# Electrophysiological investigation of infants' understanding of understanding

Balint Forgacs, Judit Gervain, Eugenio Parise, Gergely Csibra, György Gergely, Júlia Baross, Ildiko Kiraly

▶ **To cite this version:**

Balint Forgacs, Judit Gervain, Eugenio Parise, Gergely Csibra, György Gergely, et al.. Electrophysiological investigation of infants' understanding of understanding. Developmental Cognitive Neuroscience, 2020, 43, pp.100783. 10.1016/j.dcn.2020.100783 . hal-02868065

**HAL Id: hal-02868065**
**https://hal.science/hal-02868065**

Submitted on 16 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Developmental Cognitive Neuroscience 43 (2020) 100783

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn

# Electrophysiological investigation of infants' understanding of understanding

Bálint Forgács [a,b,*], Judit Gervain [c,d], Eugenio Parise [e,f], Gergely Csibra [f,g], György Gergely [f], Júlia Baross [a], Ildikó Király [a]

[a] MTA-ELTE Social Minds Research Group, Eötvös Loránd University (ELTE), Izabella utca 46, 1064, Budapest, Hungary
[b] Hungarian Academy of Sciences, Nádor utca 7, 1051, Budapest, Hungary
[c] Integrative Neuroscience and Cognition Center, Université de Paris, 45, rue des Saints-Pères, 75006 Paris, France
[d] Integrative Neuroscience and Cognition Center, CNRS, 45, rue des Saints-Pères, 75006 Paris, France
[e] Department of Psychology, Lancaster University, Lancaster LA1 4YF, United Kingdom
[f] Cognitive Development Center (CDC), Department of Cognitive Science, Central European University (CEU), Nádor utca 9, 1051 Budapest, Hungary
[g] Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom

ARTICLE INFO

*Keywords:*
Theory-of-Mind
Language acquisition
N400
ERPs
Social cognition
False belief

ABSTRACT

Social cognition might play a critical role in language acquisition and comprehension, as mindreading may be necessary to infer the intended meaning of linguistic expressions uttered by communicative partners. In three electrophysiological experiments, we explored the interplay between belief attribution and language comprehension of 14-month-old infants. First, we replicated our earlier finding: infants produced an N400 effect to correctly labelled objects when the labels did not match a communicative partner's beliefs about the referents. Second, we observed no N400 when we replaced the object with another category member. Third, when we named the objects incorrectly for infants, but congruently with the partner's false belief, we observed large N400 responses, suggesting that infants retained their own perspective in addition to that of the partner. We thus interpret the observed social N400 effect as a communicational expectancy indicator because it was contingent not on the attribution of false beliefs but on semantic expectations by both the self and the communicative partner. Additional exploratory analyses revealed an early, frontal, positive-going electrophysiological response in all three experiments, which was contingent on infants' computing the comprehension of the social partner based on attributed beliefs.

## 1. Introduction

Recent findings suggest that some of the cognitive mechanisms allowing infants to interact socially with others are subserved by the same neural mechanisms whether they are applied to the self or to the mental state of others. For example, 8-month-olds produce gamma-band electroencephalographic (EEG) activity over the temporal cortex, a potential marker of sustaining the representation of an object during occlusion, both when tracking objects themselves and when taking into account the beliefs of another person about occluded objects (Kampis et al., 2015). Similarly, 6-month-olds produce motor activation, as indicated in the EEG by sensorimotor alpha suppression over central areas, when they expect an actor to reach in accordance with her false belief (Southgate and Vernetti, 2014). These findings suggest that the same underlying neural mechanisms are recruited for representing objects or preparing actions from first- and third-person perspectives.

Shared mechanisms between self-related and other-related computations might be indispensable for linguistic communication. This is because language comprehension and production also depends on the social context, and in particular on the mental states of communicative partners (Grice, 1957; Sperber and Wilson, 1986). Learning the sound patterns of their native language(s), segmenting speech into words (Gervain et al., 2008), and even understanding something about some words (e.g., Bergelson and Swingley, 2012) are not sufficient for young infants to be involved in sophisticated communicative exchanges with others. Yet around 6 months of age they assume that speech transmits information between communicative partners even about perceptually not observable things such as preferences (Vouloumanos et al., 2014).

* Corresponding author at: MTA-ELTE Social Minds Research Group, Eötvös Loránd University (ELTE), Izabella utca 46, 1064, Budapest, Hungary.
*E-mail address:* forgacs.balint@ppk.elte.hu (B. Forgács).

https://doi.org/10.1016/j.dcn.2020.100783
Received 8 July 2019; Received in revised form 1 April 2020; Accepted 9 April 2020
Available online 28 April 2020
1878-9293/© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Soon after, around 9 months of age, the neural marker of semantic processing, the N400 event-related potential (ERP), can be evoked by mislabeling objects, even though infants need some situational aides (Junge et al., 2012; Parise and Csibra, 2012) Semantic processing becomes less dependent on contextual cues around 14 months of age (Friedrich and Friederici, 2005, 2008). The functional role of the N400 has been debated for decades, but recent reviews suggest that it is an indicator of retrieval effort from semantic memory (Brouwer et al., 2012; Kutas and Federmeier, 2011).

Around the same age when infants begin to exhibit the N400 consistently, they start to produce it to process the meaning of words on behalf of others. Forgács and colleagues (2019) measured the ERPs of 14-month-olds in a live puppet-theater experiment, where objects were labelled in the presence of an adult Observer. The objects were always named correctly for babies, but sometimes incorrectly from the Observer's perspective. In such cases the object – unbeknownst to her – was replaced with another object before naming. Thus she must have falsely believed that the label she heard was incongruent with the object. Infants produced an N400 effect in this situation, even though the label was congruent with the object from their own perspective. Because of its social nature, similar effects are sometimes referred to as the 'social N400' (Jouravlev et al., 2019; Rueschemeyer et al., 2015; Westley et al., 2017). Such processes might not only help adults to track the comprehension of communicative partners but could aid infants to disambiguate and learn language – even though mentalization in communication should be useful beyond misunderstandings per se. The observation of the social N400 in infants (Forgács et al., 2019) suggests that their semantic expectations might be generated by the same neural system for the self and when attributed to a communicative partner.

