Cognition 258 (2025) 106084

Contents lists available at ScienceDirect

Cognition

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

Abstractness impacts conversational dynamics

Claudia Mazzuca [a,*,1], Caterina Villani [b,1], Tommaso Lamarra [b], Marianna Marcella Bolognesi [b], Anna M. Borghi [a,c]

[a] Department of Dynamic and Clinical Psychology, and Health Studies, Sapienza University of Rome, via degli Apuli 1, Rome 00185, Italy
[b] Department of Modern Languages, Literatures, and Cultures, University of Bologna, via Cartoleria 5, Bologna 40124, Italy
[c] Institute of Cognitive Sciences and Technologies, Italian National Research Council, via Romagnosi 18/A, Rome 00196, Italy

ARTICLE INFO

Keywords:
Abstractness
Conversation
Interaction
Abstract sentences

ABSTRACT

Conversation topics may vary in abstractness. This might impact the effort required by speakers to reach a common ground and, ultimately, an interactive alignment. In fact, people typically feel less confident with abstract concepts and single-words rating studies suggest abstract concepts are more associated with social interactions than concrete concepts—hence suggesting increasing levels of abstractness enhance inner and mutual monitoring processes. However, experimental studies addressing conversational dynamics afforded by abstract concepts are still sparse. In three preregistered experiments we ask whether abstract sentences are associated with specific constructs in dialogue, i.e., higher uncertainty, more curiosity and willingness to continue a conversation, and more questions related to causal and agency aspects. We do so by asking participants to evaluate the plausibility of linguistic exchanges referring to concrete and abstract concepts. Results support theories proposing that abstract concepts involve more inner monitoring and social dynamics compared to concrete concepts and suggest that reaching alignment in dialogue is more effortful with abstract than with concrete concepts.

1. Introduction

Humans are exceptionally good at talking to each other. Indeed, it has been proposed that we are hard-wired for dialogue (Pickering & Garrod, 2021). To be successful, a conversation rests upon the shared understanding of the interlocutors, that reach a common ground through a cooperative joint activity (Clark, 1996; Pickering & Garrod, 2021).

Although language seems to be optimised for communication (Fedorenko, Piantadosi, & Gibson, 2024), and conversations generally flow smoothly (Galantucci & Roberts, 2014), miscommunication in social interactions is rather frequent (Healey, De Ruiter, & Mills, 2018). Language is a complex system, in which mutual understanding is co-built iteratively and interactively in conversations through the negotiation of meanings and the constant monitoring of common ground (Dingemanse, 2017). Overall, clarifications requests like "Huh?" or "Who?" have been estimated to occur on average every 84 *s* in ordinary conversations (Enfield, 2017), and cross-linguistic findings showed that these requests are all but uncommon, occurring about once every 1.4 *min* across targeted languages (Dingemanse et al., 2015; see also Fay et al., 2018). Furthermore, research on sign languages and gesture-mediated communication highlighted these employ similar mechanisms for communicating conversation troubles (Manrique & Enfield, 2015; Skedsmo, 2020). Linguistic markers of this kind are known as interactive repairs, and have been proposed to be a universal communication device, foundational to mutual understanding (Dingemanse & Enfield, 2024; Dingemanse & Enfield, 2015).

Across spoken languages, interactive repairs have been proposed to take mainly three forms. Open requests point to general trouble with understanding and request for clarifications (e.g., "Huh?"); restricted requests ask for specification or precise clarification of a specific component of the sentence (e.g., "Who?"); restricted offers ask for confirmation of what has just been said, offering a potential candidate answer (e.g., "She had a boy?") (Dingemanse et al., 2015; Digemanse & Enfield, 2024). To indicate there are no mutual understanding problems, and the recipient is paying attention, instead, speakers often employ backchannels (Gardner, 2001; Yngve, 1970). These usually consist of short words like "ok", "uh-uh", and can also act as "continuers"

* Corresponding author at: Via degli Apuli 1, 00185 Rome, Italy.
E-mail address: claudia.mazzuca@uniroma1.it (C. Mazzuca).
[1] Shared first co-authorship

https://doi.org/10.1016/j.cognition.2025.106084
Received 12 September 2024; Received in revised form 7 February 2025; Accepted 8 February 2025
Available online 14 February 2025

(Schegloff, 1982) to make the interlocutor aware that the recipient expects more information. While interactive repairs and backchannels seem to underpin speakers' ability to sustain a conversation, their use might vary depending on specific conversation characteristics (Dideriksen, Christiansen, Tylén, Dingemanse, & Fusaroli, 2023), among which—perhaps—semantic aspects.

