issues that need to be considered in order to fully understand Agent preference in cognition. One possible factor that explains this preference might be animacy. In the events we witness in our everyday lives, Agents are usually people or animals (animate), but this is not always the case for Patients. The visual system is especially tuned to recognize animate image content (Thorpe et al., 2001). When viewing natural scenes, attention is automatically drawn to animate Agents such as animals relative to inanimate objects such as vehicles and plants, even when lower-level stimulus differences are controlled (New et al., 2007). Further, functional magnetic resonance imaging (fMRI) studies suggest that our visual pathways are tuned to the animate-inanimate distinction (Konkle & Caramazza, 2013; Sha et al., 2015).

Visual features like faces play a role in animacy's salience (e.g., Frank et al., 2009). This effect is thought to arise because the brain has specific routines for face processing that do not extend to other body parts (Kanwisher, 2010; Sinha et al., 2006; Sugita, 2009; Wilmer et al., 2010). This selective processing may also operate at a finer level, with certain facial features, like the eyes, receiving even more specialized attention (e.g., Emery, 2000; see also Scholl & Gao, 2013 on the salience of animacy).

In some of the above-mentioned Agent preference studies, Agents are animate, and Patients are inanimate objects (e.g., Ünal et al., 2021; 2024). Moreover, the studies that included animacy as a variable yielded different results for animate/inanimate Agents and Patients (e.g., Brocard et al., 2024). Thus, the results of the studies reviewed above can be difficult to interpret without understanding the role of animacy in Agent preference.

To our knowledge, there have been three studies that considered the effects of animacy while investigating Agent preference. The stimuli of Cohn and Paczynski (2013) included mainly animate Agents and Patients but also a few inanimate Agents and Patients. They reran their analysis by excluding all inanimate entities. In this case, the higher predictive power of Agents over Patients was only marginally significant. All other main findings of the study remained significant in the animate-entities-only analyses. The second study that considered animacy as a factor was by Isasi-Isasmendi and colleagues (2023). The Patient was an animate entity (human) in half of their stimuli, and an inanimate object in the other half. They included animacy as a nuisance predictor in their analyses. Their Agent preference findings held true independent of Patient animacy. The last study that investigated Agent preference while considering animacy was conducted by Brocard and colleagues (2024). As discussed above, their results showed a stronger

Agent preference in animate-acting-on-inanimate events and a weaker preference in animate-acting-on-animate events. The literature reviewed on the salience of animacy and the nuisances that arise from Agent preference-animacy studies points to the importance of including animacy as a variable when investigating Agent preference.

There might also be other low-level features that can pose as confounds. One of these may be size as literature shows that objects that are bigger attract more attention compared to smaller ones (Wolfe & Horowitz, 2017). Furthermore, bigger and smaller real-life sizes activate cortical zones with distinct response preferences (Konkle & Caramazza, 2013). Therefore, the sizes of event participants should be kept similar or controlled when studying Agent preference.

There may be other confounding factors that contribute to findings in support of Agent preference. Researchers have used a few different methods to disrupt event structure and demonstrate that Agent preference is likely due to event participants and not other factors. For example, in their last experiment, Ünal and colleagues (2024) removed event structure by replicating the change blindness task (explained above in detail) with single objects (e.g., Patient in the same location with the same size as in the previous task but without other event participants). The results showed no difference in change detection times between single Agents, Patients, Goals, and Instruments. Moreover, the change detection reaction times for single Agents and Patients were similar to Agents and Patients in a coherent event structure. Based on these results, the authors concluded that the differences in reaction times observed in the organized events are not due to the physical features of the individual event participants, such as shape, size, or color. However, comparing organized events to single object scenes introduces the problem of scene complexity. Hence, it may not be the most useful method to determine whether observed differences are due to event participants.

Another method to disrupt event structure is using the upside-down versions of the same events as controls. This method ensures that the visual features and complexity remain the same. It has been used in a study by Goupil and colleagues (2022). When comparing infants' looking times at facing and non-facing body dyads, they used upside-down versions of the images as well. They found that infants looked longer at non-facing than facing dyads in right-side-up images, but no difference between upside-down dyads was observed. Two studies by Papeo and colleagues have utilized this method to study interaction perception as well (Papeo et al., 2017; Papeo et al., 2019; see also Ünal et al., 2024 for converging evidence from adults). Using upside-down versions allows many visual features, such as size, color, shape, and complexity, to be kept constant; however, it introduces a new confounding variable: orientation.

A more useful method to disrupt event structure is using new versions of the stimuli in which event participants are back-to-back as controls. As briefly mentioned above, Papeo and colleagues have compared face-to-face body dyads to their back-to-back versions and found a higher inversion effect for face-to-face dyads (Goupil et al., 2022; Papeo et al., 2017; Papeo, 2019). This line of research shows that face-to-face body dyads are perceived as meaningful units and inverting them causes the coherence to be disrupted. On the other hand, the inversion effect is lower for back-to-back dyads, indicating that they were not perceived as coherent scenes to begin with. Further support for this method comes from an EEG study that investigated intermodulation responses to face-to-face and back-to-back dyads (Goupil et al., 2023). Face-to-face dyads showed higher intermodulation in higher-order frequencies which indicates integration. However, back-to-back dyads showed higher individual frequencies, which indicates a less integrated and more individual perception of bodies in this configuration. To our knowledge, this method has only been utilized by one study investigating Agent preference. Dobel and colleagues (2007) found that Agents were identified more accurately in face-to-face than in back-to-back images. Given this robust control in place, this finding indicates that Agent preference is likely due to the significance of Agents in event perception and not due to other visual factors. This method of disrupting event structure provides a means to maintain the same orientation, visual complexity, and low-level features of the individual entities.

1.4 The Present Study

In the present study, we seek evidence for Agent preference in causative events among young children. While prior studies have explored Agent preference among adults (e.g., Cohn & Paczynski, 2013; Isasi-Isasmendi et al., 2023), as well as infants' and children's understanding of Agent and Patient roles (e.g., Leslie & Keeble, 1987), only a few have investigated Agent preference with young children (see Brocard et al., 2024; Ünal et al., 2021). One of these studies (Brocard et al., 2024) had a wide age range for their child sample (2 to 12 years of age) which makes it difficult to draw a conclusion about Agent preference among young children. Another child study (Ünal et al., 2021), used human Agents and non-human/inanimate Patients in all their stimuli. Although this is parallel to many everyday events, the animate/inanimate difference makes it difficult to directly compare the results for Agents and Patients as thematic roles. Hence, in our study, we seek to address this open question by examining how cognitively salient Agent-Patient roles are for preschoolers across two levels of Patient animacy.

Our first aim is to assess whether children show a preference for Agents over Patients in causative scenes. We address the limitations of previous research by providing a systematic examination of Agent and Patient roles in causative event depictions. To this end, we use a change-detection task, in which preschoolers view causative events. Previous studies have used reaction time measurements for change detection with adults (e.g., Ünal et al., 2024), but only accuracy measures have been utilized with children due to the challenge in obtaining reliable reaction time measures in this age group (e.g., Ünal et al., 2021). Because accuracy may provide coarse measurements, we also use an eye tracking paradigm to get finer and more reliable measurements with young children.

Our second aim is to determine whether children's Agent preference is affected by whether the Patient is animate or inanimate. For instance, prior work indicates that attention to Agents may vary when the Agent interacts with an animate versus an inanimate Patient (Brocard et al., 2024). By comparing responses to events in which Agents interact with animate Patients to those in which Agents interact with inanimate Patients, we aim to understand how Agent preference is impacted by Patient animacy.

Our third aim is to examine whether any difference we find in saliency is due to the conceptual roles of event participants and not due to other visual features. Prior research with adults and children has shown that eliminating the standard face-to-face orientation can disrupt event coherence (e.g., Goupil et al., 2022; Dobel et al., 2007; Vettori et al., 2023). Thus, we present each causative event in both orientations, face-to-face and back-to-back, allowing us to examine whether back-to-back configurations diminish the preference for Agents.

By investigating these gaps in the literature, the present study aimed to advance our understanding of the factors that influence causative event perception in young children. Specifically, we believe our results provide insights into Agent preference in development and factors that play a role in this cognitive bias. Lastly, because our participants are native speakers of Turkish, which is a pro-drop language (Göksel & Kerslake, 2005), our results contribute to the generalizability of Agent preference across languages. To this end, we used a change detection task (Rensink et al., 1997) with Turkish-speaking preschoolers. In each trial, preschoolers were presented with causative events that flicker back and forth between original and modified images, with color changes applied to either the Agent or the Patient. Their task was to detect which character changed color, allowing us to investigate their preference for Agents versus Patients across two levels of Patient animacy (animate and inanimate) and two levels of event coherence (face-to-face and back-to-back). Higher change detection accuracy for an event participant would indicate easier detection and thus higher salience of the event participant. We recorded their eye movements and calculated the proportion of fixations to the changing participant as well as calculated the accuracy of their responses to 'Which one changed color?'.

Based on the literature reviewed above, we hypothesized that children would demonstrate a stronger Agent preference in the face-to-face events that include inanimate Patients and a weaker Agent preference in the face-to-face events that include animate Patients. We expected more accurate detection of change in Agents than in Patients and differing timecourse of fixation proportions to Agents and Patients in these conditions. Lastly, we anticipated these Agent/Patient asymmetries to disappear in the back-to-back events regardless of Patient animacy**.** Alternatively, if Agent preference is due to animacy only, we would expect an Agent preference in the face-to-face and back-to-back events that include inanimate Patients but not in the events with animate Patients. Further, if Agent preference is due to other visual features of the entities, such as color, shape, or complexity, we would expect face-to-face and back-to-back trials to reveal similar results since these features remain the same between face-to-face and back-to-back trials.

2. METHOD

2.1 Participants

Data were collected from preschoolers (*n* = 34, 26 females, M_{age} = 4.65 years, SD_{age} = 0.55, range = 3.42 – 5.33). Sample size was determined based on power analysis. All were native speakers of Turkish. Participants were recruited from various preschools in Istanbul, Turkey. Ethical approval for this study was obtained from Özyeğin University Research Ethics Committee (2021/14/02). None of the participants had teacher-reported language, speech, or developmental problems. An additional 14 preschoolers were tested but excluded from the analysis because they did not follow the instructions (*n* = 2), did not want to complete the experiment (n = 1), had less than 50% accuracy throughout the entire task (n = 4) or did not fit the age criteria (n = 7).