While the studies by Forgács et al. (2019) demonstrated the tracking of semantic comprehension of social partners by 14-month-olds, they did not clarify the nature of this process. In particular, they did not reveal whether infants tracked the Observer's representation of the referent as a particular object or as a member of an object kind. In principle, the second option would have been sufficient to detect the mismatch between the label and the attributed representation. It was also unclear whether the social N400 was due to false beliefs represented in the semantic system created at the appearance of the object, or due to the evaluation of an episodic memory representation by the semantic system, prompted by the labeling.

We performed three experiments to answer these questions. First, as a baseline, we sought to replicate Forgács et al.'s (2019) main finding in a different laboratory, using a different language (Experiment 1). Second, to investigate the question about the functional role of the N400 in mental state attribution and the nature of the representation attributed to the Observer (whether it is on a type or a token level), we used the original paradigm, but instead of replacing the object with a new object from a different category (e.g., a shoe replaced by a cup), we now replaced the object with a new one from the same category (e.g., a red shoe replaced by a blue shoe). This manipulation allowed us to investigate whether false beliefs are attributed on a type or a token level, and whether the N400 is functionally necessary for tracking misunderstandings (Experiment 2).

Third, to investigate whether the social N400 was due to the evaluation of an episodic memory representation by the semantic system, prompted by labeling, we inverted our original paradigm: we always labelled the objects incorrectly for babies, but half of the times congruently with the false belief of the Observer (Experiment 3). If the first object (the attributed false belief) is represented in infants' semantic memory and plays a direct role in modulating semantic activation, as could be suspected based on our previous study, the N400 effect should be reduced or absent in the inverted paradigm. Alternatively, false belief representations could interact with the semantic system only at the labeling. These manipulations of semantic congruity and false beliefs could shed light on how infants establish word meaning, for themselves and for others, in communicative situations.

## 2. Experiment 1

Our previous study (Forgács et al., 2019) was conducted with French infants in Paris, while the current series of studies were performed with Hungarian infants in Budapest. Since French and Hungarian are different in their grammatical structure, with Hungarian having heavily suffixing morphology, word learning could follow a different trajectory than in French, given the distinct segmentation and morphological decomposition tasks. Thus, we first sought to replicate the original result in order to test the generalizability of the effect across languages and laboratories. We kept every aspect of the current study identical to the previous one, except for a few minor technical modifications that are reported below.

### 2.1. Material and methods

**Participants.** Eighteen typically developing, full-term 14-month-olds (mean age: 443 days; range: 427–456 days) were included in the final sample of this study. They were all Hungarian monolingual babies, who did not hear any other languages for more than 15 % of the time according to parental report. An additional 69 infants were excluded (21 % inclusion rate), just as in our previous study (Forgács et al., 2019), for the following reasons: fussiness ($N = 27$), refusing wearing the electrode net ($N = 2$), technical error ($N = 2$), not producing at least 10 artifact free trials (Stahl et al., 2010) ($N = 20$), or providing overall too noisy ERPs ($N = 18$) as determined by the authors (cf. Forgács et al., 2019). Participant exclusion was carried out in batch, prior to statistical analyses, was blind to any systematic individual differences that would be relevant for language comprehension and Theory-of-Mind (ToM) abilities, and it was mainly driven by infants looking away from the object or the Observer at the moment of object naming, which led to excessive eye-movements.

**Materials.** Fifteen objects (e.g., a bunny, a cup, a shoe, a duck, etc., adapted from Parise and Csibra, 2012), and the audio recording of the corresponding Hungarian object labels were used in the experiment. We kept the teddy from the original list (reported by Parise and Csibra, 2012), which was replaced by a nursing bottle for the French study (Forgács et al., 2019). We also recorded a number of attention getting exclamations ("Look!", "How interesting!", "Listen!" etc.) in an infant-directed intonation, which were used on demand.

**Apparatus.** Infants were seated on their parent's lap facing a puppet theater stage. On the other side of the stage an adult was seated, who played the role of the Observer. There was a curtain right in front of her, which she could operate manually and a 30 cm tall occluder in the middle of the stage, which was operated by another individual, the Experimenter, who was sitting on the left side of the stage, hidden from the infant. The Experimenter placed objects on the stage in front of the occluder, visible for infants at all times but for the Observer only when the occluder was lowered. Infants were video-recorded using a camera that was placed beneath the stage to allow for later visual analysis and confirmation of eye-blinks and other movement artifacts. Audio playback, object order, and EEG triggering was controlled using Psychotoolbox-3 running on MATLAB 2017 (MathWorks Inc, Natick, Massachusetts, USA).

**Procedure.** As a family arrived in the lab, the experimenters played briefly with the infant, they then explained the procedure to the parents and obtained their informed consent. We attained the ethics approval from the Ethics Committee of Eötvös Loránd University. During the experiment, parents wore headphones playing music, and they were requested to look downwards to the floor, not to follow events on the stage and not to talk to or interact with their children, unless they wished to take a break to calm and help them engage again.

The procedure was identical to that of Forgács et al. (2019), where we provided a detailed description; the sequence of events for all experiments is provided in Fig. 1. The critical event in each trial occurred when an object was pointed and verbally labelled, and the crucial