Conversational topics differ and some of them might require more effort. For instance, converging evidence indicates abstract words like "democracy" are *harder* to learn than concrete words like "hammer" (Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). Overall, abstract concepts and words are processed slower and less accurately than concrete concepts (Kroll & Merves, 1986; Paivio, 1991; Schwanenflugel, 1991; Schwanenflugel, Harnishfeger, & Stowe, 1988), are acquired later (Gilhooly & Logie, 1981; Della Rosa et al., 2010; Villani, Lugli, Liuzza, & Borghi, 2019; Bellagamba et al., 2022), they often collect under a single label multiple situations and experiences rather than pointing to a single, bounded referent (Davis, Altmann, & Yee, 2020; Villani, Loia, & Bolognesi, 2024), and people typically feel less confident in mastering their meaning and more prone to negotiate it with others (Borghi, 2022; Falcinelli et al., 2024; Falcinelli, Fini, Mazzuca, & Borghi, 2024; Fini et al., 2023; Mazzuca, Falcinelli, Michalland, Tummolini, & Borghi, 2022; Mazzuca & Santarelli, 2023). In addition, a wide body of evidence shows people need more contextual information to understand abstract words compared to that needed for concrete words (Schwanenflugel & Stowe, 1989; Schwanenflugel, Akin, & Luh, 1992). Consistently, research shows that abstract concepts are processed and represented differently from concrete concepts (Banks et al., 2023; Binder, Westbury, McKiernan, Possing, & Medler, 2005; Bolognesi & Steen, 2018; Borghi, 2023; Conca, Borsa, Cappa, & Catricalà, 2021; Dove, 2022; Henningsen-Schomers & Pulvermüller, 2022; Mazzuca et al., 2021). These differences might be the result of differing acquisition pathways, and the scientific literature in this domain has emphasised distinct aspects that might drive abstract word acquisition, such as emotional valence, interoception, and social interaction (Wauters, Tellings, Van Bon, & Van Haaften, 2003; Ponari, Norbury, & Vigliocco, 2018; Connell, Lynott, & Banks, 2018; Borghi, 2023; Dove, 2022, Reilly et al., 2024). Whatever the specific mechanisms supporting the acquisition of abstract words, abstract language is ubiquitous—at least in English (Lupyan & Winter, 2018). However, the relationship between conversation dynamics and abstractness is still largely underexplored, with a few exceptions.

It has been proposed that whereas in the case of concrete concepts speakers can reach an alignment more easily and simply by attending to a perceptual entity or by recalling it, with abstract concepts they need more "mutual monitoring" in conversations (Gandolfi, Pickering, & Garrod, 2023). Along similar lines, research focusing on the interplay between conversational devices and conversational contexts suggests the use of specific strategies might vary depending on the precision of information required by the topic. To illustrate, Dideriksen et al. (2023) showed that affiliative conversations (e.g., conversations about TV series or roommate experiences) afforded more backchanneling and less repair than task-oriented conversations (elicited through symmetric and asymmetric cooperative games). In addition, there were also differences in the specificity of repair afforded by the two types of conversations, with the former eliciting more open repairs, and the latter eliciting more specific repairs. On top of that, research showed that other conversational devices that are thought to sustain dialogical interactions like linguistic alignment (i.e., the entrainment of linguistic behaviours between speakers, Pickering & Garrod, 2004) are differently modulated by the conversation goals (see Fusaroli, Rączaszek-Leonardi, & Tylén, 2014). Against this background, one might hypothesise that conversations with varying levels of abstractness might elicit different and specific conversational devices.