2.2 Materials

Target stimuli consisted of 12 illustrations that depicted causative events in which a person (Agent) acted on another person or object (Patient). Sample events are provided in Figure 2.1. In these events, a woman (Agent) tickles a man (Patient; Fig.2.1 a), and a man (Agent) knocks down stacked boxes (Patient; Fig.2.1 d). The Agent was always an adult human. The Patient was an adult human in half of the target events (Fig.2.1 a) and an inanimate object in the other half (Fig.2.1 d). When selecting the events and creating the illustrations with inanimate Patients, we ensured that the objects were similar in size to humans to eliminate the possible effect of size on salience (Konkle & Caramazza, 2013; Wolfe & Horowitz, 2017). Gender was equally distributed and counterbalanced for Agents and Patients. Agents and Patients face each other in all the illustrations. Additional versions of the images, in which part of the Agent's clothing color was different, were created (Fig.2.1 b and e). The same was done for the Patients (for inanimate Patients, parts of the objects similar in size to the humans' clothing color were changed; Fig.2.1 c and f).

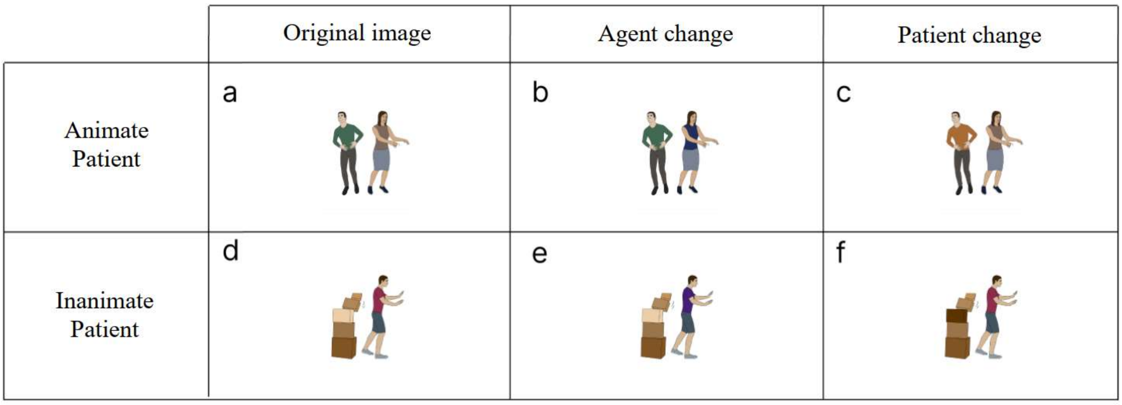
Figure 2.1 *Examples of Face-to-Face Event Illustrations*

	Original image	Agent change	Patient change
Animate Patient	a 	b 	c 
Inanimate Patient	d 	e 	f 

Note. (a) With animate Patient, (b) color change in Agent version of an event with an animate Patient, (c) color change in Patient version of an event with an animate Patient, (d) with inanimate Patient, (b) color change in Agent version of an event with an inanimate Patient, (c) color change in Patient version of an event with an inanimate Patient.

We also created new versions of the images in which each event participant's vertically symmetrical image was taken (See Figure 2.2 a and d). This was done to disrupt the event structure while keeping the same visual features such as body configurations and sizes. These images do not depict events, therefore the people or objects in them are not participants of an event. However, for ease of reference throughout the paper, we will refer to them as Agents or Patients. Color change in Agent and Patient versions of the back-to-back images were created as well (Fig.2.2 b, c, e and f). The stimuli were created using Procreate 5.3.10 and edited using Adobe Photoshop 2024. The illustrations depicted the midpoints of the events on a white background.

Figure 2.2 *Examples of Back-to-Back Event Illustrations*



Note. (a) With animate Patient, (b) color change in Agent version of an event with an animate Patient, (c) color change in Patient version of an event with an animate Patient, (d) with inanimate Patient, (b) color change in Agent version of an event with an inanimate Patient, (c) color change in Patient version of an event with an inanimate Patient.

There were also filler illustrations that depicted noncausative events involving two people (e.g., two men shaking hands). These were included because while they depict events, they do not show an Agent acting on a Patient. This provided an opportunity to check that the manipulation worked due to event participants and not low-level explanations (see Appendix A for a complete list of events).

2.2.1 Selection of the Stimuli

The stimuli used in the main task were selected based on two pilot studies. The first pilot study had two goals: (1) to ensure that the roles of the depicted people or objects (i.e., who was the Agent and who was the Patient) could be clearly identified from the causative event illustrations and (2) to ensure that the events could be understood from the target illustrations.

The participants in the pilot study were 21 adult native speakers of Turkish. None of the participants who took part in the pilot study participated in the main experiment.

Ratings were obtained for a larger set of 25 events: Nine causative events with animate Patients, nine causative events with inanimate Patients, and eight noncausative events. Data were collected online via Qualtrics. Participants saw each original face-to-face illustration individually in a randomized order. For each causative event illustration, participants responded to two close-ended questions: "*Which one is performing the action?*", and "*Which one is affected by the action?*" (options: man and woman or man and boxes etc.). They also responded to one open-ended question for each causative or noncausative event: "*Describe the action depicted in the image with one verb*." All questions and response options were presented in Turkish. The task lasted approximately 10 minutes.

Based on these data, we selected 12 target and 6 filler images: Six causative events with animate Patients, six causative events with inanimate Patients, and six noncausative events. The selection was based on the highest percentage of the participants using the target verb and correctly identifying the Agent and the Patient. The accuracy and agreement ratings for the final set were as follows: Agent identification accuracy ranged from 0.90 to 1.00 (M = 0.97, SD = 0.04), Patient identification accuracy ranged from 0.90 to 1.00 (M = 0.97, SD = 0.03), verb agreement ranged from 0.62 to 1.00 (M = 0.85, SD = 0.14).

A second norming study was conducted to ensure that the face-to-face depictions were perceived as events and the back-to-back depictions were not. The participants were 20 adult native speakers of Turkish. None of the participants who took part in the pilot study participated in the main experiment.

Ratings were obtained for 12 target and 6 filler images, as well as their back-to-back versions. Participants either saw all the face-to-face images or the back-to-back images. For each image, they rated on a scale of 1 to 5 whether "*The people and/or objects in the image form a coherent event.*" (1 = *Strongly disagree*, 5 = *Strongly agree*). All questions and response options were presented in Turkish. The task lasted approximately 5 minutes.

Across all categories (causative-animate Patient, causative-inanimate Patient, noncausative), the face-to-face images (M = 4.07, SD = 1.19) were rated significantly higher than back-to-back images (M = 2.43, SD = 1.32) as depicting events, $t(30.765) = -8.493$, $p < .001$. Event ratings by category can be found in Appendix B.

2.3 Procedure

Participants were tested individually in a quiet room in their preschool. Participants were seated approximately 60 cm away from a DELL Precision M4800 laptop with the SMI RED 250 eye-tracker (SensoMotoric Instruments) mounted underneath the screen. The stimuli were presented via NBS Presentation software (Version 23.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Each version of the event (change in Agent or Patient and face-to-face or back-to-back) was seen by four different participants, but each participant saw each event only once. Further, each participant saw 3 instances of Agent and 3 instances of Patient change in each category. The side of the color change was counterbalanced within events[1]. Each participant saw a total of 12 trials and 6 filler trials. The trials were presented in a single randomized order.

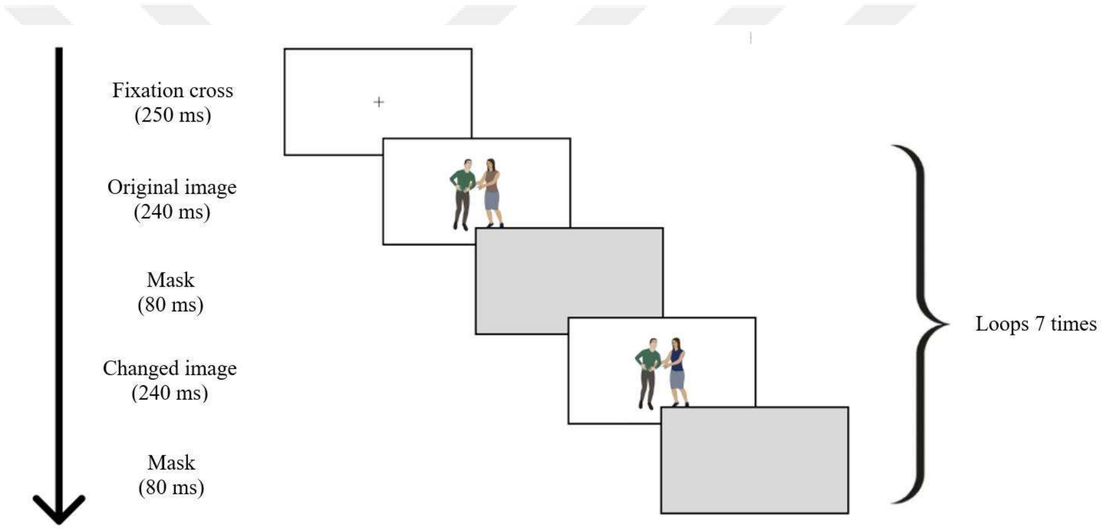
The change blindness procedure was adapted from the established paradigm of Rensink and colleagues (1997) (see also Aginsky et al., 2000; Hollingworth & Henderson, 2000). It was also used with children aged 3-12 in other studies (Ünal et al.,

---

[1] Due to an error in counterbalancing lists, back-to-back and face-to-face events were distributed unevenly across participants but the other variable levels (changing role and Patient animacy) were distributed evenly.

