AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General


ISSN: 0096-3445

2025, Vol. 154, No. 3, 787–798
https://doi.org/10.1037/xge0001657

# A Perceptual Cue-Based Mechanism for Automatic Assignment of Thematic Agent and Patient Roles

Sofie Vettori, Catherine Odin, Jean-Rémy Hochmann, and Liuba Papeo
Centre National de la Recherche Scientifique, Institut des Sciences Cognitives Marc Jeannerod–UMR5229, Université Claude Bernard Lyon1

Understanding social events requires assigning the participating entities to roles such as agent and patient, a mental operation that is reportedly effortless. We investigated whether, in processing visual scenes, role assignment is accomplished automatically (i.e., when the task does not require it), based on visuospatial information, without requiring semantic or linguistic encoding of the stimuli. Human adults saw a series of images featuring the same male and female actors next to each other, one in an agentlike (more dynamic/leaning forward) and the other in a patientlike (static/less dynamic) posture. Participants indicated the side (left/right) of a target actor (i.e., the woman). From trial to trial, body postures changed, but the roles, defined by the type of posture, sometimes changed, sometimes not. We predicted that if participants spontaneously *saw* the actors as agent and patient, they should be slower to respond when roles switched from trial $n - 1$ to trial $n$, than when they stayed the same (role switch cost). Results confirmed this hypothesis (Experiments 1–3). A role switch cost was also found when roles were defined by another visual relational cue, the relative positioning (where one actor stands relative to another), but not when actors were presented in isolation (Experiments 4–6). These findings reveal a mechanism for automatic role assignment based on encoding of visual relational information in social (multiple-person) scenes. Since we found that roles in one trial affected the processing of the subsequent trial despite variations in postures and spatial relations, this mechanism must be one that assigns entities in a scene, to the *abstract* categories of agent and patient.

***Public Significance Statement***

This study demonstrates that individuals automatically encode multiple-person scenes in terms of who is the agent and who is the patient, leveraging visuospatial information such as people's body postures and their relative spatial positioning. Thus, a mechanism for the assignment of agentlike and patientlike roles, independent from verbal or deliberate processing of the stimuli, aids a key mental operation for understanding events.

*Keywords:* thematic roles, event semantics, scene perception, social interaction, action understanding

*Supplemental materials:* https://doi.org/10.1037/xge0001657.supp

Imagine a situation where you see a man, a dog and a biting. Representing who is biting (the agent) and who is being bitten (the patient) is key to understanding the event and thus deciding whether to call the animal control services or psychiatric workers. Assigning the thematic roles of agent and patient supports the ability to represent *who does what to whom.* Thematic roles have been extensively studied in linguistics and psycholinguistics as they define the relationship between verb phrase and noun phrases,



This article was published Online First December 12, 2024.
Karen Schloss served as action editor.
Sofie Vettori https://orcid.org/0000-0002-3190-514X
Preliminary results and the main ideas of this article have been presented at the European Conference on Visual Perception (Nijmegen, the Netherlands, August 2022) and at the workshop on concepts, actions and objects (Rovereto, Italy, May 2023).

The author has no conflicts of interest to disclose. This work was supported by a H2020 Marie Skłodowska-Curie Actions individual fellowship (MSCA-IF 101108756) to Sofie Vettori and by a European Research Council Starting Grant to Liuba Papeo (Grant THEMPO-758473). This research was funded in whole, or in part, by the European Commission (Grants MSCA-IF 101108756 and THEMPO-758473). 

Sofie Vettori played a lead role in data curation, formal analysis, investigation, methodology, visualization, and writing–original draft. Catherine Odin played a supporting role in conceptualization, data curation, formal analysis, investigation, methodology, and visualization. Jean-Rémy Hochmann played a lead role in conceptualization, methodology, and writing–original draft and played a supporting role in formal analysis and supervision. Liuba Papeo played a lead role in funding acquisition, conceptualization, supervision, methodology, visualization, writing–original draft, and writing–review and editing and played a supporting role in formal analysis and investigation.

Correspondence concerning this article should be addressed to Sofie Vettori or Liuba Papeo, Centre National de la Recherche Scientifique, Institut des Sciences Cognitives Marc Jeannerod–UMR5229, Université Claude Bernard Lyon1, 67 Boulevard Pinel, 69675 Bron, France. Email: sofievettori@gmail.com or liuba.papeo@isc.cnrs.fr

determining a sentence meaning (Dowty, 1991; Gruber, 1965; Rissman & Majid, 2019). Thematic roles are understood as abstract categories that apply to every event that instantiates them: whether a man bites a dog, a girl strokes a monkey, or a cat licks a baby, all of these very different events share a common abstract structure with *an agent acting upon a patient*.

The fact that young preverbal infants (Galazka & Nyström, 2016; Leslie & Keeble, 1987; Papeo et al., 2024) and nonhuman species (Abdai et al., 2017; Mascalzoni et al., 2010; see also Wilson et al., 2022) discriminate between two similar events in which agent and patient are swapped (e.g., *A pushes B* vs. *B pushes A*) suggests that—at least—precursors of agent and patient representations appear prior to (i.e., independently from) human language, and possibly rely on encoding of specific visuospatial relational cues. A well-studied case is that of causal launching events. In the typical scenario, one ball (A, agent) moves until it collides with another (P, patient) that then starts moving. Here, observers experience a transfer of force from A to P, as if A caused P to move (Michotte, 1963). There is evidence that the spatiotemporal contingency between the motion of the two balls, which is responsible for the representation of causality, is encoded as a basic feature in visual perception. Accordingly, causal launching events enter awareness faster than noncausal launching events (Moors et al., 2017) and yield retinotopic visual adaptation (Rolfs et al., 2013; see also Choi & Scholl, 2006; Kominsky & Scholl, 2020; Scholl & Tremoulet, 2000).

A compelling question is whether, in perceiving causality from spatiotemporal contingencies, the observer also represents that A is the agent and P is the patient of the event; in other words, the question is whether agent and patient representations are driven by the stimulus structure as automatically as the representation of causality. Suggesting an affirmative answer, Hafri et al. proposed that agent/patient role assignment in prototypical human social interaction events is automatic, profiting from postural cues that are reliably associated with agent and patient (Hafri et al., 2013). Toward this hypothesis, they showed that after observing a social interaction (e.g., punching) in which Person A was the agent and Person B was the patient, participants were slower to process a different interaction (e.g., kicking) in which B was the agent and A was the patient, than another (e.g., pushing) in which A remained the agent and B, the patient (Hafri et al., 2018). This suggested that participants automatically and implicitly assigned the actors to the abstract categories of agent and patient and paid a cost when roles from trial $n - 1$ to trial $n$ conflicted. The authors argued for the automaticity of this process, emphasizing that the so-called *role switch cost* (RSC) was found even though the task required the participants to attend to features of the stimuli (i.e., shirt color or gender), other than the roles of the actors in the event.

In Hafri et al. (2018), stimuli carried a mixture of perceptual cues of interaction depicting familiar, readily identifiable and easy to label social events. Thus, it is possible that participants first categorized the actions and then, based on this semantic analysis, they assigned the roles to the actors. The role of action understanding –that is, whether automatic role assignment can occur even if one cannot readily identify the action—as well as the perceptual cues that drive automatic role assignment during the encoding of visual social events remain unclear.

Building on Hafri et al.'s work, we set the conditions to study the basic information on which role assignment can be implemented; particularly, we carried out a systematic investigation of the perceptual cues to automatic role assignment in processing visual scenes that may, but also may not, depict meaningful, familiar social interactions.

Relating role assignment to visuospatial information processing, independently of the semantic processing of actions, will support the hypothesis that individuals automatically extract rich information on event structures from visual properties of the stimuli. In this endeavor, we began to investigate the boundary conditions for such visually driven role assignment –that is, in which visual contexts and by virtue of which visual properties, social entities are readily seen as agentlike or patientlike.