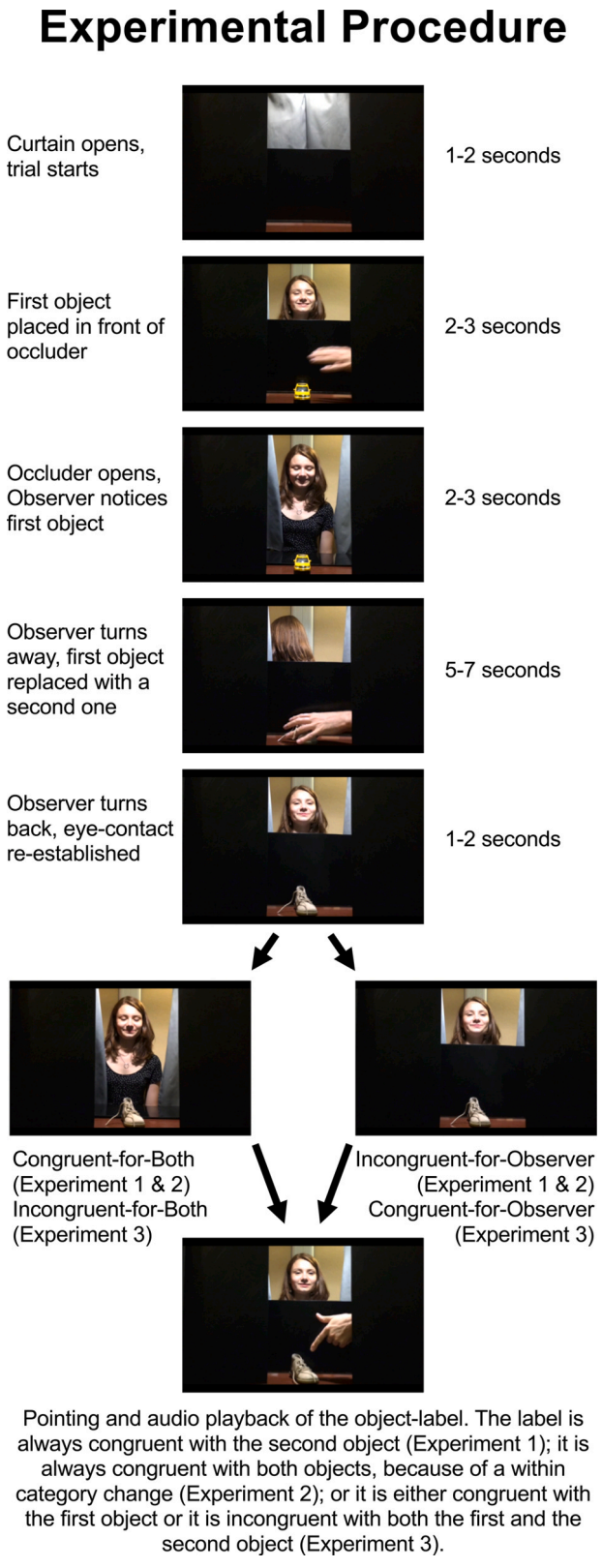


**Fig. 1.** Experimental procedure for all three experiments. Infants viewed a live puppet-theater show mixing an object naming and a false belief paradigm. Across experiments objects were labelled either correctly or incorrectly while an Observer, sitting in front of the infant, had either a true or a false belief about the identity of the objects. (Adapted from Forgács et al., 2019 with minor modifications describing the experiments reported here).

manipulation of the experiment was whether this label matched (Congruent-for-Both condition) or did not match (Incongruent-for-Observer condition) the belief that the Observer should have had about the identity of the object. Each trial started with the Observer opening the curtain and establishing eye-contact with the infant. The Experimenter placed an object (e.g., a toy car) in front of the occluder at the center of the stage and lowered the occluder. The Observer re-established eye--contact by looking at the object and then in the eyes of infants to make sure infants noticed that she saw the object. Then the Experimenter raised the occluder, the Observer turned away 90° to her right, and the Experimenter replaced the first object with a second one (e.g., a plush bunny). After the object change, the Observer turned back, at which point the two conditions diverged. In the Incongruent-for-Observer condition the Experimenter pointed at the object, and the Observer triggered an audio recording of the label of the object using a hidden gamepad. In the Congruent-for-Both condition the Experimenter lowered the occluder, and the Observer again made sure that the infant noticed that she saw the second object by establishing joint attention with him/her. Only then the Experimenter raised the occluder, pointed at the object, and the Observer triggered the audio label. The Observer made sure that the infant paid attention and that her own face was always identical and neutral in both conditions. Two seconds after sound playback, the Experimenter removed the object, and the Observer closed the curtains for two seconds before the next trial began. Trials were presented in a pseudo random order: a particular object appeared only once in every five trials, the same condition did not repeat for more than two consecutive trials, and in Incongruent-for-Observer trials the first object was replaced with an object from a different category. Individual pseudo-random stimulus lists were generated using Python 3.6 (Python Software Foundation, Beaverton, Oregon, USA).

**EEG recording and analysis.** A continuous EEG signal was recorded using a 128-channel Geodesic Sensor Net (EGI Inc., Eugene, Oregon, USA) at a 500 Hz sampling rate. The data was filtered with a 0.3 Hz high-pass and 30 Hz low-pass filter, and then segmented into epochs starting 200 ms before and 1200 ms after audio onset. Automatic artifact detection algorithms for blinks, eye-movements (both confirmed using the video recording), and bad channels were employed to mark bad segments. Infants had to produce at least 10 clean, artifact-free trials per condition (Stahl et al., 2010) to have their data averaged and considered for inclusion. We replaced bad channels using spherical spline interpolation, averaged segments separately for the two conditions, applied a 200 ms pre-stimulus baseline correction, and re-referenced to the average reference. We quantified ERP responses over Regions-of-Interests (ROIs), consisting of 13 frontal electrodes (3, 4, 5, 9, 10, 11, 12, 15, 16, 18, 19, 22, 23) and 13 parietal ones (62, 65, 66, 67, 70, 71, 72, 75, 76, 77, 83, 84, 90), exactly as in Forgács and colleagues (2019). Statistical reports follow the recommendations of Lakens (2013). All raw EEG data are available on-line at the Open Science Framework (https://osf.io/p52ey/).

### 2.2. Results

To quantify the N400 effect, we averaged the ERP amplitudes over the parietal ROI in the 400−600 ms time window, where we observed a significant difference between the Incongruent-for-Observer ($M = –13.5$ μV, $SD = 6.72$ μV) and the Congruent-for-Both conditions ($M = –10.6$ μV, $SD = 6.32$ μV), $t(17) = –2.46$, $p = .025$, 95 % CI of the difference [–5.32 μV, –0.41 μV], Hedges' $g_{av} = 0.42$ (Fig. 2). This result replicates that of Forgács et al. (2019). In contrast, we have not replicated the late frontal negativity reported by Forgács et al. (2019), as we found no evidence for a difference between the Incongruent-for-Observer ($M = 20.3$ μV, $SD = 10.9$ μV) and the Congruent-for-Both conditions ($M = 16.7$ μV, $SD = 10.2$ μV) in the 700−1000 ms time window, $t(17) = 1.25$, $p = .23$, 95 % CI [–2.46 μV, 9.57 μV], Hedges' $g_{av} = 0.32$.