2021; Fletcher-Watson et al., 2009; Shore et al., 2006). Each trial began with a fixation cross that was presented in the middle of the screen for 250 ms. Then, the original event appeared on the screen for 240 ms, followed by a grey screen mask that appeared for 80 ms. After the mask, the color-changed version of the event appeared for 240 ms, again followed by a mask for 80 ms (Figure 2.3). This 640 ms loop was repeated 7 times, which amounted to 4.48 seconds in total.

Figure 2.3 *Trial Structure*



Participants were instructed that they would see flashing pictures on the screen and that something in the picture would change color. They were told they would be asked on what side the color change was after the pictures were done flashing. Participants were instructed to raise their left or right hand based on the color change side. They were asked to sit as still as possible and keep their hands on the table for accurate eye tracking (See Appendix C for the instructions in the original Turkish and the English translation).

At the beginning of the experiment, participants completed two practice trials, which depicted two cartoon animals standing opposite each other, one of which changed color. The animal on the left changed color in the first practice trial, and the animal on the right changed color in the second one. Then, the participants completed a 5-point calibration and validation procedure. Participants were given an opportunity to ask questions before they began the main task with 18 trials.

3. RESULTS

Our research questions were: **(1)** Do preschool-aged children show Agent preference? **(2)** What role does animacy play in the salience of Agents and Patients? **(3)** Does the salience of Agents and Patients differ across coherent (face-to-face condition) and incoherent (back-to-back condition) events?

To address these questions, we first examined whether accuracy varied based on the changing role (Agent vs. Patient) and the animacy of the Patient (animate vs. inanimate). We further examined the variation in participants' target fixations depending on the changing role and the animacy of the Patient. Analyses were conducted for both face-to-face and back-to-back trials to confirm whether these differences were event-specific.

3.1 Accuracy Data

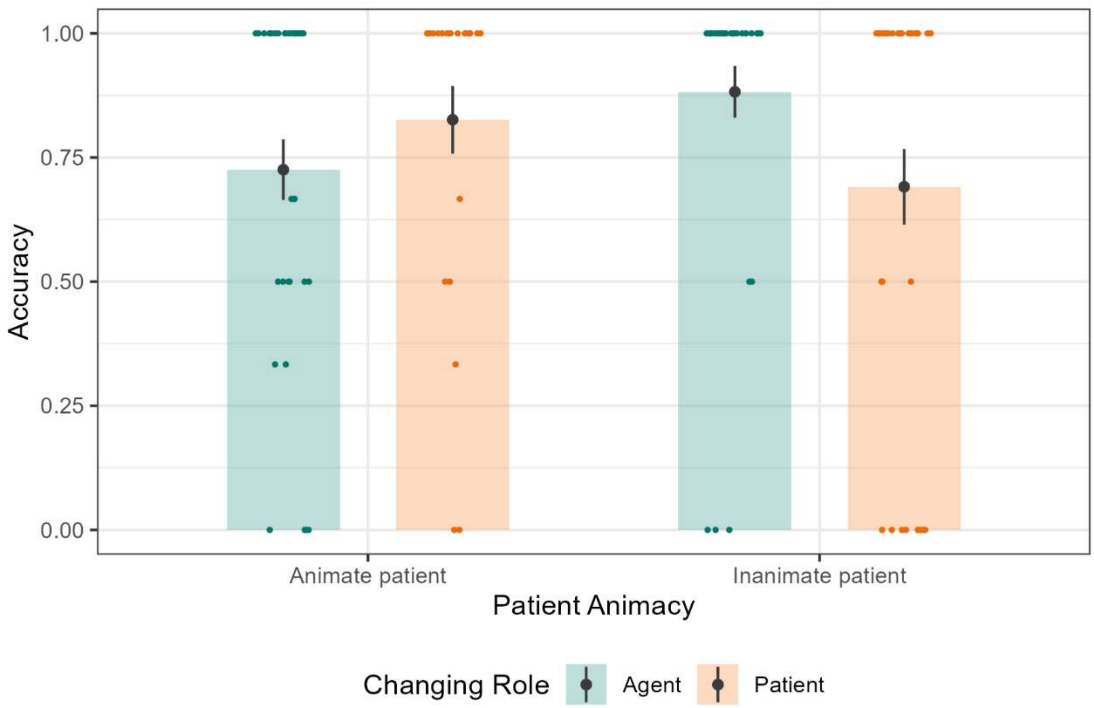
First, we checked whether participants had difficulty understanding the task and complying with the instructions. To do so, we identified those participants who had less than 50% accuracy throughout the entire task. Four participants who met this criterion were excluded from further analyses. All exclusions were done prior to data analysis.

Then, we examined the variation in change detection accuracy depending on the changing role and animacy of the Patient. We first analyzed the face-to-face trials to determine whether Agent preference emerged in events. We, then, examined the back-to-back trials to make sure any difference we found in face-to-face trials was due to event roles and not visual features.

3.1.1 Face-to-Face Events

Accuracy data were analyzed using generalized binomial linear mixed effects models. Models were fit using the *glmer* function in the *lme4* package (version 1.1.33; Bates et al., 2015) in R (version 4.2.3; R Core Team, 2023). The figure was produced using *ggplot2* package (Wickham, 2016). The face-to-face model tested the fixed effect of changing role (Agent, Patient) and animacy of the Patient (animate Patient, inanimate Patient) on the binary dependent variable of accuracy (1 = *correct*, 0 = *incorrect*) at the trial level (see Figure 3.1). The fixed effect of changing role and animacy was tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for Subjects only (Baayen, 2008; Baayen et al., 2008). A more complex model that also included random intercepts for Items produced a singular fit error, hence this term was excluded from the model.

Figure 3.1 *Accuracy across Changing Role and Patient Animacy in Face-to-Face Events*



The parameter estimates for the fixed effects from the glmer model for the accuracy data are presented in Table 3.1. The model revealed a significant interaction between changing role and animacy of the Patient. In the inanimate Patient condition, participants were more accurate when detecting changes to Agents (M = 0.88, SD = 0.33) than Patients (M = 0.73, SD = 0.45) (β = 1.2647, SE = 0.6145, z = 2.058, p = .04). There was no difference between detecting changes to Agents (M = 0.73, SD = 0.45) and Patients (M = 0.85, SD = 0.36) in the animate Patient condition (β = -0.8895, SE = 0.5464, z = -1.628, p = .10).

Table 3.1 *Parameter Estimates for the Fixed Effects from the Glmer Model for the Mean Accuracy of Face-to-Face Events*

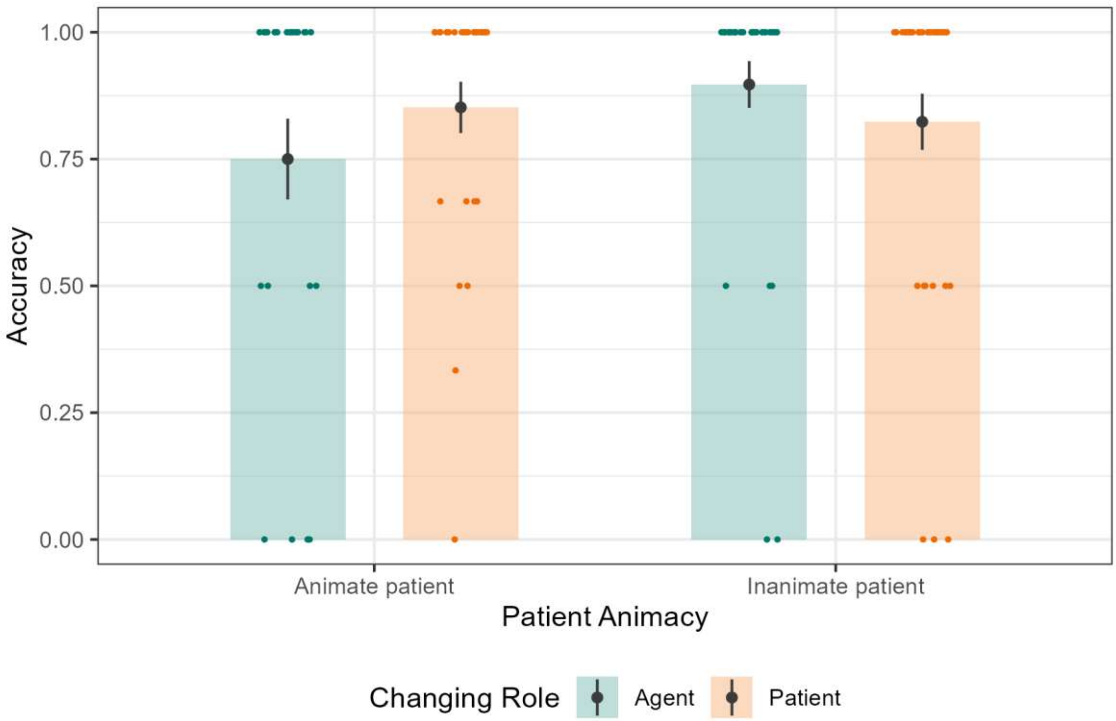
Effect	ß	*SE*	*z*	*p*-value
Intercept	1.820	0.343	5.301	<.001
Animacy (Animate Patient vs. Inanimate Patient)	0.150	0.423	0.354	.724
Changing Role (Agent vs. Patient)	-0.188	0.407	-0.461	.644
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient)	-2.154	0.831	-2.591	.010

*Note. Formula in R: Mean Accuracy ~ Animacy*Changing Role + (1 | Subject).*

3.1.2 Back-to-Back (non-)Events

The back-to-back model tested the fixed effect of changing role (Agent, Patient) and animacy of the Patient (animate Patient, inanimate Patient) on the binary dependent variable of accuracy (1 = *correct*, 0 = *incorrect*) at the trial level (see Figure 3.2). The fixed effects of changing role and animacy were tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for both Subjects and Items (Baayen, 2008; Baayen et al., 2008).

Figure 3.2 *Accuracy across Changing Role and Patient Animacy in Back-to-Back Events*



Parameter estimates for the fixed effects from the glmer model for the accuracy model are presented in Table 3.2. No fixed or interaction effects were significant. Participants were equally accurate at detecting changes to Agents and Patients in back-to-back events both with an animate or an inanimate Patient.