Adapting the above paradigm from Hafri et al. (2018), we presented participants with sequences of images featuring a man and a woman, face-to-face. The two were in different postures, which changed from trial to trial. Body postures were created and assigned to one of two types, agentlike or patientlike, based on the previous work showing that a person in a dynamic posture, leaning forward, with outstretched extremities, is readily recognized as the agent of an event, while a person in a less dynamic or static posture, with limbs resting close to or along the body, is more likely to be recognized as the patient (Hafri et al., 2013). In each dyad, one actor was always in an agentlike posture and the other, in a patientlike posture, but agentlike and patientlike postures were paired randomly, so that the scene did not necessarily depict any obvious, familiar, meaningful interactions (as confirmed by independent ratings). Participants were only asked to report the side (left/right) of a target (e.g., the woman). We measured the role switch cost, that is, the delay in the response time to a trial with a different (vs. the same) role assignment than the previous trial. This effect would indicate that participants automatically assign the two actors in a scene to the abstract categories of agent and patient, so that when roles switch, the assignment process starts over, yielding a delay in the response. Here, without explicit requirement for role assignment and without always an obvious semantic categorization of actions in the event, role assignment would mainly rely on visual information such as body posture. In subsequent experiments, we investigated whether this effect, initially tested with spatially close, face-to-face dyads (Experiments 1–2), could also be found with back-to-back dyads (Experiments 3–4), with bodies in agentlike or patientlike postures presented in isolation (Experiments 5–6), and when roles were not cued as much by postures as by the relative positioning of the two characters, that is, where one stands in relation to another (Experiment 7).

# Method

## Transparency and Openness Statement

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. Data, material, and analysis code can be found at https://osf.io/3w59v. This study was not preregistered.

## Experiment 1: Face-to-Face Dyads

### *Participants*

Twenty-four participants (19 participants reported their gender as male, five as female; $M_{\text{age}} = 24.8$ years, min = 19, max = 35) were recruited and tested on the online platform at https://www.testable

.org/ (Rezlescu et al., 2020). The sample size was the same as in Hafri et al. (2018) and was larger than the minimal sample size ($n$ = 14) required to obtain a RSC with effect size comparable to Hafri et al. (2018), Experiment 1a: Cohen's $d = 1.07$; two-tailed; $\alpha = .05$; $\beta = .95$; G*Power 3.1. Participants gave informed consent prior to the study and were paid $1 for participation. The local ethics committee (CPP Sud Est II) approved this and the following experiments.

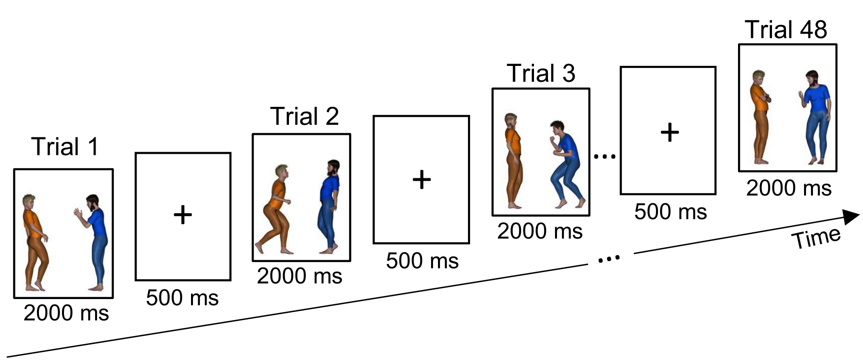
### Stimuli

Renderings of human bodies were created with Daz3D (Daz Productions, Salt Lake City, Utah) and the image-processing toolbox of MATLAB (The MathWorks, Natick, Massachusetts) and represented four models (a male body dressed in blue, the same male body dressed in orange, a female body dressed in blue, and the same female body dressed in orange) in 48 different body postures, 24 agentlike and 24 patientlike. Following Hafri et al. (2013), agentlike postures were fairly dynamic, leaning forward with outstretched limbs; patientlike postures were less dynamic or static, standing still or leaning back, with limbs close to or along the body. Postures did not necessarily represent familiar, recognizable actions. For each posture, we created a mirror-flipped version, yielding a total of 384 individual body postures (four models, 48 postures, original and mirror flipped). By combining the 384 single bodies, 192 facing dyads were created, each featuring one male body and one female body, facing each other, one in an agentlike and the other in a patientlike posture (Figure 1). The side (50% left/50% right), color (50% orange/50% blue), and posture (50% agentlike/50% patient-like) of the man and of the woman were counterbalanced across stimuli. Agentlike and patientlike postures were randomly paired with no consideration of the meaningfulness/familiarity of the resulting interaction. Meaningfulness of the stimuli was measured in an online rating study with external participants, who rated the stimuli on a 1–7 Likert scale (from *meaningless* to *very meaningful*). Mean meaningfulness score was 4.22 ($SD$ = 1.60, range 1–5.3), showing that stimuli were not judged as very meaningful, although they could not be said to be meaningless either. The stimuli and a detailed report of this rating study are provided as Supplemental Material. Each stimulus subtended a visual angle of 11° at a distance of 60 cm (image size: 379 × 500 pixels), with two bodies at an equal distance (25 pixels) from the center of the image, considering the point that was the closest to the center.

### Procedure

Participants were tested on the online platform at https://www.testable.org/ (Rezlescu et al., 2020). Prior to the experiment, participants were asked to sit 60 cm away from the screen (about the length of an arm), align their eyes with the center of the screen, avoid moving for the duration of the experiment, and follow instructions for calibrating the screen and recording its frame rate. Next, they were instructed to place the left and the right index fingers on the [e] and the [o] key of their keyboard, respectively, and to read the task instructions. For each trial, they had to report the location of the target, as fast and as accurately as possible. The target (the man or the woman) was indicated at the beginning of each block. It was randomly selected in the first block and then counterbalanced across blocks to have 50% of blocks with the woman as target and 50% of blocks with the man as target. The instructions did not mention roles or social interaction. Each trial began with a blank screen showing a central fixation cross (500 ms), followed by the stimulus (the dyad), which remained on the screen until the participant responded, or up to 2000 ms (Figure 1). The experiment consisted of 10 blocks of

**Figure 1**
*Illustration of the Paradigm Used in Experiments 1 (and 4)*




*Note*. Each stimulus-image was preceded by a central fixation screen (500 ms) and was shown up to 2000 ms. In each trial, participants had to press a button to indicate the position of the target actor (left or right). From a trial to the subsequent one the two actors could switch type of posture (e.g., Trial 1: woman is the agent, man is the patient; Trial 2: man is the agent, woman is the patient) or maintain the same type of posture (e.g., Trial 2: man is the agent, woman is the patient; Trial 3: man is the agent, woman is the patient). There could also be a switch of side (left/right) and of color of clothes. The same paradigm was used in the following experiments but with variations in the relative positioning of the two actors (back-to-back rather than face-to-face in Experiments 3–4; face-to-back in Experiment 7) or in the number of actors and task instructions (Experiments 5–6). See the online article for the color version of this figure.

48 trials each, for a total of 480 trials. Stimuli were shown in a random order, and each stimulus was shown two to four times during the experiment. The posture (agentlike/patientlike), the side (left/right) and the color (orange/blue) of the target was counterbalanced and changed randomly across trials within a block. Thus, a trial could differ from the previous trial for the posture of the target (agentlike/patientlike), its side (left/right) and/or the color of the clothes (blue/orange). The change (or switch) of interest was the type of posture, which defined the role (trial $n-1$: the target was agentlike; therefore, the other body was patientlike; trial $n$: the target was patientlike, therefore, the other body was agentlike). A role switch (relative to trial $n-1$) occurred in 50% of the trials. The color change was introduced only to increase the variability of the stimuli and ensure that the gender did not coincide with a color. Since the target was in orange in 50% of trials, and in blue in the remaining trials, participants could not perform the gender-search task by attending to the color. The experiment lasted ∼13 min.