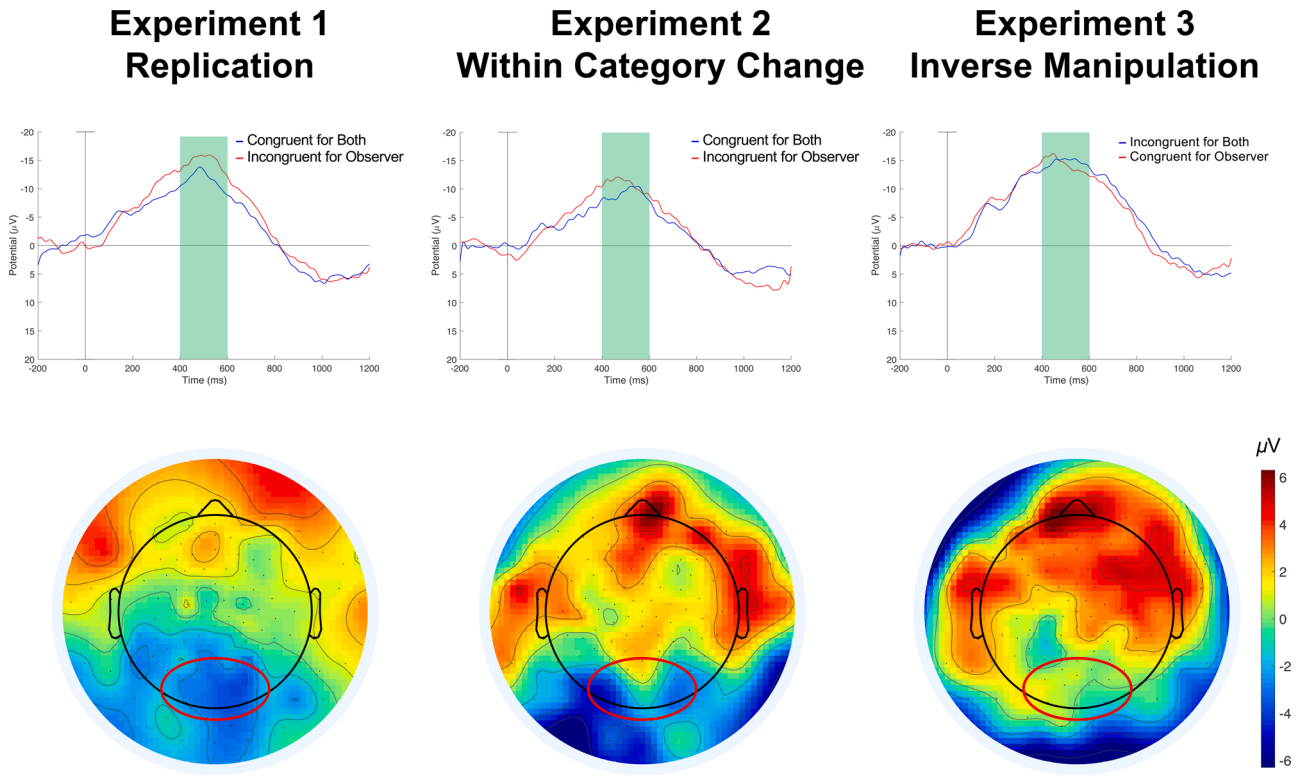


**Fig. 2.** The N400 effect in Experiment 1, 2, & 3. Upper panels show ERP plots of the grand average of the electrodes over the parietal ROI. Negative is plotted upwards, and time 0 is the onset of the audio playback. Green shadings indicate the time window of the infant N400 (400-600 ms). Lower panels show the topographical maps of the ERP difference of the conditions in the N400 time window for all studies. Colder colors indicate greater negativities and red rings illustrate the ROI over which the N400 amplitude was quantified. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

### 2.3. Discussion

We successfully replicated our previous finding that 14-month-old infants exhibit an N400 effect when another person in their presence could have thought that a label did not match an object, even though from the infants' perspective it did. Both Hungarian and French babies appear to follow the linguistic comprehension of a social partner. We did not replicate the late frontal negativity reported by Forgács et al. (2019). In Section 5, we will come back to this matter using combined exploratory statistical analyses of all three studies.

Since the emergence of the social N400 effect that we replicated in Experiment 1 relies on the attribution of false beliefs to the communicative partner, it can also be exploited to ask questions about the content of the beliefs that infants attribute to others. Experiment 2 was designed to answer such a question.

## 3. Experiment 2

We designed an experiment that was identical with Experiment 1, except that in each trial we replaced the object with another one from the same category (e.g., a toy car with another toy car). As a result, labels were always correct for both parties, but for the infant the referent was the second object, while for the Observer it was the first object. If beliefs are tracked at the kind level, no semantic mismatch for the Observer should occur and we should find no difference in the N400s between the two conditions. In contrast, if beliefs are tracked at the level of individual objects, we should find diverging ERP responses. Specifically, if the N400 effect in Experiment 1 was driven by the mismatch between the referent of the Observer's belief and the referent of the word label, we should find an N400 effect in this paradigm as well. Even though the label was congruent with both the object in front of the baby and the object representation entertained by the Observer, the pointing gesture and the overall presentation made the referent of the label unambiguously incongruent with the referent of the false belief of the Observer in the Incongruent-for-Observer condition.

### 3.1. Material and methods

**Participants.** Eighteen full-term, Hungarian monolingual 14-month-olds (mean age: 442 days; range: 429–454 days) participated. A further 15 infants were excluded (53 % inclusion) because of fussiness ($N = 2$), technical error ($N = 1$), not providing a minimum of 10 artifact free trials ($N = 11$), and providing too noisy grand averages ($N = 1$). The lower exclusion rate in this Experiment could be due to the design that involved no semantic incongruity for either parties, which could have made the study less frustrating for the infants.