Table 3.2 *Parameter Estimates for the Fixed Effects from the Glmer Model for the Mean Accuracy of Back-to-Back Events*

Effect	ß	*SE*	*z*	*p*-value
Intercept	2.083	0.420	4.955	<.001
Animacy (Animate Patient vs. Inanimate Patient)	0.560	0.569	0.984	.325
Changing Role (Agent vs. Patient)	-0.113	0.452	-0.249	.803
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient)	-0.872	0.906	-0.963	.336

*Note. Formula in R: Mean Accuracy ~ Animacy*Changing Role + (1 | Subject) + (1 | Item).*

3.2 Eye Gaze Data

Participants' fixations to two areas of interest (AoI) were computed using the SMI BeGaze software: the Agent and the Patient. We analyzed eye gaze data only for the trials in which the participants correctly identified the changing component (81% of the data). We also assessed whether participants experienced more than 45% trackloss throughout the entire task or if individual trials had over 50% trackloss. No participant had to be excluded based on the criterion of having more than 45% trackloss across all trials. However, 15% of the accurate trials that had more than 50% trackloss within a trial were excluded from further analysis.

We then computed the proportion of fixations to changing event participant (target). This was done by first examining whether the participant looked at the target AoI, non-target AoI and remaining white space in each 100 ms time window. Then, for each time window, we calculated the proportion of target fixations as *target fixation / (target + non-target + white space fixations)* (See Figure 3.3 and Figure 3.4). From this, we aggregated the proportion of target fixations in 1000 ms time windows for the first 2000 ms of the trial. We hypothesized that eye movements during the earlier time windows would more precisely reflect the visual identification of target event roles since event participants are detected rapidly (Hafri et al., 2013; 2018). Furthermore, according to the literature, 1000 ms is enough to make a fixation on the changing target (Hirose & Hancock, 2007; Hochhauser et al., 2022). Thus, we adopted this approach to analyze the timeframe in which cognitive and perceptual processes in recognizing event roles occur, and to align our methodology with prior studies and theoretical expectations.

Figure 3.3 *Proportion of Fixations at Target across Changing Role and Patient Animacy in Face-to-Face Events*

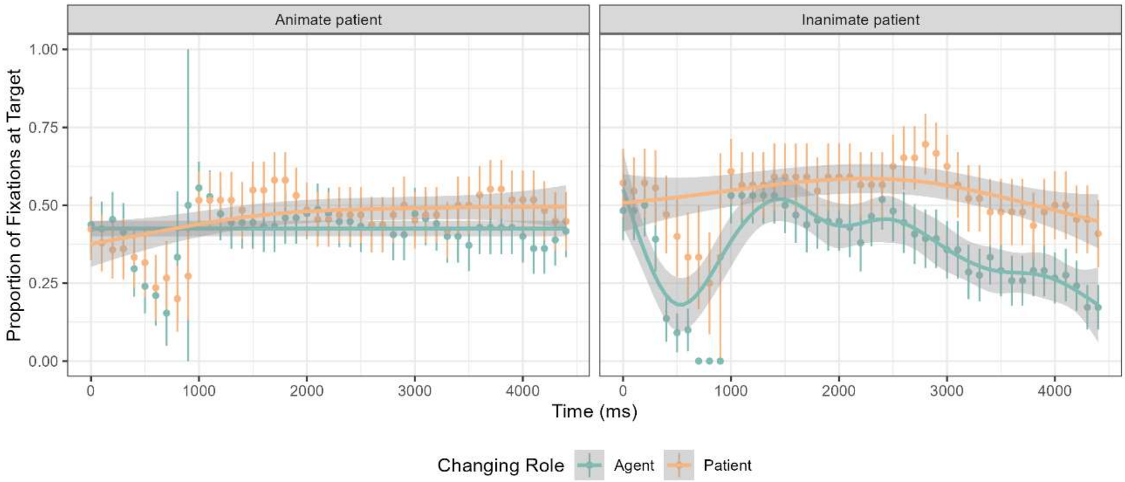
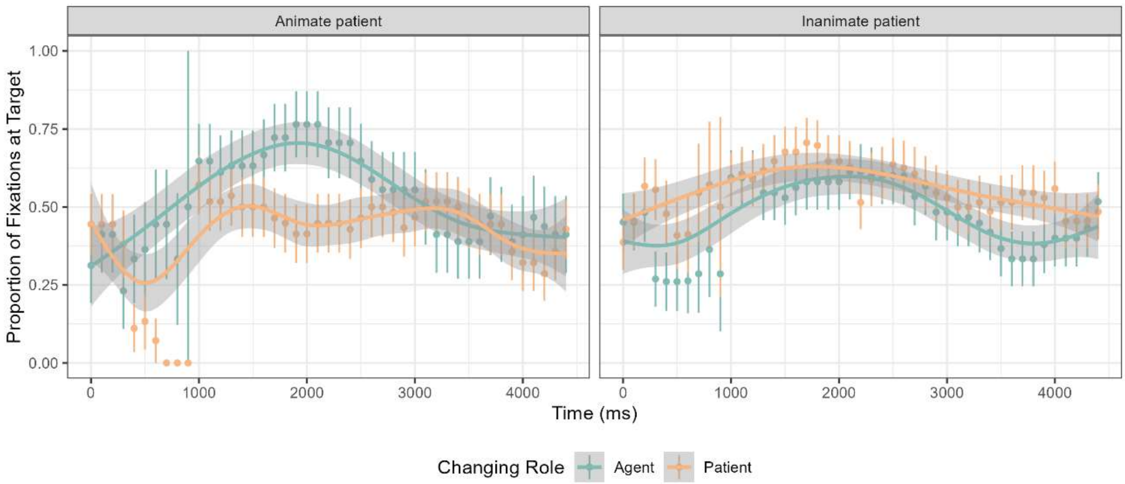


Figure 3.4 *Proportion of Fixations at Target across Changing Role and Patient Animacy in Back-to-Back Events*

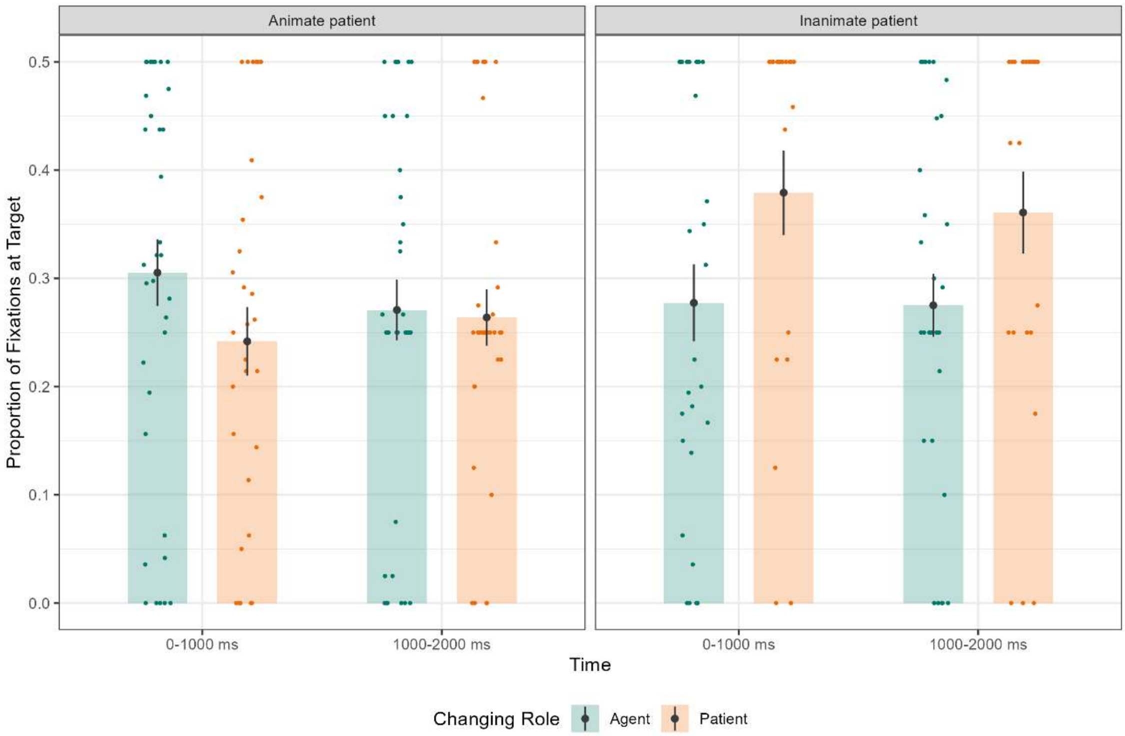


Eye gaze data were analyzed using linear mixed effects models. Models were fit using the *lmer* function in the *lme4* package (version 1.1.33; Bates et al., 2015) in R (version 4.2.3; R Core Team, 2023). The figures were produced using *ggplot2* package (Wickham, 2016). We examined the variation in participants' target fixations depending on the changing role, animacy of the Patient and time window. Face-to-face and back-to-back trials were analyzed separately.

3.2.1 Face-to-Face Events

The face-to-face model tested the fixed effect of changing role (Agent, Patient), animacy of the Patient (animate Patient, inanimate Patient) and time window (0-1000 ms, 1000-2000 ms) on the dependent variable of proportion of target fixations at the trial level (see Figure 3.5). The fixed effects were tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for Subjects only (Baayen, 2008; Baayen et al., 2008). A more complex model that also included random intercepts for Items produced a singular fit error, hence this term was excluded from the model.

Figure 3.5 *Proportion of Fixations at Target across Changing Role, Patient Animacy, and Time Windows in Face-to-Face Events*



Parameter estimates for the fixed effects from the *lmer* model for the proportions of target fixations model are presented in Table 3.3. The model revealed a significant interaction effect of changing role and animacy of the Patient. In the inanimate Patient condition, participants fixated more on the target when detecting changes to Patients (*M* = 0.37, *SD* = 0.18) compared to when detecting changes to Agents (*M* = 0.28, *SD* = 0.18) (β = -0.0869, *SE* = 0.0328, *z* = -2.646, *p* = .009). There was no difference between fixations durations when detecting changes to Agents (*M* = 0.29, *SD* = 0.18) and Patients (*M* = 0.25, *SD* = 0.16) in the animate Patient condition (β = 0.0182, *SE* = 0.0288, *z* = 0.63, *p* = 0.53).