### *Data Analysis*

We measured the RSC, as the reaction times (RTs) difference between two consecutive trials, distinguishing between the condition in which two consecutive trials had the same role assignment versus the condition in which the second trial introduced a role switch. With the same schema, we also measured the effect of side switch, that is, the difference in RTs between two consecutive trials distinguishing between the condition in which two consecutive trials had the target on the same side versus the condition in which they showed the target on different sides. Thus, data were analyzed as a function of role switch (role switch vs. no role switch trials) and side switch (side switch vs. no side switch). All participants were included in the analysis as they all performed within 2.5 *SD* from the group mean for both accuracy rates and RTs ($M = 568$ ms, $SD = 141$ ms). For each participant, we discarded trials in which the response was incorrect or too fast (<200 ms). Moreover, as the dependent variable considered the difference between two consecutive trials, we also discarded the trials that followed a trial with an incorrect response. Finally, we discarded trials with RT ± 2.5 *SD* from the individual condition's mean. Following this procedure for data cleaning, a total of 12.52% of trials was removed from the RT analysis ($M = 15.9\%$, $SD = 7.05$ per participant). Details on data cleaning for this and the following experiments are provided in Supplemental Table S1. Clean RTs were analyzed with linear mixed effects modeling (LMM) using lme4 (Version 1.1.27.1; Bates et al., 2015) and afex (Version 1.3; Singmann et al., 2020) with R (Version 4.0.2; R Core Team, 2021). *P* values for fixed effect model terms (main effects and interactions) were estimated with the Kenward–Roger method and Type 3 tests obtained by comparing a model in which only the corresponding effect is missing with the full model containing all effects. The following factors were included in models as fixed effects: role switch (role switch vs. no role switch) and side switch (side switch vs. no side switch). For all experiments, we tried to fit the structure of maximal complexity (Barr, 2013), including correlated random intercepts and slopes of role switch and side switch by participants and by stimulus and by block. When models failed to converge (due to over-parameterization; e.g., Bates et al., 2018; Matuschek et al., 2017), we progressively simplified the random-effect structure by first removing correlations among random terms (i.e., fitting zero-correlation models), then by removing random slopes (or intercepts) associated with the smallest amount of variance (often corresponding to 0). Full details of random effects structures of models are available as the Supplemental Materials. Follow-up tests and estimated marginal means from LMMs were calculated using the R package emmeans (Lenth et al., 2019). Effects sizes were calculated with Cohen's *d*, and bayes factors were calculated using the BayesFactor package (Morey et al., 2011). Following Hafri et al. (2018), we expected effects in RTs. Data, material, and analysis code can be found at https://osf.io/3w59v.

## Results

Accuracy was overall high ($M = 96.9\%$, $SD = 4.18\%$). RTs showed a main effect of role switch, $F(1, 8.69) = 9.77$, $p = .013$, $BF = 19.81$, $d = 0.71$, and a main effect of side switch, $F(1, 27.80) = 8.13$, $p = .008$, $BF = 7.20$, $d = .61$, but no interaction, $F(1, 29.55) = .00$, $p = .946$ (Figures 2A and 3). Participants were faster when the target switched side than when the target stayed on the same side ($M_{\mathrm{Diff}} = -16.20$, $SE = 5.39$). Moreover, we found an RSC, showing that participants were slower in role switch trials than in no role switch trials (RSC: $M_{\mathrm{Diff}} = 8.90$ ms, $SE = 2.54$).

The RSC implied that, when instructed to search for a target based on gender, participants also automatically assigned the actors to agentlike and patientlike roles: across visually different scenes, the agent role was consistently attributed to the actor in the more dynamic and leaning-forward posture, and the patient role was consistently attributed to the one in the less dynamic posture.

## Experiment 2: The Effect of Meaning

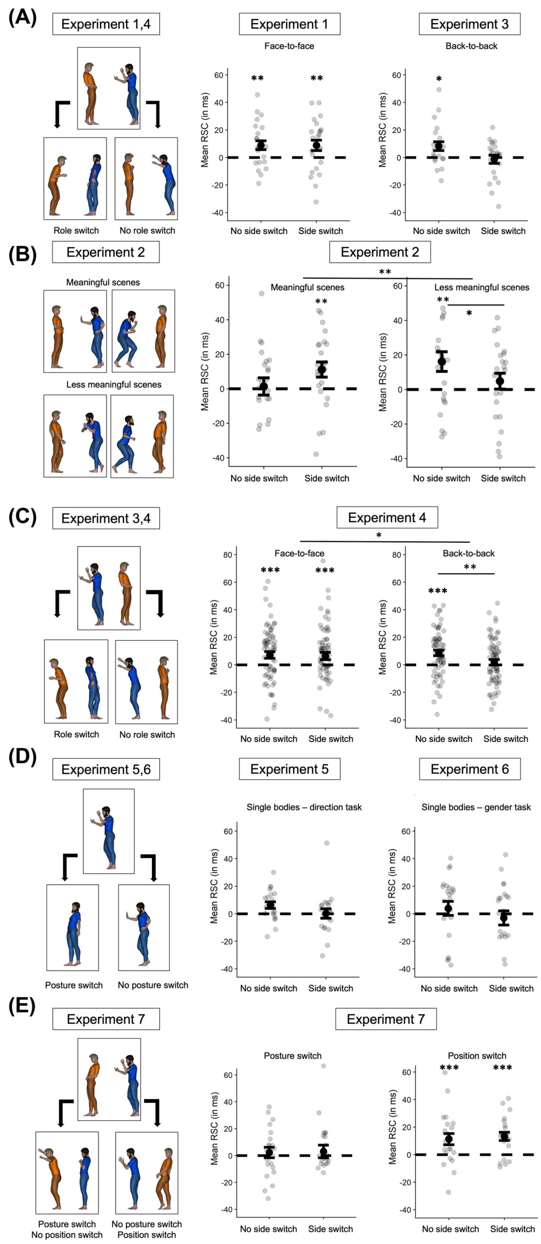
A goal of the present study was to test automatic role assignment independently of action understanding. To this end, we selected body postures with scant consideration of their relationship with readily identifiable actions. These postures gave rise to dyadic scenes that did not necessarily depict familiar, meaningful interactions, as confirmed by independent ratings. However, ratings also showed that stimuli were not so meaningless after all (mean meaningfulness score: 4.22 ± 1.60 *SD*, range 1–5.3; see above). Moreover, people have a tendency to assign meaning to virtually *anything*, and they do so even more when seeing people in configurations (e.g., face-to-face) that cue a meaningful relationship (i.e., an interaction; see Paparella & Papeo, 2022). Experiment 2 investigated to what extent automatic role assignment depended on the meaningfulness of the stimuli, that is, on how much stimuli depicted actions and events that could be readily identified. We reasoned that, if role assignment relies on the semantic coding of the stimuli (i.e., action understanding) more than, or as much as, coding of visuospatial information, a larger RSC should be found for more meaningful relative to less meaningful stimuli.

### *Participants*

Twenty-four new participants (18 participants identified their gender as male, six participants identified their gender as female; $M_{\mathrm{age}} = 25.5$ years, min = 18, max = 30) were recruited and tested on the https://www.testable.org/. They gave informed consent and were compensated $1 for their participation.

**Figure 2**

*Illustration of Stimuli and Results for Each Experiment*




*Note.* (A)–(C) Left: Examples of role switch and no role switch conditions with face-to-face dyads (Experiments 1, 4), and back-to-back dyads (Experiments 3, 4). Right: Results of Experiments 1, 3, and 4. Mean RSC (in ms) was computed as the difference between the RT for trials with role switch and the RT for trials with no role switch. Each dot represents a participant. (B) Left: Examples of stimuli in the meaningful block and the less meaningful blocks for Experiment 2. Right: Results of Experiment 2. (D) Left: Examples of posture switch and no posture switch conditions in the single-body stimuli (Experiments 5–6). Right: Results of Experiments 5–6. (E) Left: Examples of role switch and no role switch conditions with face-to-back dyads (Experiment 7), where role assignment could be biased by the postures (agentlike or patientlike) or by the positioning (behind or ahead). Right: Results of Experiment 7. RSC = role switch cost; RT = reaction time. See the online article for the color version of this figure.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

### Stimuli

Based on the rating study described above (see Experiment 1; see also Experiment S1 in Supplemental Material), we selected the eight face-to-face dyads with the highest rate of meaningfulness (MF+ set) and the eight with the lowest rate (MF− set). We verified that meaningfulness statistically differed between the two sets, with a *t* test (one-tail). Rates of the MF+ set were significantly higher ($M = 4.94$, $SD = 0.34$) than those of the MF− set ($M = 3.49$, $SD = 0.41$) ($t_{13} = 7.745$, $p < .001$).