**Materials.** We extended the set of objects with an additional 15 toys in order to have two tokens of each type.

**Apparatus and EEG recording and analysis** were the same as in Experiment 1.

**Procedure.** The procedure was identical to that of Experiment 1, except that every object was replaced by a different category member. Every object served as first and second object in equal number of times.

### 3.2. Results

First, N400 amplitudes were calculated over the parietal ROI in the 400−600 ms time window as in Experiment 1. The comparison of the two conditions revealed no evidence for a significant difference, $t(17) = -0.81$, $p = .43$, 95 % CI [–5.37 μV, 2.39 μV], Hedges' $g_{av} = 0.16$ (Fig. 2), between the Incongruent-for-Observer ($M = -9.85$ μV, $SD = 9.38$ μV) and the Congruent-for-Both conditions ($M = -8.36$ μV, $SD = 8.83$ μV). As

the Congruent-for-Both condition was the same as the Congruent-for-Both condition of Experiment 1, which involved no false beliefs and labels congruent for both parties, the lack of an amplitude difference (i.e. N400 effect) between the two conditions of Experiment 2 also implies that no N400 response (i.e. amplitude increase) was evoked by either conditions in this paradigm.[1]

Just like in Experiment 1, we did not find a greater frontal negativity in the 700−1000 ms time window for the Incongruent-for-Observer ($M$ = 17.2 μV, $SD$ = 10.8) compared to the Congruent-for-Both condition ($M$ = 19.8 μV, $SD$ = 10.4), $t(17) = –0.83$, $p = .42$, 95 % CI [–9.44 μV, 4.11 μV], Hedges' $g_{av} = 0.24$.

### 3.3. Discussion

Labels were congruent for both parties, and the only potential difference between the two conditions was the specific content of the beliefs attributed to the Observer. The lack of an N400 effect (no N400 response in the Incongruent-for-Observer condition either) is not unexpected in this experiment according to the N400 literature: it suggests that the social N400 was not the result of a referent violation in Experiment 1 either, but it likely reflects semantic priming based on category information. We will discuss the lack of the frontal effect in Section 5.

## 4. Experiment 3

According to our previous findings (Experiment 1 and Forgács et al., 2019), the social N400 was linked to infants' representation of the Observer's false belief. However, those studies did not clarify whether the social N400 effect was due to the mismatch between the label and the representation attributed to the Observer, or to the mismatch between the infants' own representation and the one they attributed to the Observer, prompted by the labeling. In Experiment 3, we investigated whether the content of the attributed mental state of the Observer plays a direct role in evoking an N400 effect in infants, by inverting the original paradigm: we named objects always incongruently from the perspective of infants, but in half the trials with the label of the preceding object. Consequently, the label was incongruent with the object but it was congruent with the Observer's false belief about the object. If infants simply retained the label of the first object in their semantic memory, one would expect a reduced N400 effect, even on the classical account of the N400 (Brouwer et al., 2012; Kutas and Federmeier, 2011). If, however, the N400 effect indexes any semantic mismatch when episodic memories of attributed representations are evaluated, a similar magnitude N400 response could be predicted.

### 4.1. Material and methods

**Participants.** In this experiment 18 full term, Hungarian monolingual 14-month-olds (mean age: 441 days; range: 427–456 days) participated. A further 63 infants were excluded (22 % inclusion) because of fussiness ($N = 24$), refusing the cap ($N = 3$), technical error ($N = 6$), not providing a minimum of 10 artifact free trials ($N = 25$), and providing too noisy individual averages ($N = 5$).

**Materials, apparatus, and EEG recording and analysis** were identical to those of Experiment 1.

**Procedure.** Generally, the procedures were identical to those of Experiment 1. The only difference was the audio playback. In the Congruent-for-Observer condition, in which the occluder was not lowered, the Observer triggered the label of the first object (i.e., the one that had been removed) while Experimenter was pointing at the second object (which was different from the first one). In the Incongruent-for-Both condition, in which the occluder was lowered and the Observer saw the second object as well, she played back a label that was not congruent with either objects.

### 4.2. Results

Statistical analysis of the N400 time window over the parietal ROI did not reveal any evidence for a significant difference between the Congruent-for-Observer ($M = –13.0$ μV, $SD = 8.24$ μV) and Incongruent-for-Both conditions ($M = –13.3$ μV, $SD = 8.01$ μV), $t(17) = 0.22$, $p = .82$, 95 % CI [–2.79 μV, 3.46 μV], Hedges' $g_{av} = 0.04$. The latter condition is equivalent to the incongruent condition of Experiment 1 in Forgács et al. (2019), which compared it to a Congruent-for-Both condition. Since that study yielded a significant effect, the lack of an amplitude difference (i.e. an N400 effect) between the conditions in the present experiment indicates that both conditions evoked an N400 response or component (i.e. amplitude increase) (Fig. 2).

We did not find an enhanced frontal negativity in the 700−1000 ms time window for the Incongruent-for-Both ($M = 19.7$ μV, $SD = 11.0$ μV) relative to the Congruent-for-Observer condition ($M = 17.3$ μV, $SD = 10.3$ μV), $t(17) = 0.76$, $p = .46$, 95 % CI [–4.12 μV, 8.83 μV], Hedges' $g_{av} = 0.21$.

### 4.3. Discussion

When infants experienced incongruity between what they saw and what they heard, we found that the congruity of a label with the Observer's false belief in the Congruent-for-Observer condition did not reduce infants' N400 response, and we did not observe an N400 difference / effect in comparison to the Incongruent-for-Both condition. Such an outcome is incompatible with the classical interpretation of the N400 (Brouwer et al., 2012; Kutas and Federmeier, 2011): if infants remembered the label of the first object, due to belief tracking or otherwise, they should have had sufficient semantic memory pre-activation of target words in the Congruent-for-Observer condition, reducing their N400 amplitudes. Based on our earlier findings, however, we assume that infants did track the contents of the Observer's mental representation in this experiment, too. But this did not impact the N400. We, therefore, argue that the cognitive mechanisms behind the N400 may not track the match between the understanding of the participant and an interlocutor. Rather, the N400 reflects semantic incongruity of a word with any representation of the referent, whether it is the first-order representation of the infant (Friedrich and Friederici, 2008), an episodic representation attributed to a communicative partner (Experiment 1 in this study and Experiment 2 in Forgács et al., 2019), or both (the present, Experiment 3). In other words, any semantic incongruity may manifest itself as an increment of the N400, when an episodic memory trace is prompted for semantic evaluation by language.