Table 3.3 *Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Proportion of Target Fixations in Face-to-Face Events*

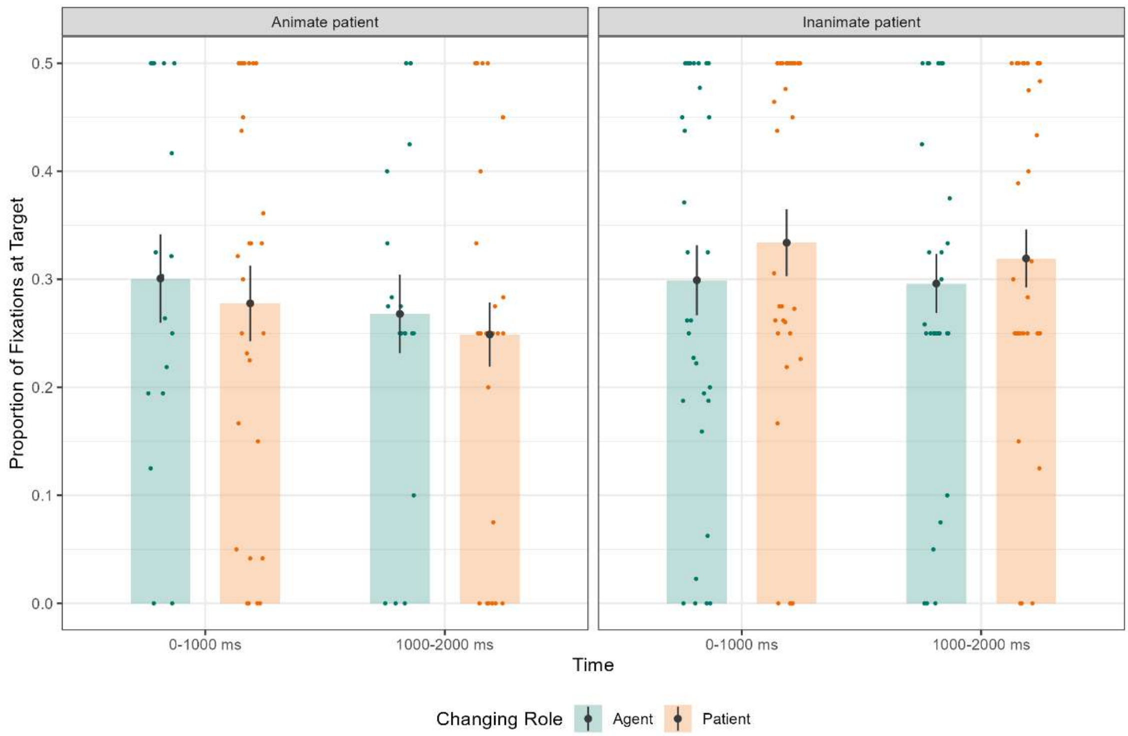
Effect	ß	*SE*	*df*	*z*	*p*-value
Intercept	0.297	0.016	30.038	18.727	<.001
Animacy (Animate Patient vs. Inanimate Patient)	0.040	0.023	246.796	1.772	0.078
Changing Role (Agent vs. Patient)	0.034	0.022	241.084	1.590	0.113
Time Window (0-1000 ms vs. 1000-2000 ms)	-0.008	0.021	215.482	-0.369	0.713
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient)	0.105	0.044	247.259	2.379	0.018
Animacy (Animate Patient vs. Inanimate Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	-0.005	0.042	215.482	-0.124	0.902
Changing Role (Agent vs. Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	0.019	0.042	215.565	0.462	0.645
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	-0.071	0.083	215.565	-0.846	0.399

*Note. Formula in R: Mean Proportion of Target Fixations ~Animacy*Changing Role*Time Window + (1 | Subject).*

3.2.2 Back-to-Back (non-)Events

The back-to-back model tested the fixed effect of changing role (Agent, Patient), animacy of the Patient (animate Patient, inanimate Patient) and time window (0-1000 ms, 1000-2000 ms) on the dependent variable of proportion of target fixations at the trial level (see Figure 3.6). The fixed effects were tested with sum-to-zero contrasts as the face-to-face model (Schad et al., 2020). The model included random intercepts for Subjects and Items (Baayen, 2008; Baayen et al., 2008).

Figure 3.6 *Proportion of Fixations at Target across Changing Role, Patient Animacy, and Time Windows in Back-to-Back Events*



Parameter estimates for the fixed effects from the *lmer* model for the proportions of target fixations model are presented in Table 3.4. The fixed effects or interactions were not statistically significant. Participants equally fixated on the target when detecting changes to Agents and Patients in back-to-back events both with an animate or an inanimate Patient.

Table 3.4 *Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Proportion of Target Fixations in Back-to-Back Events*

Effect	ß	*SE*	*df*	*z*	*p*-value
Intercept	0.293	0.024	17.329	11.962	<.001
Animacy (Animate Patient vs. Inanimate Patient)	0.049	0.040	8.897	1.216	0.255
Changing Role (Agent vs. Patient)	-0.003	0.022	184.507	-0.152	0.880
Time Window (0-1000 ms vs. 1000-2000 ms)	-0.018	0.018	189.482	-0.957	0.340
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient)	0.016	0.043	196.164	0.364	0.716
Animacy (Animate Patient vs. Inanimate Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	0.0143	0.037	189.402	0.388	0.698
Changing Role (Agent vs. Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	-0.007	0.037	188.781	-0.182	0.856
Animacy (Animate Patient vs. Inanimate Patient): Changing Role (Agent vs. Patient): Time Window (0-1000 ms vs. 1000-2000 ms)	-0.003	0.074	188.766	-0.042	0.967

*Note. Formula in R: Mean Proportion of Target Fixations ~ Animacy*Changing Role*Time Window + (1 | Subject) + (1 | Item).*

4. DISCUSSION

Accumulating evidence with adults suggests that people show a preference towards Agents above other event participants in causative events (Cohn & Paczynski, 2013; Isasi-Isasmendi et al., 2023; Ünal et al., 2024). Studies with non-human animals indicate that this preference might have an evolutionary basis (Brocard et al., 2024; Meewis et al., 2024). If that is the case, it would be expected to emerge early in cognitive development. However, the current evidence does not offer sufficient support for this supposition. Differences in animacy of the Agent and Patient, as well as coarse measurements, pose as limitations of previous studies. The present study addressed several unresolved questions about Agent preference in young children's perception of causative events. Prior work with children used stimuli mostly with animate Agents and inanimate Patients (e.g. Ünal et al., 2021), which may be parallel to everyday events, nevertheless complicates direct comparisons between roles due to animacy differences. Furthermore, previous research with children relied on accuracy (Ünal et al., 2021), which may not capture nuanced differences in role saliency.

This study addressed these gaps by examining preschoolers' preference for Agents versus Patients during a change detection task. We systematically varied Patient animacy to explore whether animacy modulates Agent preference. We also tested this preference in back-to-back conditions, which disrupts event coherence. This orientation control helped us draw conclusions about the factors affecting Agent preference. We were able to determine whether visual features such as size, shape and complexity play a role in Agent preference. Utilizing eye tracking allowed us to gain a deeper understanding by analyzing the proportion of fixations to target as well as accuracy.

4.1 Agent Preference in Development and the Role of Animacy

Our main goal was to assess whether children showed a preference for Agents over Patients in coherent causative scenes and whether this was modulated by Patient animacy. We expected higher change detection accuracy for Agents in face-to-face conditions than for Patients. We hypothesized that this difference between Agents and Patients would be stronger for the face-to-face events with inanimate Patients and diminish for the face-to-face events with animate Patients. In line with our expectations, children were more accurate at detecting changes to Agents than to Patients in face-to-face events with an inanimate Patient. However, no such asymmetry was observed in events with an animate Patient, as children were similarly accurate at detecting changes to Agents and Patients. Furthermore, as expected, no asymmetry between Agent and Patient change detection accuracy was observed in back-to-back conditions. This eliminates the possibility that the asymmetry in face-to-face events with inanimate Patients is due to animacy alone. If the results were due to animacy alone, we would expect higher change detection accuracy for Agents in the back-to-back events with inanimate Patients as well.

We had expected a difference in the proportion of fixations to target between Agents and Patients in face-to-face events and this difference to disappear in back-to-back events. Results of the eye tracking were in line with this expectation and showed a higher proportion of fixations to the target when detecting changes to Patients compared to when detecting changes to Agents but only in the face-to-face events with inanimate Patients. While this asymmetry was consistent for both time windows in this condition (0-1000ms and 1000-2000ms), no difference was observed either in animate Patient or back-to-back conditions.

The eye gaze patterns can be interpreted in two ways. One way to interpret them could be more looks indicating higher salience. If this were true, our results would go against previous literature. As reviewed above, we would expect agency and/or animacy to be salient (e.g., Cohn & Paczynski, 2013; New et al., 2007) and, therefore, receive higher proportions of looks. Since our results show higher proportions of fixations to inanimate Patients, more looks at an entity must not mean paying more attention to it due to higher salience. Alternatively, keeping the animacy and agency salience in mind, a reaction-time-like explanation might be more fitting with our results and the literature. Insofar as more looks (as in longer reaction times) may indicate the need to pick up more visual information or the need to process for longer to accurately detect the changes in inanimate Patients because they are less salient. This interpretation of increased looking times as an indicator of greater cognitive processing has also been adopted in previous studies (e.g., Birulés et al., 2023; Candan et al., 2012). Furthermore, the accuracy results of our study would also support this explanation. The changes to inanimate Patients were detected less accurately than those to animate Agents, which might mean that changes to inanimate Patients were more challenging to detect and, thus, inanimate Patients are less salient. Assuming they were less salient for the participants who accurately detected the changes to them, these participants may have needed to look longer at inanimate Patients to gather more information. According to this possible explanation, the accuracy of detecting changes to animate Agents is higher because animate Agents are more salient. Thus, participants might have detected the changes to them easily and quickly, then looked at other things, which led to lower proportions of fixation to animate Agents[2]. To

[2] An exploratory examination of the timecourse of fixations towards target participant in non-causative events support this interpretation. In these events, fixations toward changing role started high and they decreased at around 1000 ms.

summarize, we believe that the higher proportion of looks towards inanimate Patients in our study might be reflecting increased cognitive effort in processing to accurately detect the changes due to the inanimate Patients' lower saliency.