### Procedure

Task and instructions were identical to Experiment 1. The procedure was identical to the one in Experiment 1 except that here, we included five blocks with MF+ stimuli and five with MF− stimuli (Figure 2C). The order of the blocks was pseudorandomized such that for half of the participants, the experiment started with a MF+ block, and for the other half of the participants the experiment started with a MF− block. Within each block, stimuli (MF+ or MF−) were presented randomly.

### Data Analysis

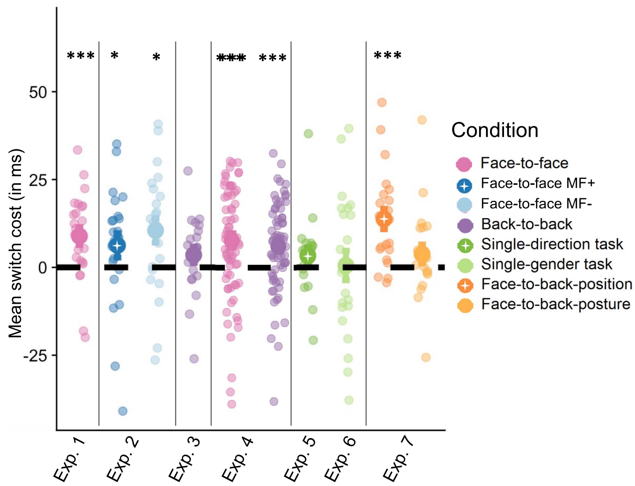
After data cleaning (see Supplemental Table S1), RTs were analyzed with an LMM with role switch (role switch vs. no role switch) and side switch (side switch vs. no side switch) and meaningfulness (MF+ vs. MF−) as fixed factors. The final random-effect structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes of role switch and side switch by participant, stimulus and block.

### Results

Accuracy was overall high ($M = 94.7\%$, $SD = 4.02\%$). Results showed an effect of role switch, $F(1, 9697.82) = 14.94$, $p = .00011$, $BF = 5.62$, $d = .59$, and an effect of side switch, $F(1, 9620.05) = 39.41$, $p < .0001$, $BF = 9.45$, $d = .64$ (Figures 2C and 3). Participants were slower when there was a role switch from trial $n - 1$ to Trial 1 (RSC: $M_{\text{Diff}} = 8.36$, $SE = 2.87$), and when there was a side switch ($M_{\text{Diff}} = -13.60$, $SE = 4.48$). There was no main effect of meaningfulness, $F(1, 25.44) = 3.49$, $p = .073$, $BF = .74$, $d = .34$, no significant interaction between meaningfulness and role switch, $F(1, 9699.15) = .39$, $p = .53$, and no significant interaction between meaningfulness and side switch, $F(1, 9620.24) = .078$, $p = .78$. The interaction between role switch and side switch was not significant, $F(1, 9627.19) = .024$, $p = .88$. However, results showed a significant three-way interaction between role switch, side switch, and meaningfulness, $F(1, 9621.92) = 7.47$, $p = .0063$.

To explain the three-way interaction, RTs for the MF+ and the MF− conditions were analyzed separately, as function of role switch and side switch. For the MF+ condition, there was a main effect of role switch, $F(1, 22.12) = 5.15$, $p = .033$, $BF = 0.88$, $d = .37$, and a main effect of side switch, $F(1, 33.42) = 4.59$, $p = .040$, $BF = 3.19$, $d = .53$. Participants were slower when there was a role switch (RSC: $M_{\text{Diff}} = 6.24$, $SE = 3.44$) and a side switch ($M_{\text{Diff}} = -13.10$, $SE = 5.19$). The interaction was not significant, $F(1, 53.81) = 1.70$, $p = .20$. For the MF− condition, there was a main effect of role switch, $F(1, 31.34) = 5.86$, $p = .021$, $BF = 7.94$, $d = .62$, and a main

**Figure 3**
*Summary Plot*




*Note.* In pink, the mean RSC, RSC (RT role switch—RT no role switch) is shown for face-to-face dyads (Experiments 1, 4). In blue, the mean RSC (RT role switch—RT no role switch) is shown for the meaningful (MF+) and less meaningful (MF−) face-to-face dyads (Experiment 2). In purple, the mean RSC (RT role switch—RT no role switch) is shown for back-to-back dyads (Experiments 3, 4). In green, the mean posture switch cost (RT posture switch—RT no role switch) is shown for single actors (Experiments 5, 6). Dark green = direction task, light green = gender task. In dark orange, the mean switch cost based on position (RT position switch—RT no position switch) is shown for face-to-back dyads (Experiment 7). In light orange, the mean switch cost based on posture (RT posture switch—RT no posture switch) is shown for face-to-back dyads (Experiment 7). RSC = role switch cost; RT = reaction time. See the online article for the color version of this figure.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

effect of side switch, $F(1, 23) = 7.24$, $p = .013$, $BF = 4.17$, $d = .56$. Again, participants were slower when there was a role switch (RSC: $M_{\text{Diff}} = 10.5$, $SE = 3.43$) and a side switch ($M_{\text{Diff}} = -14.10$, $SE = 5.23$). The interaction between role switch and side switch was significant, $F(1, 4382.52) = 4.60$, $p = .032$, revealing that the RSC was significant only in the absence of a concurrent side witch ($t_{23} = 2.84$, $p = .0092$), but not when in addition to the role, the target also changed side ($t_{23} = 1.05$, $p = .31$).

In summary, Experiment 2 showed a significant RSC for both types of dyads (MF+ and MF−), suggesting that the low meaningfulness of the stimuli did not prevent automatic role assignment. A post hoc analysis on the data of Experiment 1 confirmed this circumstance. Briefly, we analyzed the RTs of Experiment 1 including the type of transition (from MF+ to MF+ trial and from MF− to MF− trial) as factor. We reasoned that, if automatic role assignment critically depended on how easy actions could be categorized, RSC should be larger in the MF+ than MF− transition. However, no difference in RSC was found between the two, $F(1, 4641.4) = 1.45$, $p = .23$.

At the same time, the results of Experiment 2 showed that, in the MF− condition, the RSC was significant only when there was no side switch. The sensitivity of the RSC to a side switch suggests that role assignment may be more susceptible to changes in perceptual properties of the stimuli (i.e., side switch) for the less meaningful stimuli. Thus, meaning could further promote role assignment, although a clear unambiguous meaning does not seem to be a necessary condition to observe automatic RSC: across visually and semantically different scenes, individuals automatically assigned the actors to role categories, attributing an agentlike role to the actor in the more dynamic and leaning-forward posture, and a patientlike role to the actor in the less dynamic posture. In Experiment 3, we asked whether this automatic posture-based agent/patient role assignment could also occur for two actors that are close together in space but not in the prototypical face-to-face configuration for interaction.

## Experiment 3: Back-to-Back Dyads

Facilitating mutual perceptual access, and therefore a number of basic social behaviors (e.g., communication, gaze following, and shared attention), the face-to-face positioning of two people is a reliable cue of interaction (Papeo, 2020). Experiments 1–2 showed that two people presented face-to-face are automatically *seen* as agentlike and patientlike, based on their body postures. Here, we investigated whether such cue-based role assignment generalized to a condition where the two bodies appeared in the same visual scene, close together in space, but back-to-back, a positioning that implies no interaction or is, at least, less prototypical of an interaction.

### *Participants*

Twenty-four new participants were recruited (16 reported their gender as male, eight as female; $M_{\text{age}} = 24.9$ years, min = 19, max = 35, $SD = 4.6$) and tested online, on https://www.testable.org/. The samples size matched that of Experiments 1–2. Participants gave informed consent before participation and were compensated with $1.

### *Stimuli and Procedure*

Stimuli and procedures were identical to Experiment 1 except that bodies in the dyads were positioned back-to-back (Figure 2B). Back-to-back dyads were obtained by merely flipping the bodies in the face-to-face dyads of Experiment 1.

### *Data Analysis*

After data cleaning (see Supplemental Table S1), RTs were analyzed with an LMM with role switch (role switch vs. no role switch) and side switch (side switch vs. no side switch) as fixed factors. The final random-effect structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes of role switch and side switch by participant, stimulus and block.