Consistent with this, in a study where participants were asked to imagine being involved in a real conversation and to respond to target sentences composed of different kinds of abstract and concrete concepts, responses to abstract and concrete sentences qualitatively differed (Villani, Orsoni, Lugli, Benassi, & Borghi, 2022). Specifically, abstract sentences elicited more expressions of uncertainty (e.g., "mmmh", "how is that?"), questions aimed at knowing more about the topic (e.g., "tell me more"), and more *why* and *who* questions compared to concrete sentences. Concrete sentences, instead, elicited less uncertain expressions, fewer questions aimed at knowing more, and more *where* and *when* questions. In other words, abstract sentences afforded more open repairs and backchannelling than concrete sentences. Restricted requests, on the other hand, differed across types of sentences in terms of the information they targeted, with abstract sentences eliciting more clarifications on causal mechanisms and concrete sentences eliciting more clarifications on spatiotemporal aspects. So, abstract concepts indirectly hinted at interpersonal mechanisms rather than pointing to locations and time. Collectively, these results have been explained suggesting that abstract concepts, in general, evoke more social validation than concrete concepts. To illustrate, we might feel the need for other people's aid to understand—or simply to frame—better an abstract sentence compared to a concrete sentence (e.g., Borghi, Barca, Binkofski, & Tummolini, 2018).

Rating studies on single words support this intuition, highlighting the entwinement between abstract concepts and social interaction. On average, more abstract words score higher in Social Metacognition, i.e., the need to rely on others to understand word meanings, than more concrete words (Borghi et al., 2018; Villani et al., 2019). Along the same lines, people believe their meaning is more debatable, and they are more open to negotiate it with others (Fini, Era, et al., 2023; Mazzuca & Santarelli, 2023). In addition, a recent norming study broadening the notion of *socialness* (Diveica, Pexman, & Binney, 2023) and collecting ratings for more than 8000 English words found that socialness is negatively correlated with concreteness, Body-Object-Interaction (i.e., the extent to which a human body can physically interact with the referent of a concept, Tillotson, Siakaluk, & Pexman, 2008), and Imageability—hence suggesting that more concrete words are considered as less relevant on the social dimension. Additionally, abstract concepts are consistently more associated with internal, mental, and interoceptive states (Barsalou & Wiemer-Hastings, 2005; Connell et al., 2018; Villani, Lugli, Liuzza, Nicoletti, & Borghi, 2021), whereas concrete concepts are typically more associated with perceptual and sensorimotor states (Lynott, Connell, Brysbaert, Brand, & Carney, 2020). Finally, recent evidence suggests people prefer starting a conversation with abstract concepts—more specifically, with emotional concepts and concepts related to the self and sociality—than with concrete concepts (Fini, Era, et al., 2023). Along similar lines, a recent study analysing conversations shows that what makes a good starting topic in conversations is the extent to which a topic is connected to other topics (Templeton, Chang, & Wheatley, 2024). Topics that resulted to be very well-connected with other topics in the study, therefore qualifying as optimal conversation starters, include "hometowns", "playing sports", "classes", and "exams"—all topics somehow related to the self and sociality. Taken together, these findings underline the importance of social interaction and mutual monitoring in abstract concepts' representation.

However, the extent to which this pattern of results can be generalised remains an open question. Indeed, while the advantages of single-word methods for studying concepts are well-known (see Banks et al., 2023 for an overview), they also present several downsides. Norming studies might, in fact, conceal important conceptual information that presenting a word in isolation fails to convey. Concepts are not stable and immutable entities in the minds of speakers, but their representation is flexibly tailored to specific situations (Lebois, Wilson-Mendenhall, & Barsalou, 2015; Yee & Thompson-Schill, 2016). This is particularly true for abstract concepts, which due to their indeterminate character are more variable across people and cultures (Borghi & Mazzuca, 2023; Mazzuca et al., 2024; Mazzuca, Majid, Lugli, Nicoletti, & Borghi, 2020; Wang & Bi, 2021), but also across contexts (Hoffman, 2016; Schwanenfluegel & Stowe, 1989; Davis et al., 2020). Along these lines, McRae,

Nedjadrasul, Pau, Lo, and King (2018) found that lexical decisions for abstract words were facilitated when participants were first presented with images depicting related—rather than unrelated—situations, thereby confirming the importance of contextual support for abstract concepts' processing. These aspects are even more critical when studying words embedded in conversations, as words do not exist in a vacuum but are always entrenched in a context—most prominently, a linguistic one.