Nevertheless, it is clear that there is a difference between animate and inanimate Patient conditions and this might be due to factors other than event roles. As discussed above, animacy by itself is not enough to explain the results. One could argue that the complexity of the entities may have played a role. It might be that inanimate Patients are visually more complex compared to animate entities, and thus, the participants were less accurate and needed to gather more information to detect changes to them. Similarly, inanimate objects being rare in the stimuli set might have led participants to look more at the inanimate Patients. Another possible explanation could be that changing the color of a person's clothes and changing the color of a part of an object might not be conceptual counterparts. The inanimate Patients eliciting lower accuracy and higher looks in the face-to-face events could be explained by these factors. However, no single factor, nor their combination, is enough to explain the full pattern of results. This is due to the similarity in accuracy and fixation proportions for Agents and Patients in the back-to-back events. If the asymmetry was due to one or more of these factors only, we would see similar results in the back-to-back condition as the face-to-face and back-to-back events did not differ in Patient animacy, visual complexity, rarity, or color-changing parts of the entities. Given that the difference between face-to-face and back-to-back depictions is the coherence of the events, the observed differences are best explained by the roles filled by the people or objects participating in the event.

To sum up, because we observed a difference in looks towards Agent and Patient between face-to-face and back-to-back events, our results support the idea that Agents are more salient due to their conceptual role in the event. Not only that, but since we found an asymmetry between Agents and Patients in face-to-face events with an inanimate Patient but not in those with an animate Patient, we can conclude that Patient animacy also plays a role in Agent preference.

4.2 Theoretical Implications

Taken together, our findings suggest that neither agency nor animacy alone fully accounts for the observed Agent preference, which was specific to face-to-face events with inanimate Patients. Instead, the results point to an interaction between event roles and animacy in shaping attentional priorities. This aligns with Brocard and colleagues' (2024) conclusion that "...the Agent preference is a specific mechanism combining agency with animacy." Such interplay highlights that Agent preference arises not from isolated features but from their specific combination within causative contexts.

The current study advances our understanding of Agent preference in several significant ways. First, we examined Agent preference in preschool-aged children using eye tracking, a more sensitive measure than accuracy-based methods, and demonstrated an asymmetry in their perception of Agents and Patients, corroborating prior findings (Brocard et al., 2024; Ünal et al., 2021). Second, we systematically varied Patient animacy and found that while animacy influences Agent preference, it does not account for it entirely, highlighting the complexity of this phenomenon. Third, by replicating Agent preference in Turkish-speaking children, we contribute to the cross-linguistic generalizability of Agent preference, expanding upon findings in languages with different syntactic properties (Isasi-Isasmendi et al., 2023; Ünal et al., 2021). These contributions deepen our understanding of the cognitive mechanisms influencing Agent preference and its developmental and linguistic contexts.

Our results align with the Thematic Hierarchy —at least for inanimate Patients— which predicts a parallel between linguistic and cognitive representations of event roles (Baker, 1997; Dowty, 1991; Grimshaw, 1981; Pinker, 1989; see also Rissman & Majid, 2019; Strickland, 2017). Our results demonstrate that Turkish-speaking preschoolers exhibit Agent preference in two-participant causative events. This is notable given that Turkish, a pro-drop language, leads children to mention Agents less frequently in event descriptions than English-speaking peers (Ünal et al., 2021). Despite such linguistic differences, the observed Agent preference suggests that Thematic Hierarchy patterns for Agents and Patients in language stem from cognitive foundations (Jackendoff, 1983, 1990). While broader conclusions about other event roles require investigations with relatively more complex events, this study provides methodologically robust support for homology between linguistic and cognitive saliency of the thematic roles.

We also demonstrate that Agent preference emerges early in human cognitive development, suggesting its foundational role in shaping event perception. This echoes findings from animal studies showing that it evolved earlier than humans (Brocard et al., 2024; Meewis et al., 2024). These results point to the possibility that it is an essential cognitive bias for event perception. Other than the literature reviewed above, this notion is supported by neuroimaging studies. A study using fMRI shows distinct neural encoding for Agent and Patient roles (Wang et al., 2016). Researchers trained machine-learning classifiers on fMRI data based on videos of Agent-verb-Patient interactions. The classifiers trained on one subset of the participants' data were able to identify the thematic role of an object in the subset of left-out participant data. Additionally, animate Agent action also seems to be a specific mechanism in the brain. A subregion in the right superior temporal sulcus has been found to be particularly tuned to human-Agent actions more than inanimate actions (Akbıyık et al., 2021). Thus, it may be that our brains are especially equipped for agency and animacy, the critical features in Agent preference, which is consistent with the idea that it is an essential foundation for event cognition.

Our results support a parallel between language and cognition by demonstrating an Agent preference among preschoolers as predicted by the Thematic Hierarchy. Our findings also support a broader notion that cognitive representation of events reflects linguistic structure (see Wagner & Lakusta, 2009) and connect to an influential proposal on language development (Gleitman, 1990), according to which learning the connections between verbs and event structure is essential. Indeed, by 2 years of age, toddlers can use syntactic cues —even with novel, made-up words— to map transitive sentences onto causative events (Arunachalam & Waxman, 2010; Naigles, 1990; Savage et al., 2003) and predict the roles of Agent and Patient (Noble et al., 2011). Moreover, research using an act-out paradigm demonstrated that they could also actively assign thematic roles, not just passively recognize them (Lidz et al., 2003). A similar paradigm, when used with Turkish-speaking 2- to 5-year-olds showed that children make use of syntactic and morphological information, such as the number of noun phrases and accusative case-marking, to infer the causativity of the event structure, highlighting the early connection between verb learning and event structure (Göksun et al., 2008). Further support for this notion, comes from studies of pronoun resolution, showing that children as young as 2 years old can resolve ambiguous pronouns and specifically that they exhibit a subject bias, favoring subjects over objects as referents in ambiguous contexts (Song & Fisher, 2005, 2007; Pyykkönen et al., 2009). This linguistic subject preference might reflect the cognitive bias toward Agents that is demonstrated in our study. To sum up, our findings lend support to the idea of homology between linguistic and cognitive processes by providing cognitive evidence for the possibility of a shared foundation for the representation of Agents.

4.3 Methodological Implications, Limitations, and Future Directions

This is the first time –to our knowledge– that eye tracking has been used in combination with a change detection paradigm to explore Agent preference in children. Our study highlights the benefits of using finer measurements with children this young because relying on coarse measures such as accuracy might not be enough to discover the nuances in Agent preference. Moreover, using similar eye tracking paradigms with adults could provide useful comparisons for interpreting looking-time data. Investigating this phenomenon in infants or younger toddlers could provide valuable insights into when and how this asymmetry emerges. Future studies could adapt Agent preference studies to younger age groups by using eye tracking and pupillometry measurements, enabling researchers to trace the earliest indicators of the asymmetry. Such research would deepen our understanding of the developmental trajectory and cognitive significance of this preference. Furthermore, eye tracking can be instrumental in cross-linguistic studies, for example by enabling comparisons between spoken languages and sign languages to investigate whether language modality has an effect Agent preference. Even though research so far points towards the generalizability of Agent preference across speakers of syntactically distinct languages, eye tracking data can offer a more nuanced understanding of Agent preference and reveal smaller differences that might otherwise go undetected.

One limitation of the current study is an error in the counterbalancing of lists, which resulted in an uneven distribution of back-to-back and face-to-face events across participants. Although we do not believe this error significantly impacted the results, it is crucial to replicate the paradigm with correct counterbalancing to ensure that the findings are not influenced by this imbalance. Proper counterbalancing would provide stronger confidence in the observed effects and their generalizability.

Even though controlling for Patient animacy is a key strength of our study, our results do not provide a complete picture of animacy effects, as we did not vary Agent animacy. Although existing research on Agent animacy is limited and less definitive than for Patient animacy (e.g., Brocard et al., 2024), it represents a promising direction for future work. In linguistics, animacy is known to influence both sentence production and comprehension. For instance, Mam-Maya (a language from Guatemala) requires the transitive subject to be higher in the animacy hierarchy than the object (Minkoff, 2000). In one study, speakers of Tzeltal (a Mayan language) and Dutch tended to assign human characters as the subject of their sentences even though this is not dictated by the languages themselves (i.e., participants mostly used active voice when the Agent was human, and passive voice when the Patient was human; Norcliffe et al., 2015). Among German speakers, animacy was shown to be a prominent cue for thematic role assignment (Kyriaki et al., 2021). Based on these linguistic studies and the Thematic Hierarchy formulation that suggests thematic roles in language map onto event participants in cognition, it is reasonable to assume Agent animacy might also play a role in cognition. Thus, systematically manipulating Agent animacy could shed light on its role in shaping Agent preference and provide a fuller understanding of the role animacy plays in event perception. This would help clarify whether animacy's effects are exclusive to Patients or whether Agent animacy also contributes to the observed asymmetry in Agent preference.

4.4 Conclusion

The current study investigated whether Turkish-speaking 3- to 5-year-olds show a preference for Agents over Patients in two-participant causative events. By systematically manipulating Patient animacy and using both behavioral and eye tracking measures, we revealed that children are more accurate in detecting changes to Agents than to Patients, but this asymmetry is modulated by Patient animacy. Specifically, Agent preference emerged in face-to-face events with inanimate Patients but disappeared in events with animate Patients or disrupted (back-to-back) contexts. Interestingly, eye tracking data showed higher proportions of looks to inanimate Patients, which may reflect increased cognitive effort rather than salience. This interpretation aligns with findings that changes to inanimate Patients were less accurately detected compared to changes to animate Agents. These findings highlight the interplay of event roles and animacy in shaping attentional priorities. By examining Agent preference in a non-Western, subject-drop language and with preschool-aged children, our findings contribute to the cross-linguistic and developmental generalizability of Agent preference. These findings advance our knowledge of the factors related to Agent preference in event cognition and provide insights into its developmental trajectory.