### *Results*

Accuracy was overall high ($M = 95.3\%$, $SD = 2.99\%$). Results on RTs showed no significant main effect of role switch, $F(1, 169.40) = 1.31$, $p = .25$, $BF = 0.61$, $d = 0.33$ (Figures 2A and 3), and no main effect of side switch, $F(1, 26.55) = 0.03$, $p = .872$, $BF = 0.22$, $d = 0.02$. There was a trend for the interaction between role switch and side switch, $F(1, 162.71) = 3.73$, $p = .055$, showing a role switch cost when there was no side switch ($M_{\text{Diff}} = 8.27$, $SE = 3.25$, $t_{21} = 2.549$, $p = .019$) but not when there was a side switch ($M_{\text{Diff}} = -1.26$, $SE = 3.00$, $t_{21} = .420$, $p = .679$).

## Experiment 4: Face-to-Face and Back-to-Back

In Experiment 4, we replicated Experiments 1 and 3 with a larger sample size, to afford the power for comparing performance between groups tested with face-to-face versus back-to-back dyads.

### Participants

A total of 161 participants were tested in Experiment 4, 80 with face-to-face dyads (53 participants identified their gender as male, 27 as female; $M_{\text{age}} = 25.5$ years, min = 18, max = 35, $SD = 4.9$ years) and 81 with back-to-back dyads (49 identified their gender as male, 32 as female; $M_{\text{age}} = 26.5$ years, min = 18, max = 35, $SD = 5.1$ years). Data from seven participants in the former group and from nine participants in the latter were discarded from the final analysis, which resolved into two groups of 73 and 72, respectively. This sample size was higher than the sample size ($n = 64$) required to observed a between-group difference of medium effect size (comparison between RSC in Experiments 2 and 3 in Hafri et al., 2018, Cohen's $d = 0.5$; 80% power; $\alpha = 0.05$; GPower 3.1.9.7). Participants were recruited and tested online, at https://www.testable.org/, gave informed consent before starting the experiment, and received $1 for participation.

### Stimuli and Procedure

Stimuli and procedures were identical to Experiments 1 (face-to-face dyads) and 3 (back-to-back dyads). Body positioning was manipulated between groups. Participants were randomly assigned to the face-to-face or back-to-back condition.

### Data Analysis

We followed the same data trimming procedure as in Experiments 1–3 (Supplemental Table S1). RTs were analyzed separately for each group, in two LMMs with role switch (role switch vs. no role switch) and side switch (side switch vs. no side switch), and then together in a mixed effects model with factors: group, side switch, and role switch. The final random effects structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes by participant, stimulus, and block.

### Results

Accuracy was overall high ($M = 92.3\%$, $SD = 7.19\%$).

**Face-to-Face Group.** Replicating Experiment 1, we found a main effect of role switch, $F(1, 24.49) = 9.22$, $p = .006$, $BF = 70.00$, $d = .44$, with slower RTs in role switch versus no role switch trials (RSC, $M_{\text{Diff}} = 6.86$, $SE = 1.80$; Figures 2C and 3). There was no effect of side switch, $F(1, 55.61) = 3.45$, $p = .068$, $BF = 0.40$, $d = .18$, and no interaction, $F(1, 118.75) = .02$, $p = .88$.

**Back-to-Back Group.** There was a main effect of role switch, $F(1, 23649.72) = 17.54$, $p < .001$, with slower RTs in role switch versus no role switch trials (RSC, $M_{\text{Diff}} = 5.29$, $SE = 1.33$, $BF = 160.84$, $d = .48$; Figures 2C and 3). There was no effect of side switch, $F(1, 91.91) = 1.95$, $p = .17$, $BF = 0.37$, $d = .18$, but, like in Experiment 3, a significant interaction between the two factors, $F(1, 24433.60) = 8.63$, $p = .003$. The interaction revealed a significant RSC when there was no side switch (RSC, $M_{\text{Diff}} = 8.67$, $SE = 1.97$, $t_{70} = 4.93$, $p < .0001$, $BF = 494.92$, $d = .52$), but not when the target switched side ($M_{\text{Diff}} = 1.91$, $SE = 2.03$, $t_{70} = 0.95$, $p = .35$, $BF = 0.20$, $d = .11$).

**Between-Group Comparison.** Results showed a main effect of role switch, $F(1, 352.60) = 23.28$, $p < .001$ ($M_{\text{Diff}} = 6.07$, $SE = 1.13$, $BF = 14559.63$, $d = 0.5$), a main effect of group, $F(1, 148.36) = 6.28$, $p = .013$, $BF = 2.98$, $d = .41$, but no main effect of side switch, $F(1, 193.29) = .13$, $p = .72$, $BF = 0.11$, $d = 0.04$. The interaction was significant between group and side switch, $F(1, 193.29) = 5.47$, $p = .020$, but not between role switch and side switch, $F(1, 143) = 2.62$, $p = .11$, $\eta_p^2 = .018$, or between role switch and group, $F(1, 352.51) = 0.53$, $p = .47$. All main effects and interactions were qualified by the three-way interaction between role switch, side switch, and group, $F(1, 37708.69) = 5.03$, $p = .025$ (Figure 2B). This interaction reflected the fact that the RSC was consistently found in the face-to-face group, irrespective of whether there was a side switch or not (Role Switch × Side Switch: $F(1, 118.75) = 0.02$, $p = .882$), while it was less consistent in the back-to-back group. In this group, an RSC was found in no side switch but not in side switch trials (Role Switch × Side Switch: $F(1, 24433.60) = 8.63$, $p = .003$).

In sum, results of Experiments 1–4 showed that participants automatically assigned agent or patient roles to two actors seen next to each other, whether the two were face-to-face, as in a prototypical interaction, or not. Role assignment, however, appeared less consistent in the back-to-back condition: In this condition, an RSC was only significant when there was no side switch.

We note that a trend for a side switch by role switch interaction was also observed in Experiment 3: also there, with back-to-back dyads, an RSC occurred in trials with no side switch, but not in trials with side switch. One possibility is that, in the back-to-back condition, where bodies did not seem to interact, and therefore an event structure could be faint, the representation of roles was more tied to specific perceptual information (i.e., a specific type of posture in a specific location/side) and failed to generalize to a similar posture presented on a different side of the scene. In sum, role representations accessed in back-to-back dyads could be less general (or generalized) and more susceptible to changes in perceptual properties of the stimuli, such as side. A similar interpretation may apply to the processing of stimuli with low meaningfulness. As reported above, a significant side switch by role switch interaction was found in the MF− (but not in the MF+) condition of Experiment 2, implying that roles generalized to postures in the same or a different spatial position (i.e., side), in the context of readily identifiable actions (MF+), but not in the context of semantically ambiguous actions.

Notwithstanding variations in the generalization of roles to the side switch condition, all experiments and conditions so far (face-to-face MF+ and MF− and back-to-back) have consistently showed an automatic role assignment in the processing of relational stimuli, such as visual social scenes involving two people spatially close to one another.

## Experiment 5: Single Bodies

The automatic role assignment documented in Experiments 1–4 implies the generation of a relational, syntacticlike representational scheme with agent and patient roles, during visual scene perception.

However, since we have selected body postures so to create a visual categorical distinction, the above effects of role switch could just reflect visual discrimination of individual actors based on body postures. In this interpretation, individuals would assign the target-actor to either type of body postures and would be surprised when the same actor switches to the other type of posture (e.g., from dynamic/agentlike in trial $n-1$ to static/patientlike in trial $n$). In this case, performance would reflect the encoding of single-body postures, rather than a relational process. If so, a cost should also be observed when stimuli depict a single body that from trial $n-1$ to trial $n$ switches from one type of posture to another. To address this, we presented images of single bodies in an agentlike or a patientlike posture (Figure 2D) and measured the cost (in RTs) of the posture switching from one type to another, between consecutive trials.

### Participants

Twenty-four new participants were recruited (19 participants identified their gender as male, five as female; $M_{age} = 25.17$ years, $SD = 5.33$ years, min = 19, max = 35) on https://www.testable.org/. In this absence of previous data with this new task, we chose the sample size of Experiments 1–2. Participants gave informed consent and were paid $1 for participation.