## 5. Combined Exploratory Statistical Analyses

While the frontal negativity, which was not predicted but found by Forgács et al. (2019), was not replicated in any of the experiments in the present study, upon visual examination of the frontal ROI we noticed an early positive difference between the conditions in all three experiments with Hungarian samples. This effect was most pronounced in the 300−500 ms time window (Fig. 3). While the experimental manipulations differed across the experiments, in all three cases the Observer held a false belief about the identity of the referent in one condition (Experiment 1 & 2: Incongruent-for-Observer, Experiment 3: Congruent-for-Observer) and a true belief in the other one. In order to substantiate the observation that the belief of the Observer was reflected by this difference, we ran a combined 3 × 2 exploratory ANOVA on the

[1] Sometimes the expressions N400 "effect" and "response" or "component" are used interchangeably, however, there is a critical distinction: the "response" or "component" is an amplitude change from the baseline of stimulus onset, while the "effect" is the difference of this change between two conditions (cf. Molinaro et al., 2011).

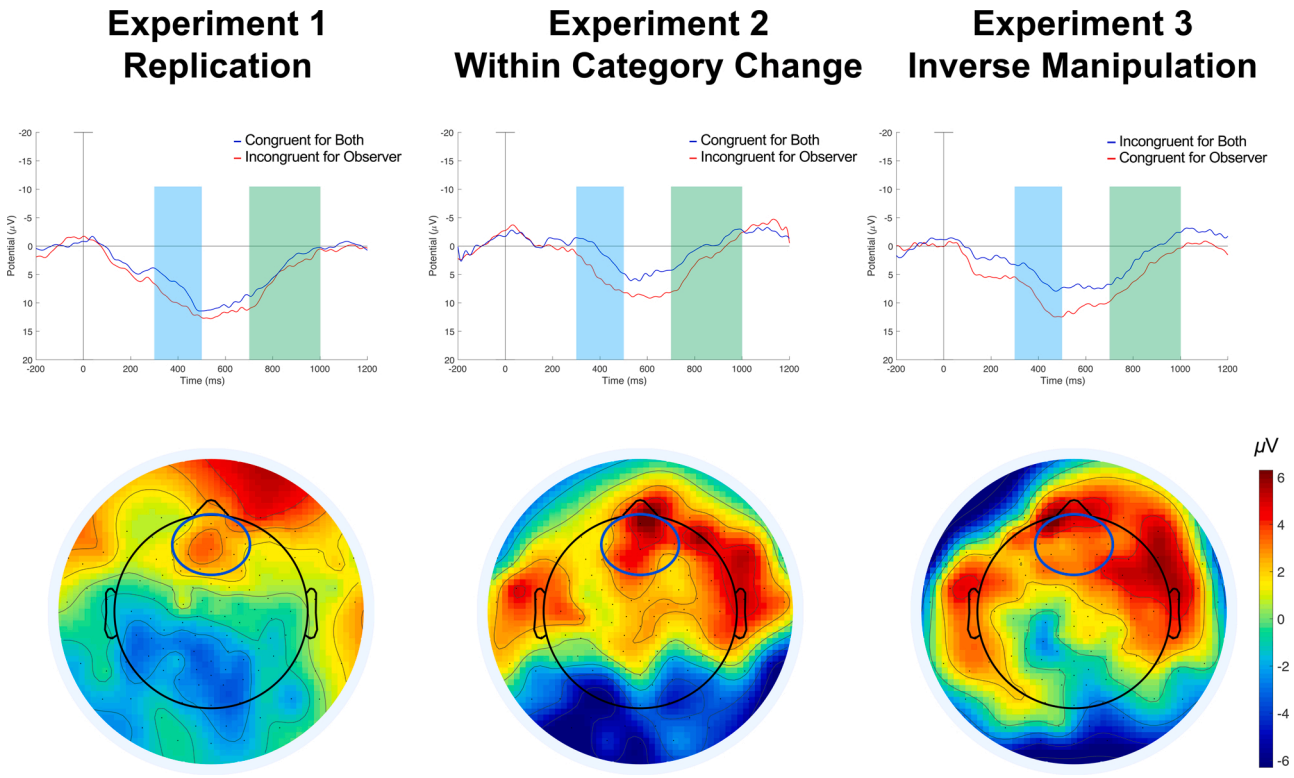


**Fig. 3.** The frontal positive going brain wave in Experiment 1, 2, & 3. Upper panels show ERP plots of grand averages of electrodes over the frontal ROI. Negative is plotted upwards, and time 0 is the onset of the audio playback. Green shades indicate the time window (700-1000 ms) of the previously reported frontal effect (Forgács et al., 2019), blue shades where exploratory analyses were carried out (300-500 ms). Lower panels show the topographical maps of the ERP difference of the conditions in the time windows indicated by blue shades. Colder colors indicate greater negativities, and red rings illustrate the frontal ROI over which amplitudes were quantified. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

average amplitude in the 300−500 ms window over the frontal ROI with Experiment (1, 2, & 3) as a between-subject factor and Condition (True vs. False Belief) as a within-subject factor. We found only the main effect of Condition to be significant, $F(1, 51) = 14.6$, $p < .001$, $\omega_p^2 = .21$; neither Experiment, $F(2, 51) = 2.84$, $p = .068$, $\omega_p^2 = .07$, nor their interaction was, $F(2, 51) = 0.23$, $p = .79$, $\omega_p^2 = .03$. This outcome confirms that a frontal effect was evoked across all three experiments to a similar degree.