REFERENCES

Aginsky, V., University, B., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition*, *7*(1–3), 147–162. https://doi.org/10.1080/135062800394739

Akbıyık, S., Caramazza A., & Wurm, M. (2021). Shared and distinct neural representations of human-agent actions and inanimate events *Journal of Vision, 21*, 2490. https://doi.org/10.1167/jov.21.9.2490

Altmann, G. T. M. (1999). Thematic role assignment in context. *Journal of Memory and Language, 41*(1), 124-145. https://doi.org/10.1006/jmla.1999.2640

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442-446. https://doi.org/10.1016/j.cognition.2009.10.015

Baayen, R. H. (2008). *Analyzing linguistic data a practical introduction to statistics using R*. New York, NY: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, M.C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.) *Handbook of generative syntax* (pp. 73–137). Dordrecht: Kluwer.

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Birulés, J., Bosch, L., Lewkowicz, D. J., Pons, F. (2024). Time course of attention to a talker's mouth in monolingual and close-language bilingual children. *Developmental Psychology, 60*(1), 135-143. https://doi.org/10.1037/dev0001659

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language, 34*(6), 774-806. https://doi.org/10.1006/jmla.1995.1034

Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience, 27*(6), 109996. https://doi.org/10.1016/j.isci.2024.109996

Candan, A., Küntay, A. C., Yeh, Y., Cheung, H., Wagner, L., & Naigles, L. R. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, *27*(3), 205-221. https://doi.org/10.1016/j.cogdev.2011.12.001

Carlson, G. N., & Tanenhaus, M. K. (1989). Thematic roles and language comprehension. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing: Studies in theoretical psycholinguistics* (pp. 413). Dordrecht; Boston: Kluwer Academic Publishers.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology, 29*(3), 421-433. https://doi.org/10.1037/0012-1649.29.3.421

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of agents: The processing of semantic roles in visual narrative. *Cognitive Psychology, 67*(3), 73–97. https://doi.org/10 .1016/j.cogpsych.2013.07.002

Comrie, B. (2013). Alignment of case marking of full noun phrases. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures* (p. 399). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Davis, A. R. (2011). Thematic roles. In C. Maienborn, K. von Heusinger, & P. H. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 1, pp. 399–420). Berlin: Mouton de Gruyter.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*(2), 129–143. https://doi.org/10.1016/j.actpsy.2006.07.004

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*(3), 547–619. https://doi.org/10.2307/415037

Dryer, M. (2013). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures* (pp. 330-333). New York: Oxford University Press.

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science, 33*(4), 547–582. https://doi.org/10.1111/j.1551-6709.2009.01023.x

Emery, N. (2000). The eyes have it: The neuroethology, function, and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*(6), 582-604. https://doi.org/10.1016/S0149-7634(00)00025-7

Flaherty, M. E. (2014). *The emergence of argument structural devices in Nicaraguan Sign Language* (Publication No. 3628073) [Doctoral dissertation, The University of Chicago]. ProQuest Dissertations & Theses Global.

Fletcher-Watson, S., Collis, J. M., Findlay, J. M., & Leekam, S. R. (2009). The development of change blindness: Children's attentional priorities whilst viewing naturalistic scenes. *Developmental Science*, *12*(3), 438-445. https://doi.org/10.1111/j.1467-7687.2008.00784.x

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160-170. https://doi.org/10.1016/j.cognition.2008.11.010

Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proceedings of the National Academy of Sciences, 112*(37), 11732-11737.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge: MIT press.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165-193. https://doi.org/10.1016/0010-0277(95)00661-H

Gleitman, L. R. (1990). The structural sources of verb learning. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature, 391*, 279-281. https://doi.org/10.1038/34646

Golinkoff, R. M. (1975). A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly*, *11*(4), 623-659. https://doi.org/10.2307/747459

Golinkoff, R. M., & Kerr, J. L. (1975). Infants' perception of semantically defined action role changes in filmed events. *Merrill-Palmer Quarterly of Behavior and Development, 24*(1), 53–61. https://www.jstor.org/stable/23084058

Goupil, N., Hochmann, J. R., & Papeo, L. (2023). Intermodulation responses show integration of interacting bodies in a new whole. *Cortex*, *165*, 129–140. https://doi.org/10.1016/j.cortex.2023.04.013

Goupil, N., Papeo, L., & Hochmann, J.-R. (2022). Visual perception grounding of social cognition in preverbal infants. *Infancy, 27*(2), 210-231. doi:10.1111/infa.12453

Göksel, A. & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. New York: Routledge.

Göksun, T., Küntay, A. C., & Naigles, L. R. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, *35*(2), 291–323. doi:10.1017/S0305000907008471

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 73–113). Cambridge: MIT Press.

Grimshaw, J. (1981). Form, function and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition* (pp.165-82). Cambridge: MIT Press.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, *21*(2), 203–225.

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), 880–905. https://doi.org/10.1037/a0030045

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. https://doi.org/10.1016/j.cognition.2018.02.011

Haspelmath, M., & Siegmund, S. (2006). Simulating the replication of some of Greenberg's word order generalizations. *Linguistic Typology*, *10*(1), 74-82.

Hirose, Y., & Hancock, P. J. B. (2007). Equally attending but still not seeing: An eye-tracking study of change detection in own- and other-race faces. *Visual Cognition*, *15*(6), 647–660. https://doi.org/10.1080/13506280601069578

Hochhauser, M., Aran, A., & Grynszpan, O. (2022). Change blindness in adolescents with attention-deficit/hyperactivity disorder: Use of eye-tracking. *Frontiers in Psychiatry*, *13*, 770921. https://doi.org/10.3389/fpsyt.2022.770921

Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*(1–3), 213–235. https://doi.org/10.1080/135062800394775

Isasi-Isasmendi, A., Andrews, C., Flecken, M., Laka, I., Daum, M. M., Meyer, M., Bickel, B., & Sauppe, S. (2023). The agent preference in visual event apprehension. *Open Mind, 7*, 240–282. https://doi.org/10.1162/opmi_a_00083

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The timecourse of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America, 107*(25), 11163–11170. https://doi.org/10.1073/pnas.1005062107

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205-225. https://doi.org/10.1016/j.jml.2004.10.002

Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, *33*(25), 10235-10242. https://doi.org/10.1523/JNEUROSCI.0983-13.2013

Kyriaki, L., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2021). Thematic role assignment in non-default verb classes: A cross-linguistic comparison of English and German. *Glossa: A Journal of General Linguistics, 6*(1), 134. https://doi.org/10.16995/glossa.5728

Leslie, A. M., and Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

Levin, B. (2014). Semantic roles. In M. Aronoff (Ed.), *Oxford bibliographies in linguistics*. New York: Oxford University Press.

Levin, B., & Rappaport Hovav M. (2005). *Argument realization*. Cambridge, UK: Cambridge University Press.

Lidz, J., Gleitman, H., & Gleitman, L. R. (2003). Understanding how input matters: The footprint of Universal Grammar on verb learning. *Cognition, 87*(3), 151–178. https://doi.org/10.1016/S0010-0277(02)00230-5

Mauner, G., & Koenig, J.-P. (2000). Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language, 43*(1), 110-134. https://doi.org/10.1006/jmla.1999.2703

Meewis, F., Fagot, J., Claidiere, N., & Dautriche, I. (2024). Guinea baboons (Papio papio) show an agent preference in chasing interactions. 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* (p. 1674). Retrieved from https://escholarship.org/uc/item/75n3n3g0

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838-850. https://doi.org/10.1037/0012-1649.31.5.838

Minkoff, S. (2000). Animacy hierarchies and sentence processing. In A. Carnie & D. Laoghaire (Eds.), *The syntax of verb initial languages* (pp. 201-212). New York: Oxford University Press.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. doi:10.1017/S0305000900013817

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America, 104*(42), 16598–16603. https://doi.org/10.1073/pnas.0703913104

Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science*, *35*(5), 963-982. https://doi.org/10.1111/j.1551-6709.2011.01175.x

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, *30*(9), 1187–1208. https://doi.org/10.1080/23273798.2015.1006238

Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development, 5*(2), 193-207. https://doi.org/10.1016/0885-2014(90)90026-P

Padden, C., Meir, I., Sandler, W., & Aronoff, M. (2009). Against all expectations: Encoding subjects and objects in a new language. In D. Gerdts, J. Moore, & M. Polinsky (Eds.), *Hypothesis A/hypothesis B: Linguistic explorations in honor of David M. Perlmutter* (pp. 383-400). Cambridge: MIT Press.

Papeo, L., & Abassi, E. (2019). Seeing social events: The visual specialization for dyadic human-human interactions. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(7), 877–888. doi:10.1037/xhp0000646

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. P*sychological Science, 28*(3), 369-379. https://doi.org/10.1177/0956797616685769

Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J.-R. (2024). Abstract thematic roles in infants' representation of social events. *Current Biology, 34*(18), 4294–4300. https://doi.org/10.1016/j.cub.2024.07.081

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2009). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, *25*(1), 115–129. https://doi.org/10.1080/01690960902944014

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Rissman, L., & Goldin-Meadow, S. (2017). The development of causal structure without a language model. *Language Learning and Development, 13*(3), 286-299. https://doi.org/10.1080/ 15475441.2016.1254633

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, *26*, 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2661-2665. https://doi.org/10.1073/pnas.0405448102

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557-567. https://doi.org/10.1111/1467-7687.00312

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197-229). Cambridge: MIT Press.

Sha, L., Haxby, J. V., Abdi, H., Guntupalli, J. S., Oosterhof, N. N., Halchenko, Y. O., & Connolly, A. C. (2015). The animacy continuum in the human ventral vision pathway. *Journal of Cognitive Neuroscience*, *27*(4), 665-678. https://doi.org/10.1162/jocn_a_00733

Shore, D. I., Burack, J. A., Miller, D., Joseph, S., & Enns, J. T. (2006). The development of change detection. *Developmental Science*, *9*(5), 490-497. https://doi.org/10.1111/j.1467-7687.2006.00516.x

Siewierska, A. (2013). Alignment of verbal person marking. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures* (pp. 406-409). New York: Oxford University Press.