### Stimuli and Procedure

Stimuli consisted of the 384 single bodies that were used to create the 192 dyads in Experiments 1–4. In each stimulus-image, a single body was presented at the center of the screen. The body could be a man (50% of stimuli) or a woman, in blue (50% of stimuli) or orange clothes, in an agentlike (50% of stimuli) or a patientlike posture, oriented leftward (50% of stimuli) or rightward. Like in previous experiments, during the experiment, stimuli were presented in 10 blocks of 48 trials each, for a total of 480 trials. Each stimulus was shown one to four times during the experiment, but never consecutively and never in the same block. Each trial started with a blank screen with a central fixation cross (500 ms), followed by the stimulus (the single body), which remained on the screen until the participant responded, or up to 2000 ms. Participants were instructed to perform a direction-judgment task, requiring to indicate the body orientation (leftward/rightward). The gender (male/female), posture (agentlike/patientlike), direction (left/right) and color (orange/blue) of the body varied randomly across trials within a block. Thus, a trial could differ from the previous one for one or more of the following features: gender, posture, direction, and color of the clothes. The switch of interest was the posture, which was orthogonal to the task-relevant information (i.e., body orientation). This new task allowed maintaining a left/right response for each trial, as in previous Experiments 1–4. The experiment lasted ∼13 min.

### Data Analysis

We followed the same data-cleaning procedure as in previous experiments (Supplemental Table S1). The remaining values were analyzed with LMMs with posture switch (posture switch vs. no posture switch) and direction switch (direction switch vs. no direction switch) as fixed effects. The final random-effect structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes by participant, stimulus, and block.

### Results

Accuracy was overall high ($M = 94.4\%$, $SD = 4.76\%$). Results on RTs show no effect of posture switch ($M_{Diff} = 3.32$ ms, $SE = 2.25$), $F(1, 18.09) = 1.83$, $p = .19$, $BF = 0.52$, $d = .30$, an effect of direction switch, $F(1, 24.50) = 9.15$, $p = .006$, $BF = 15.49$, $d = .75$, and no interaction between the two, $F(1, 5985.77) = 3.37$, $p = .066$ (Figures 2D and 3). The effect of direction switch showed that participants were slower to respond, when the body switched from one direction (trial $n-1$) to another (trial $n$). This makes sense as direction was the task-relevant dimension in this experiment. To further investigate the (null) effect of posture switch, we computed the Bayes factor using a Bayesian $t$ test on the RT difference between role switch trials and no role switch trials. A Bayes factor of 0.52 provided anecdotal evidence for the absence of a posture switch cost. The following experiment was designed to further assess the effect of posture switch.

## Experiment 6: Single Bodies

In Experiment 5, we found no evidence for a posture switch cost when agentlike and patientlike body postures were presented in isolation. We sought to provide additional support to this finding with a task that, like in Experiments 1–4, required participants to attend to the gender of the actors (gender discrimination task).

### Participants

Twenty-four participants were recruited (17 participants identified their gender as male, seven as female; $M_{age} = 25.96$ years, $SD = 4.35$ years, min = 20, max = 33) at https://www.testable.org/. Participants gave informed consent before participation and were compensated $1 for their participation.

### Stimuli and Procedure

Stimuli and procedures were identical to Experiment 5 except for the task instructions: Here, participants had to indicate the gender of the target-actor (male or female) by pressing one of two keys, [e] and [o]. The response-key mapping was counterbalanced across participants. The experiment lasted ∼13 min.

### Data Analysis

We followed the same data-cleaning procedure as in previous experiments (Supplemental Table S1). The remaining values were analyzed with an LMM with posture switch and direction switch as factors. The final random-effect structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes by participant, stimulus, and block.

### Results

Accuracy was overall high ($M = 93.3\%$, $SD = 4.85\%$). Results on RTs show no effect of posture switch ($M_{Diff} = .53$ ms, $SE = 4.08$), $F(1, 30.16) = .04$, $p = .85$, $BF = 0.22$, $d = .03$, no effect of direction switch, $F(1, 34.22) = .59$, $p = .45$, $BF = 0.26$, $d = .13$, and no interaction between the two, $F(1, 44.33) = .59$, $p = .45$ (Figures 2D and 3). The Bayes factor for the effect of posture switch was 0.22, indicating substantial evidence for the null hypothesis (i.e., absence of effect). Adding to the results of Experiment 5, Experiment 6 provided further support to the hypothesis that, in Experiments 1–4, automatic role assignment based on body postures reflected role assignment in the presence of relational (dyadic) stimuli, as opposed to single-body posture discrimination. Experiment 7 provides further demonstration that automatic role assignment during dyad perception is more than, and independent of, body posture discrimination.

## Experiment 7: Face-to-Back

The experiments above showed that participants automatically encoded dyadic social scenes in terms of agent and patient role categories based on prototypical postural features. One way for a system to achieve role assignment from body postures would be to track dynamicity and assign the role of agent to the more dynamic of two actors and the role of patient to the less dynamic one. However, if role assignment is the computational goal of the mechanism uncovered here, the agent/patient distinction could also be encoded based on visuospatial cues other than postures. By introducing a novel cue for role assignment, Experiment 7 speaks to this question. In Experiments 1–4, the two actors were always in a symmetric positioning (face-to-face or back-to-back). Thus, body posture was the only cue for discrimination *en route* to role assignment. In Experiment 7, we presented the actors in an asymmetric configuration, with one facing the other who faced away (face-to-back; Figure 2E). We reasoned that, in this configuration, irrespective of body postures (agentlike or patientlike), only one body (the facing body) is in the position to act upon the other, namely, to be the agent. Therefore, in this context, spatial positioning, rather than body posture, should bias role assignment.

### Participants

Twenty-four participants were recruited (19 participants identified their gender as male, five as female; $M_{age} = 25.1$ years, min = 19, max = 34) and tested at https://www.testable.org/. Participants provided informed consent and were compensated $1 for their participation.

### Stimuli and Procedure

The 192 dyads of Experiment 1 (face-to-face dyads) were modified by flipping one of the two actors along the vertical axis, so to obtain a dyad with one facing the other ("behind" actor) and the other facing away ("ahead" actor; Figure 2E). As a result, both actors were oriented in the same direction. We obtained 384 stimuli, with each stimulus from Experiment 1 appearing twice, once with the actors facing leftward, once with the actors facing rightward. The experiment consisted of 10 blocks, each with 48 trials (each transition type occurred 6 times per block), for a total of 480 trials. Stimuli were shown in a random order, and each stimulus was shown between one and four times during the whole experiment. Within each block, from one trial to the next, eight possible changes could occur together or independently: relative positioning of the target (behind/ahead), posture (agentlike/patientlike), side (left/right), and color of clothes (blue/orange). Everything else was identical to Experiments 1–4: Participants performed the gender-search task, reporting the side of the target-actor (the man or the woman). The experiment lasted ~13 min.

### Data Analysis

We followed the same data-cleaning procedure as in previous experiments (Supplemental Table S1). With an LMM, RTs were analyzed as a function of role $switch_{position}$ (same/different role as defined by the position), role $switch_{posture}$ (same/different as defined by the posture), and side switch (same/different). The final random-effect structure was the maximal structure that converged, simplified from the full one with correlated random intercepts and slopes by participant, stimulus, and block.

### Results

Accuracy was overall high ($M = 94.3\%$, $SD = 4.18\%$). Results on RTs show a main effect of role $switch_{position}$, $F(1, 5579) = 25.85$, $p < .001$, $BF = 92.01$, $d = .91$, showing slower RTs in trials in which the target switched from the agentlike to the patientlike position or vice versa ($M_{Diff} = 12.2$ ms, $SE = 2.57$; Figures 2E and 3). We found no effect of role $switch_{posture}$, $F(1, 33.52) = .55$, $p = .47$, $M_{Diff} = 2.61$ ms, $SE = 2.74$, $BF = 0.34$, $d = .21$, and no effect of side switch, $F(1, 22.44) = 2.08$, $p = .16$. $BF = 0.68$, $d = .34$. We found no interaction between role $switch_{position}$ and role $switch_{posture}$, $F(1, 5508.30) = .38$, $p = .54$, between $role_{posture}$ switch and side switch, $F(1, 28.80) = .02$, $p = .90$, and between role $switch_{position}$ and side switch, $F(1, 5535.41) = .02$, $p = .89$. However, we found a three-way interaction, $F(1, 5542.07) = 3.96$, $p = .047$, reflecting the fact that the RSC due to position was significant with a side switch ($t_{21} = -4.41$, $p = .0002$) and without it ($t_{21} = -2.80$, $p = .011$), and with a posture switch ($t_{21} = -.78$, $p = .011$) and without it ($t_{21} = -3.60$, $p = .0017$), but it was stronger when there was a side switch with no posture switch, than when there was a posture switch with no side switch. The RSC due to posture was never significant (with side switch: $t_{21} = -0.65$, $p = .53$, with no side switch: $t_{21} = -0.60$, $p = .56$, with position switch: $t_{21} = -0.22$, $p = .83$, with no position switch: $t_{21} = -1.57$, $p = .13$).