Since the frontal effect appeared to have an early onset, we explored whether the social N400 effect in Experiment 1, when it was present, started earlier as well. An exploratory statistical analysis revealed that the Incongruent-for-Observer condition ($M = –12.7$ μV, $SD = 6.32$ μV) evoked a greater negativity than the Congruent-for-Both condition ($M = –9.83$ μV, $SD = 6.11$ μV), $t(17) = –2.53$, $p = .022$, 95 % CI [–5.34 μV, –0.48 μV], Hedges' $g_{av} = 0.45$, in the 300−600 ms time window. In order to see whether this early onset is specific to the current study, we reanalyzed the data reported by Forgács et al. (2019) in the 300−600 ms time window, and we found a significant difference again between the Incongruent-for-Observer ($M = –12.5$ μV, $SD = 7.43$ μV) and the Congruent-for-Both conditions ($M = –7.66$ μV, $SD = 6.87$ μV), $t(17) = –2.59$, $p = .019$, 95 % CI [–8.84 μV, –0.90 μV], Hedges' $g_{av} = .65$ (Fig. 2). The early onset of both the N400 and the frontal effect could be related to the high predictability of events in all of our studies: in Experiment 1 and 2 infants could always expect that the object in front of them is going to be labelled correctly, whereas in Experiment 3 that it is always going to be labelled incorrectly.

These exploratory analyses suggested a frontal effect of the attributed belief, which appeared in different polarity and in an earlier time window (300−500 ms) than the one reported previously (700−1000 ms). This outcome was surprising as the experimental protocol was entirely identical with the original study, except that the replication was run in Hungarian, and in the Incongruent-for-Observer trials the two objects never came from the same superordinate category. It is not highly likely that cultural differences played a significant role in this discrepancy, but we cannot exclude such a possibility at this point. One further speculative explanation for this difference is that Hungarian is a language with free word order, and as a consequence, any word can start an utterance or sentence. In French, nouns with no articles cannot possibly start a grammatical sentence. This linguistic difference could prompt Hungarian babies to listen to noun labels with an expectancy of additional words to follow, and handle their predictions, which could involve the mental state of the observer, differently. Cross-cultural comparative studies might be necessary to provide an explanation for the observed difference.

## 6. General Discussion

If words have meaning, how is it possible to misunderstand them? And if it is possible to misunderstand words, how is it possible to learn them? We carried out three EEG experiments to better understand the way young language users understand others' misunderstanding, which is a practical implementation of the broader set of real life scenarios, where mentalization might enter linguistic communication (beyond false beliefs and semantic incongruities). With the approach of inference-based communication in mind, words might gain their contextual meaning in the cognitive environment of speakers' intentions (Sperber and Wilson, 1986), and misunderstandings could be due to misreading mental states. Disambiguation of meaning based on intentions of social partners could significantly contribute to the rapid pace of word and language learning humans exhibit during infancy. In the present study, we found an N400 effect that would not have been predicted by current models of this ERP response. Rather, this effect

**Table 1**
Summary of the results of the current study and that of Forgács et al. (2019).

| Experiment | Condition (Object congruity) | False Belief | Semantic Incongruity for Infant | Semantic Incongruity for Observer | N400 Response | Frontal Response |
|---|---|---|---|---|---|---|
| Forgács et al. (2019) Typical N400 | Congruent for Both | | | | | |
| | Incongruent for Both | | ✓ | ✓ | ✓ | |
| Forgács et al. (2019) Social N400 & Replication (Exp. 1) | Congruent for Both | | | | | |
| | Incongruent for Observer | ✓ | | ✓ | ✓ | ✓ |
| Within Category Change (Exp. 2) | Congruent for Both | | | | | |
| | Incongruent for Observer | ✓ | | | | ✓ |
| Inverse Manipulation (Exp. 3) | Incongruent for Both | | ✓ | ✓ | ✓ | |
| | Congruent for Observer | ✓ | ✓ | | ✓ | ✓ |

seems to be related to the detection of any semantic incongruity within a communicative situation, occurring either from first-person or third-person perspective. We also identified a positive going brain wave over frontal brain regions of infants whenever a label was applied to an object about which and adult Observer held a false belief. The findings of our experiments are summarized in Table 1.

Under the classical interpretation of the N400 (Brouwer et al., 2012; Kutas and Federmeier, 2011), we should not have observed an N400 effect in Experiment 1 (and in Forgács et al., 2019), because there must have been sufficient priming in the semantic system of the infants to effortlessly retrieve the label of an object right in front of them. The set of findings rather suggests that this system was also engaged when the mental state of the Observer was different from that of the infant (the social N400 effect). This was not the case in Experiment 2, where we observed no N400 response, suggesting that referent violation was not sufficient to evoke the social N400. This finding conforms to the classical account. However, if the social N400 tracks the comprehension of a communicative partner, we should have observed a reduced N400 response in Experiment 3, since infants most likely represented the Observer's belief. The content of this representation should have primed the semantic system of infants, which should have led to a reduced N400 amplitude. According to our observations, it did not, which suggests that infants produced an N400 response whenever themselves, the Observer, or both of them experienced a semantic incongruity.

Based on the pattern of results the N400 effect could be better understood as a *communicational expectancy indicator*, which fires up whenever linguistic expectations of any participant in a communicative situation are not met – irrespective of memory retrieval demands. In communication (or perhaps even in other social contexts as well), when infants track others beliefs and might have more than one representation of a referent, the N400 response is evoked if the utterance is in conflict with their own or others' (tracked) representation. In Experiment 1 it was only the Observer's predictions that were not met and the N400 marked her mismatch; in the Congruent-for-Observer-only condition of Experiment 3 infants' predictions were not met and even though Observer's were, the amplitude of the N400 response was not reduced. Alternatively, infants could exhibit an N400 whenever their own semantic representation, which they could take to be the default, is in mismatch with that of another person. Under such a scenario the source of the N400 would be a semantic mismatch detected between the representation of the infant and the Observer (or as a matter of fact, the speaker of the audio playback) in all experiments. In the Incongruent-for-Both condition of Experiment 3 the effect would be driven then by the mismatch between the infant and not the Observer but the playback voice, instead of a prediction failure. Further studies are necessary to determine if communicational expectancy or mismatch detection provide a better explanation for the phenomenon. Either way, the results appear to conform with the idea that we utilize identical systems to interact with the world (when using language) and when attributing mental states to social partners.