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE, 94*(11), 1948–1962. doi:10.1109/JPROC.2006.884093

Song, H., & Fisher, C. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua*, *117*(11), 1959-1987. https://doi.org/10.1016/j.lingua.2006.11.011

Song, H., & Fisher, C. (2005). Who's ‘‘she’’? Discourse structure influences preschoolers' pronoun interpretation. *Journal of Memory and Language, 52*(1), 29–57. doi:10.1016/j.jml.2004.06.012

Strickland, B. (2017). Language reflects “core” cognition: A new theory about the origin of cross-linguistic regularities. *Cognitive Science*, *41*(1), 70–101. https://doi.org/10.1111/cogs.12332

Sugita, Y. (2009). Innate face processing. *Current Opinion in Neurobiology, 19*(1), 39–44. https://doi.org/10.1016/j.conb.2009.03.001

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience*, *14*(5), 869-876. https://doi.org/10.1046/j.0953-816x.2001.01717.x

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language,* *33*(3), 285-318. https://doi.org/10.1006/jmla.1994.1014

Ünal, E., Richards, C., Trueswell, J. C., & Papafragou, A. (2021). Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science, 24*(6), e13116. https://doi.org/10.1111/desc.13116

Ünal, E., Wilson, F., Trueswell, J., & Papafragou, A. (2024). Asymmetries in encoding event roles: Evidence from language and cognition. *Cognition, 250*, 105868. https://doi.org/10.1016/j.cognition.2024.105868

Vettori, S., Odin, C., Hochmann, J., & Papeo, L. (2023). *A perceptual cue-based mechanism for automatic assignment of thematic agent and patient roles.* PsyArXiv. https://doi.org/10.31234/osf.io/ydgca

Wang, J., Cherkassky, V. L., Yang, Y., Chang, K. min K., Vargas, R., Diana, N., & Just, M. A. (2016). Identifying thematic roles from neural representations measured by functional magnetic resonance imaging. *Cognitive Neuropsychology*, *33*(3–4), 257–264. https://doi.org/10.1080/02643294.2016.1182480

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag.

Wilmer, J., Germaine, L., Chabris, C., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences of the United States of America, 107*(11), 5238–5241. https://doi.org/10.1073/pnas.0913053107

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*(3), 0058. https://doi.org/10.1038/s41562-017-0058

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Woodward, A. L. (2013). Infant foundations of intentional understanding. In M. R. Banaji, & S. A. Gelman (Eds), *Navigating the social world: What infants, children, and other species can teach us* (pp. 75-80). New York: Oxford University Press.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

APPENDICES

APPENDIX A. LIST OF EVENTS USED IN THE EXPERIMENT

Causative events		Noncausative events
Animate Patient	Inanimate Patient	
A woman kicks another woman	A woman opens the drawer	Two women argue
A woman kisses a man	A man knocks down boxes	Two women chat
A man pats a woman	A woman pushes the cupboard	A man and a woman dance
A woman tickles a man	A man pulls the table	Two men fight
A man pushes a woman	A man pulls/opens the curtain	Two men shake hands
A man punches another man	A woman plays the piano	A woman and a man wave hands

APPENDIX B. EVENT RATINGS BY CATEGORY

	Causative events		Noncausative events
	Animate Patient	Inanimate Patient	
Face-to-face	*M* = 3.75 *SD* = 1.34	*M* = 4.32 *SD* = 1.00	*M* = 4.15 *SD* = 1.16
Back-to-back	*M* = 2.56 *SD* =1.31	*M* = 1.78 *SD* = 1.06	*M* = 2.93 *SD* = 1.34

APPENDIX C. INSTRUCTIONS IN TURKISH AND ENGLISH

Instructions in the Original Turkish

Welcome

Merhaba! Benim adım **[Uygulayıcı İsmi]**! Senin adın ne? Merhaba **[çocuğun adı]**! Seninle bu oyunu oynayacağım için çok heyecanlıyım! Bu oyunu bilgisayarda oynayacağız. Sen daha önce bilgisayarda oyun oynadın mı?

[Evet derse] Ne güzel! Bakalım bu oyunu sevecek misin!

[Hayır derse] Birlikte oynayalım, bakalım bu oyunu sevecek misin!

Bu özel bir bilgisayar. Bu bilgisayar senin gözlerini görebiliyor!

Starting the Experiment

Önce bilgisayarın gözlerini gördüğünden emin olmalıyız.

[SMI'da gözler görününce] Aa bak bilgisayar gözlerini görebiliyor! **[Göstererek]** Bir elinle gözünü kapat, bak burada da söndü. Şimdi diğer gözünü kapat. Bak, bu sefer o söndü, gördün mü?

Bilgisayarın gözlerini görebilmesi için oyunu oynarken mümkün olduğunca sabit oturmaya çalış. Bu oyunda ekranda bazı resimler göreceksin. Bu resimlere dikkatlice bakmanı istiyorum. Şimdi başlayacağız.

Practice

Önce birkaç alıştırma yapacağız. Ekrandaki resimler yanıp sönecek ve resimlerdeki bir şeyin rengi değişecek. Resimler yanıp sönmeyi bitirince sana hangisinin değiştiğini soracağım. Hangi taraftaki değiştiyse o taraftaki elini kaldırarak cevap verebilirsin **[Ne yapmasını gerektiğini göster]**. Lütfen sadece resimlerin yanıp sönmesi bitince cevap ver, tamam mı?

Cevap verdikten sonra bir sonraki resme geçeceğiz. Bir sonrakine geçerken de yine mümkün olduğunca sabit oturmaya çalış.

Şimdi artı işareti ve ilk resim çıkacak.

[Uygulayıcıya not: Çocuğun rahatlayabilmesi için öncelikle biraz havadan sudan sohbet etmeliyiz. Eğer bir çocuk özellikle konuşmakta zorlanıyorsa practice'lerdeki hayvanlar ve sevdikleri diğer hayvanlar hakkında sorular sormak rahatlamasına yardımcı olabilir. Eğer çocuk rahat görünüyorsa bu kısım kısa tutulabilir.]

[Pratice sonrasında] Oyunun devamında da bunu yapmanı isteyeceğim. Sormak istediğin bir şey var mı?

Calibration

Oyuna devam etmeden önce bilgisayarın nereye baktığını bulabildiğinden emin olmak için minik bir oyun oynayacağız.

Ekranda bir artı işareti göreceksin. Eğer bilgisayar senin ona baktığını anlayabilirse artı hareket edecek. Sen baktıkça artı kaçacak! Bu artıyı gözlerinle takip et. Artı sabit kaldığı sürece ona bakmaya devam et. Anlaştık mı?

[Validation] Aynısını bir kere daha yapacağız.

Trials

Şimdi oyuna başlıyoruz! Oyunuma göz atmak ister misin?

Az önce yaptığımız gibi önce artı işareti sonra da ilk resmi göreceksin. Hazırsan başlayalım!

[17x] [Resim yanıp sönmeyi bitirince] Hangisinin rengi değişti?

[Çocuk cevap vermezse sor] Neyin rengi değişti?

[Cevap verdikten sonra] Tamam, bir sonrakine geçelim.

[Son resim yanıp sönmeyi bitirince] Hangisinin rengi değişti?

[Çocuk cevap vermezse sor] Neyin rengi değişti?

[Cevap verdikten sonra] Tamam.

Goodbye

Benimle oyunumu oynadığın için çok teşekkürler **[Çocuğun adı]**! Çok güzel oynadın! Arkadaşlarının yanına dönmek ister misin?

English translations of the instructions

Welcome

Hi! My name is **[name of the experimenter]**! What's your name? Hi **[name of the child]**! I'm very excited to play this game with you! We will play this game on a computer. Have you played a computer game before?

[If yes] Great! Let's see if you'll like this game!

[If no] Let's play together and see if you'll like this game!

This is a special computer. This computer can see your eyes!

Starting the Experiment

First we have to make sure the computer can see your eyes.

[When eyes appear on the SMI] Look the computer can see your eyes! **[Showing how to]** Close one of your eyes with one hand. Look it disappeared on the screen! Now close your other eye. Look, this time that one disappeared, see?

Sit as still as possible when playing the game so the computer can see your eyes.

In this game, you will see some pictures on the screen. I want you to look carefully at these pictures. We'll begin now.

Practice

First we will practice. The pictures on the screen will be flashing and the color of something in the picture will change. When the pictures stop flashing I will ask which one changed. You can answer by raising your hand on the changed side **[Show the child how to respond]**. Please answer only when the pictures stop flashing, okay?

We will move on to the following picture after you respond. Try to sit as still as possible when we're moving on to the next one.

A cross sign and the first picture will appear now.

[Note to the experimenter: We should chat a little so the child can relax. If a child especially has trouble talking, asking questions about the animals in the practice trials and other animals they like can help them relax. If the child looks comfortable enough this part can be kept short.]

[After practice trials] I will ask you to do this for the rest of the game. Is there anything you want to ask?

Calibration

Before we continue with the game, we will play a little game to make sure the computer can understand where you're looking.

You will see a cross on the screen. The cross will move if the computer can understand you're looking at it. It will move as you look at it! Follow this cross with your eyes. As long as it stays still, keep looking at it, okay?

[Validation] We'll do the same thing one more time.

Trials

Now we begin the game! Do you wanna check out my game? **[Get child's consent]**

Like we did before, first you will see the cross sign then the first picture. Let's start if you're ready!

[17x] [After the picture stops flashing] Which one changed color?

[Ask if the child doesn't respond] What changed color?

[After the child responds] Okay, let's see the next one.

[After the last picture stops flashing] Which one changed color?

[Ask if the child doesn't respond] What changed color?

[After the child responds] Okay.

Goodbye

Thank you for playing my game with me **[child's name]**! You played it well! Do you want to go back to your friends?