In sum, Experiment 7 showed that in processing visual scenes with actors in a configuration suggesting that one acts upon the other, participants assigned roles based on relative positioning. Adding to Experiments 5–6, Experiment 7 ultimately generalized the automatic role assignment process described here, from mere discrimination of body postures. Doing more than just discriminating between types of postures, the mechanism underlying automatic role assignment could weight different types of information in the visual input to build representations *en route* to event structures.

## General Discussion

The present study asked whether, during perception of visual social scenes, the observers spontaneously encode people in the scene as agents and patients, on the basis of visuospatial information such as body postures and relative positioning, and independently of the access to the semantic content of the scene (i.e., action categorization).

Results showed that, although individual postures changed all the time, from trial to trial, participants were slower to process a scene when the roles of the actors, as cued by their postures, changed, relative to when the roles stayed the same. This happened even though identifying roles was irrelevant to the task, suggesting automatic and implicit role assignment. The effect was consistently found when the scene featured two people in the prototypical configuration for interaction (i.e., face-to-face; Experiments 1 and 4) and was especially reliable for scenes depicting more meaningful (i.e., easy to identify) actions (Experiment 2). Scenes featuring two people at a social distance, but not in the prototypical configuration for interaction (i.e., back-to-back) also elicited implicit role assignment, albeit less consistently: The RSC was only found when the position of the actors (i.e., the side) did not change (Experiments 3–4). Experiments 5–6 showed that a switch cost was not found when the scenes depicted single bodies that, from trial to trial, could switch from one type of posture to another. Experiment 7 ultimately generalized the effect of RSC from a posture-based process to a positioning-based process, showing automatic role assignment based on the position of one actor relative to the other. This finding demonstrated that, in processing multiple-person scenes, representations of roles can be accessed via different types of visual relational information.

Based on these results, we propose the existence of a perception-based selectional process that, during presentation of multiple-person (social) scenes, leverages visuospatial information in the input structure (e.g., body postures, relative positioning) to assign social entities to agentlike and patientlike roles. This process appears to be *automatic*, *relational*, and *smart*. It is automatic to the extent that it takes place in the absence of explicit instructions to identify roles or interpret social interaction events, and when the observer's attention is diverted to aspects of the stimuli unrelated to roles. It is relational to the extent that it only takes place in the context of multiple actors (a dyad at least), suggesting the representation of a relational structure (e.g., *an agent acting on a patient*) that assigns the actors to different roles. Finally, we call the present role assignment process smart, as it weights different types of visual relational information (e.g., postures, positioning) to provide a rich representation of a visual scene, which embeds the entities as well as the relationships that organize those entities into a complex structure. Importantly, in this structure, agentlike and patientlike roles appear to be abstract variables that can be filled by any instantiation of agent and patient (Fodor & Pylyshyn, 1988; Frankland & Greene, 2015; Hafri et al., 2023; Quilty-Dunn et al., 2023). Abstract representation, of which generalization is a key property, would explain how role assignment in one context (trial $n - 1$) could: (a) affect the same process in a different context (from trial $n - 1$ to trial $n$), (b) generalize to variations in spatial locations (e.g., target side) and postures, and (c) be sensitive to the implied relational structure such that it was especially reliable for more canonical social interactions (face-to-face and face-to-back dyads), less consistent for less canonical or obvious social interactions (back-to-back and MF− scenes) and absent for single-body scenes. In short, an abstract format would explain how agentlike and patientlike representations could be accessed via different depictions of events (i.e., the different scenes presented sequentially during the experiment) and different types of perceptual signals (postural and spatial relations).

With the automatic perception-based process for role assignment, this study begins to unravel the foundations of the mechanism by which individuals come to understand *who-did-what-to-whom* in social interaction events. This is the kind of mechanism that might be shared with nonhuman primates and other animals, as evoked in theoretical accounts of the evolutionary origins of syntax (Wilson et al., 2022).

Our results replicate and extend previous work in the field. In Hafri et al. (2018), automatic role assignment was inferred from the RSC in two tasks that required paying attention to the gender of a target-actor or its clothes. This process, however, was uncovered during the processing of meaningful, familiar, prototypical social interactions, where perceptual and semantic information could concur to discriminate and assign roles. In a hypothetical scenario, participants in Hafri et al. automatically accessed the meaning of events—a process as rapid and automatic as the stimulus is more familiar (for review, see Oliva, 2005), and assigned roles as part of the semantic processing of the stimuli. Thus, while forwarding the intriguing idea of automatic role assignment, Hafri et al.'s study did not clarify at what level—that is, based on which type of information—this mechanism would operate and to what extent action/event categorization is required for role assignment.

Our results showed that automatic role assignment occurred for more meaningful as well as less meaningful stimuli, where meaningfulness reflected how much the actions in the scenes were easy to categorize (Experiments 1–2): Although more susceptible to changing perceptual properties of the stimuli (e.g., side switch), individuals assigned the actors to role categories based on postures, even when those postures depicted actions that were not readily identifiable (see also Hafri et al., 2013, for a discussion on the relationship between role assignment and event categorization). Thus, the present work singled out a mechanism for role assignment that would operate on visual relational information, before semantic and inferential processes kick in. With symmetric (face-to-face or back-to-back) body positioning, this mechanism appears to rely on body postures; postures become less relevant when positioning unbalances the relationship between the actors, such that only one can access and act on the other but not vice versa. In short, automatic role assignment would be solved by a computationally straightforward mechanism that relies on postural information when positioning is ambiguous (i.e., symmetric) with respect to the agent–patient discrimination.

Moreover, the sophistication of this mechanism would be such that not all types of symmetric positioning have the same effect for role assignment. We found evidence for a role switch cost when actors were arranged face-to-face, a positioning that is reliably associated with humans' prototypical social interaction (Papeo, 2020; Papeo et al., 2017), as well as when actors where still within the mutual personal or social space, but in a less prototypical configuration for social interaction, that is, back-to-back. This finding was not unexpected: Hafri et al. (2018) also reported a role switch cost when agent and patient were shown back-to-back, breaking the spatial configuration for interaction. Like Hafri et al. (2018), we also observed that the role switch cost was more consistently found with face-to-face dyads than with back-to-back dyads. In the latter condition, we only found an effect in the absence of side switch. Similar to the less meaningful stimuli, for the back-to-back stimuli, role assignment did not generalize to actors who switched position, suggesting retrieval of role representations that were less abstract and more tied to the perceptual properties of the stimuli.

Altogether, our results demonstrate that visual social scenes are automatically encoded in terms of agent–patient relations based on visual relational cues (e.g., posture and positioning), and that the representation of roles as abstract categories is facilitated the more the visuospatial cues converge toward the representation of a meaningful social interaction.

This study raises a number of interesting questions for further investigation. First, how is the role integrated in the representation of social entities in a visual event? Some authors have proposed that relations can be represented independently from the participating entities (Hafri & Firestone, 2021; see for discussion Hochmann & Papeo, 2021). New findings, however, suggest that the role of an entity in an event could be represented as a feature of that entity, analogously to other features such as gender, size, color, and so on (Yu et al., 2023).