Importantly, the above semantic expectancy indicator explanation is not based on common ground (Clark et al., 1983) in the sense that it is not the last commonly seen object that serves as the basis of interaction (Moll et al., 2008). In Experiment 3, if the label had been compared to the first object, which was experienced together, there would have been no N400, because the two matched in the Congruent-for-Observer condition. The effect does not seem to be driven solely by the mental representations computed for the Observer either. Infants seem to track the comprehension of a communicative partner in parallel to their own comprehension.

In our exploratory analyses we observed an early starting, positive going frontal wave across all three experiments whenever infants could have assumed that the Observer was misrepresenting the labelled object. The appearance of the frontal effect in Experiment 2, independent from the N400 effect, excludes the possible interpretation of Experiment 1 that it merely disambiguated the ownership of the N400. Rather, it suggests that the content of the false belief attributed to a social partner carries information about particular object tokens, not merely about the object kind – or, at minimum, mark that the attributed belief was not formed on the basis of the present exemplar. Such an interpretation is possible only when the frontal responses are taken into consideration. The frontal positivity accompanied an N400 when the Observer encountered an incongruity, the infants did not (Experiment 1), and when only the infants experienced incongruity, the Observer did not (Experiment 3), but it was present also when there was no N400 and no semantic incongruity, only the Observer's false belief regarding the token identity of the object (Experiment 2). The only overlap in these experiments is that infants could have detected that the Observer's belief about the labelled object must have been different from their own. This frontal positivity seems to be a potential candidate for an amodal component of the belief attribution processes that specify and inform domain specific representations activated at the label onset.

## 7. Conclusions

Taken together, we replicated our finding that 14-month-olds follow the comprehension of a communicative partner and process that their social partner could have misunderstood an object labeling, by producing an N400. This ERP component seems to be sensitive to the social context and might function as a communicational expectancy indicator: it shows up whenever any interlocutors' semantic expectations are not met. Finally, a frontal positivity appeared whenever the belief state of infants and their social partners about a referent of a label were in conflict, suggesting that belief attribution functions are fully at work in 14-month-olds and provide input for meaning making in social context.

## Declaration of Competing Interest

The authors have no competing interest to declare.

## Acknowledgements

This work was supported by the postdoctoral fellowship of the Hungarian Academy of Sciences (462003) and an NKFIH Young Researcher Grant (125417) to Bálint Forgács, a HFSP Young Investigator Grant (0073/2014) and an ERC Consolidator Grant ("BabyRhythm" 773202) to Judit Gervain, and an ERC Synergy Grant (609819) to György Gergely. Eugenio Parise is a researcher of the LuCiD team at Lancaster University; the support of the Economic and Social Research Council (ES/L008955/1) is gratefully acknowledged.

## References

Bergelson, E., Swingley, D., 2012. At 6–9 months, human infants know the meanings of many common nouns. Proc. Nat. Acad. Sci. U.S.A 109 (9), 3253–3258.
Brouwer, H., Fitz, H., Hoeks, J., 2012. Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. Brain Res. 1446, 127–143.
Clark, H.H., Schreuder, R., Buttrick, S., 1983. Common ground at the understanding of demonstrative reference. J. Verbal Learning Verbal Behav. 22 (2), 245–258.
Forgács, B., Parise, E., Csibra, G., Gergely, G., Jacquey, L., Gervain, J., 2019. Fourteen-month-old infants track the language comprehension of communicative partners. Dev. Sci. e12751.
Friedrich, M., Friederici, A.D., 2005. Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. Neuroreport 16 (6), 653–656.
Friedrich, M., Friederici, A.D., 2008. Neurophysiological correlates of online word learning in 14-month-old infants. Neuroreport 19 (18), 1757–1761.
Gervain, J., Macagno, F., Cogoi, S., Peña, M., Mehler, J., 2008. The neonate brain detects speech structure. Proc. Nat. Acad. Sci. U.S.A 105 (37), 14222–14227.
Grice, H.P., 1957. Meaning. Philosoph. Rev. 66 (3), 377–388.
Jouravlev, O., Schwartz, R., Ayyash, D., Mineroff, Z., Gibson, E., Fedorenko, E., 2019. Tracking colisteners' knowledge states during language comprehension. Psychol. Sci. 30 (1), 3–19.
Junge, C., Cutler, A., Hagoort, P., 2012. Electrophysiological evidence of early word learning. Neuropsychologia 50 (14), 3702–3712.
Kampis, D., Parise, E., Csibra, G., Kovács, A.M., 2015. Neural signatures for sustaining object representations attributed to others in preverbal human infants. Proc. R. Soc. B 282 (2015), 1683.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.
Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. 4, 863.
Molinaro, N., Barber, H.A., Carreiras, M., 2011. Grammatical agreement processing in reading: ERP findings and future directions. Cortex 47 (8), 908–930.
Moll, H., Richter, N., Carpenter, M., Tomasello, M., 2008. Fourteen-Month-Olds Know What "We" Have Shared in a Special Way. Infancy 13 (1), 90–101.
Parise, E., Csibra, G., 2012. Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. Psychol. Sci. 23 (7), 728–733.
Rueschemeyer, S.A., Gardner, T., Stoner, C., 2015. The Social N400 effect: how the presence of other listeners affects language com- prehension. Psychon. Bull. Rev. 22 (1), 128–134.
Southgate, V., Vernetti, A., 2014. Belief-based action prediction in preverbal infants. Cognition 130 (1), 1–10.
Sperber, D., Wilson, D., 1986. Relevance: Communication and Cognition. Blackwell, Oxford.
Stahl, D., Parise, E., Hoehl, S., Striano, T., 2010. Eye contact and emotional face processing in 6-month-old infants: advanced statistical methods applied to event-related potentials. Brain Dev. 32 (4), 305–317.
Vouloumanos, A., Martin, A., Onishi, K.H., 2014. Do 6-month-olds understand that speech can communicate? Dev. Sci. 17 (6), 872–879.
Westley, A., Kohút, Z., Rueschemeyer, S.A., 2017. "I know something you don't know": Discourse and social context effects on the N400 in adolescents. J. Exp. Child Psychol. 164, 45–54.