Second, we not only found a hierarchy in perceptual cues for role assignment, whereby relative positioning overruled posture, but we also found a numerically larger role switch cost for positioning-based versus posture-based role assignment (see Figure 3: Experiments 1 and 4 vs. 7). This circumstance suggests that the agent was easier to discriminate from the patient in the face-to-back condition. This could be because in face-to-back dyads one actor lacked a critical cue of agency, that is, the orientation of attention/body/action toward a goal (Woo & Spelke, 2023), and/or because positioning could be extracted faster/more easily based on the global shape of the bodies, while postures required the analysis of finer grained visual information. With relatively short stimulus presentation, this latter factor might have played a role, making role assignment easier in the face-to-back context. Finally, it remains possible that the difference in role switch cost reflected qualitatively different agentlike and patientlike representations triggered by reciprocal (face-to-face) versus transitive (face-to-back) event structures (but see Papeo et al., 2024).

Third, our work assumed categorical agent–patient representation, but key perceptual and nonperceptual properties of agency and patience may vary in degrees (H. M. Gray et al., 2007). Determining the parameters that make one an agent or a patient could contribute to understanding how, beyond thematic roles, people attribute responsibilities, intentionality, and other mental contents to social beings (De Freitas & Alvarez, 2018; K. Gray et al., 2012).

Fourth, future studies should establish how the automatic perceptual cue-based role assignment described here relates to role assignment in language and to what extent it is permeable to top-down influences (Rissman & Lupyan, 2022).

Last, our results show that agent/patientlike representations automatically retrieved during event perception are abstract. How early do such abstract representations emerge during development? At 6 months, infants are sensitive to perceptual features underlying role assignment in causal events. For instance, in launching events, they distinguish between *launcher* and *launchee* based on spatiotemporal contingencies of the objects' motion (Leslie & Keeble, 1987; Oakes & Cohen, 1990). An open question remains whether infants also assign the *launcher* and the *launchee* to the abstract categories of *agent* and *patient*, respectively, or abstraction takes time (and/or other processes) to develop (for recent results see Papeo et al., 2024).

In conclusion, we showed that, in processing multiple-person visual scenes, an abstract agent–patient structure is automatically generated based on a hierarchy of visual cues, where the observer relies on body postural relations when spatial relations are ambiguous about who acts on whom. This mechanism may unravel the basis and precursors of the agent–patient social dyad, a human universal that, by defining *who-does-what-to-whom*, determines how individuals understand, think, and talk about social events.

## Constraints on Generality

Since the instructions were in English and we wanted to ensure that all participants understood the instructions correctly, we only included participants who indicated English as their native language. Although this choice may limit the generalization of results, we included participants from a wide range of nationalities (Indian, Venezuelan, British, Portuguese, Kenyan, American, Latvian, Zimbabwean, Indian, Filipino, South African, Polish, Pakistani, Romanian, Malaysian, Serbian, Chilean, Kashmiri, Sri Lankan, Spanish, Canadian, Nigerian). Moreover, since all participants were between 18 and 35 years old, results may not generalize to populations in other age ranges.

## References

Abdai, J., Ferdinandy, B., Terencio, C. B., Pogány, Á., & Miklósi, Á. (2017). Perception of animacy in dogs and humans. *Biology Letters*, *13*(6), Article 20170156. https://doi.org/10.1098/rsbl.2017.0156

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, Article 328. https://doi.org/10.3389/fpsyg.2013.00328

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious mixed models*. arXiv. https://doi.org/10.48550/arXiv.1506.04967

Bates, D., Mächler, M., Bolker, B., Walker, S., Christensen, R. H., Singmann, H., & Dai, B. (2015). *lme4: Linear mixed-effects models using Eigen and S4* (R package, Version 1.1–7. 2014) [Computer software]. https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf

Choi, H., & Scholl, B. J. (2006). Measuring causal perception: Connections to representational momentum? *Acta Psychologica*, *123*(1–2), 91–111. https://doi.org/10.1016/j.actpsy.2006.06.001

De Freitas, J., & Alvarez, G. A. (2018). Your visual system provides all the information you need to make moral judgments about generic visual events. *Cognition*, *178*, 133–146. https://doi.org/10.1016/j.cognition.2018.05.017

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619. https://doi.org/10.1353/lan.1991.0021

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1–2), 3–71. https://doi.org/10.1016/0010-0277(88)90031-5

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 of the United States of America*, *112*(37), 11732–11737. https://doi.org/10.1073/pnas.1421236112

Galazka, M., & Nyström, P. (2016). Infants' preference for individual agents within chasing interactions. *Journal of Experimental Child Psychology*, *147*, 53–70. https://doi.org/10.1016/j.jecp.2016.02.010

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619. https://doi.org/10.1126/science.1134475

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101–124. https://doi.org/10.1080/1047840X.2012.651387

Gruber, J. (1965). *Studies in lexical relations* [Doctoral dissertation, Department of Modern Languages]. Massachusetts Institute of Technology.

Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences*, *25*(6), 475–492. https://doi.org/10.1016/j.tics.2021.01.006

Hafri, A., Green, E. J., & Firestone, C. (2023). Compositionality in visual perception. *Behavioral and Brain Sciences*, *46*, e277. https://doi.org/10.1017/S0140525X23001838

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

Hochmann, J.-R., & Papeo, L. (2021). *How can it be both abstract and perceptual? Comment on Hafri, A., & Firestone, C. (2021), the perception of relations, trends in cognitive sciences*. PsyArXiv. https://doi.org/10.31234/osf.io/hm49p

Kominsky, J. F., & Scholl, B. J. (2020). Retinotopic adaptation reveals distinct categories of causal perception. *Cognition*, *203*, Article 104339. https://doi.org/10.1016/j.cognition.2020.104339

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *Package "emmeans."* (R Package, Version, 1 3.2) [Computer software]. https://cran.r-project.org/web/packages/emmeans/index.html

Leslie, A. M., & 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

Mascalzoni, E., Regolin, L., & Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4483–4485. https://doi.org/10.1073/pnas.0908792107

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Michotte, A. (1963). *The perception of causality*. Basic Books.

Moors, P., Wagemans, J., & de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ*, *5*, Article e2932. https://doi.org/10.7717/peerj.2932

Morey, R. D., Rouder, J. N., Pratte, M. S., & Speckman, P. L. (2011). Using MCMC chain outputs to efficiently estimate Bayes factors. *Journal of Mathematical Psychology*, *55*(5), 368–378. https://doi.org/10.1016/j.jmp.2011.06.004

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

Oliva, A. (2005). CHAPTER 41—Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). Academic Press. https://doi.org/10.1016/B978-012375731-9/50045-8

Paparella, I., & Papeo, L. (2022). Chunking by social relationship in working memory. *Visual Cognition*, *30*(5), 354–370. https://doi.org/10.1080/13506285.2022.2064950

Papeo, L. (2020). Twos in human visual perception. *Cortex*, *132*, 473–478. https://doi.org/10.1016/j.cortex.2020.06.005

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion Effect. *Psychological 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

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, *46*, Article e261. https://doi.org/10.1017/S0140525X22002849

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. *Progress in Brain Research*, *253*, 243–262. https://doi.org/10.1016/bs.pbr.2020.06.005

Rissman, L., & Lupyan, G. (2022). A dissociation between conceptual prominence and explicit category learning: Evidence from agent and patient event roles. *Journal of Experimental Psychology: General*, *151*(7), 1707–1732. https://doi.org/10.1037/xge0001146

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, *26*(6), 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology*, *23*(3), 250–254. https://doi.org/10.1016/j.cub.2012.12.017

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., & Højsgaard, S. (2020). *afex: Analysis of factorial experiments, 2021* (R Package, Version, 1–0) [Computer software]. https://cran.r-project.org/web/packages/afex/index.html

Wilson, V. A. D., Zuberbühler, K., & Bickel, B. (2022). The evolutionary origins of syntax: Event cognition in nonhuman primates. *Science Advances*, *8*(25), Article eabn8464. https://doi.org/10.1126/sciadv.abn8464

Woo, B. M., & Spelke, E. S. (2023). Toddlers' social evaluations of agents who act on false beliefs. *Developmental Science*, *26*(2), Article e13314. https://doi.org/10.1111/desc.13314

Yu, X., Li, J., Zhu, H., Tian, X., & Lau, E. (2023). *Electrophysiological hallmarks for event relations and event roles in working memory*. bioRxiv. https://doi.org/10.1101/2023.05.08.539845

Received October 11, 2023
Revision received July 16, 2024
Accepted July 24, 2024 ■