In keeping with recent perspectives, online interactive methods eliciting social and linguistic interactions are best suited for tackling abstract conceptual knowledge (Banks et al., 2023; Borghi, 2023; Borghi, Osińska, Roepstorff, & Raczaszek-Leonardi, 2023). While steps towards this direction have already been taken (see e.g., Fini et al., 2023; Fini, Era, Da Rold, Candidi, & Borghi, 2021; Zdrazilova, Sidhu, & Pexman, 2018), interactive paradigms pose a further whole set of methodological issues that cannot be overlooked. The first, and perhaps more important concern relates to experimental control. Freeing participants from strict experimental constraints when asking them to chat about abstract or concrete concepts, for instance, can lead to sudden shifts in topics and topics' abstractness—often scarcely informative for the research questions at stake (Fini, Falcinelli, et al., 2023). On the other hand, linguistic production studies that can be considered intermediate in terms of constraints imposed on speakers, like the one devised by Villani et al. (2022), might only reflect speakers' metacognitive awareness of the appropriateness of linguistic exchanges. Altogether, the challenges related to norming, interactive, and linguistic production studies call for a tailored, flexible approach to the investigation of abstract conceptual knowledge that adapts to the needs of specific research questions and predictions.

Here we take this stance building on previous findings from Villani et al. (2022) with the aim of assessing their generalisability in a more controlled experimental setting. We do so by investigating the relation between abstractness and sociality using carefully controlled sentences and asking speakers to evaluate the plausibility of different types of follow-up sentences as quickly and accurately as they could. We are aware that this procedure might not be considered as fully interactive, as judging a sentence differs from producing one in a real conversation. However, we aimed to push one step further the study of abstract semantic knowledge by not relying on single words while at the same time overcoming the limitations of more interactive paradigms. If abstract sentences evoke more uncertainty and clarification expressions and, to some extent, more social interaction than concrete sentences, then this should also be observed in behaviour-related measurements.

1.1. The present study

This study tackles these questions with three different experiments, preregistered at https://osf.io/erxd6. Ethical approval was provided by the Ethics Committee of the Department of Dynamic, Clinical Psychology and Health of Sapienza University of Rome [Prot. n. 0000688]. Each experiment targets a specific aspect of the relation between abstractness and social interactions leveraging on excerpts of simulated conversations. Participants are presented with sentences varying in abstractness, followed by three different types of possible follow-up expressions or questions, and are asked to judge whether the sentence–follow-up combinations are plausible linguistic exchanges as if they were embedded in a real conversation. Follow-ups represent expressions of uncertainty vs. expression of certainty (Experiment 1); expressions of curiosity/signalling the willingness to know more vs. the willingness to end the conversation (Experiment 2); *why* and *who/for whom* questions vs. *where* and *when* questions (Experiment 3). The choice of these three categories of follow-up expressions is twofold. From a theoretical perspective, they can be considered proxies for aspects that are known to characterise abstractness. Specifically, people feel more uncertain about the meaning of abstract concepts compared to that of concrete concepts (Borghi, 2022; Falcinelli, Fini, Mazzuca, Alessandri, et al., 2024; Falcinelli, Fini, Mazzuca, & Borghi, 2024; Fini, Era, et al., 2023; Mazzuca et al., 2022); they prefer starting a conversation with abstract rather than concrete concepts (Fini, Era, et al., 2023), and abstract concepts' meaning is more often the remit of negotiation (Falcinelli, Fini, Mazzuca, & Borghi, 2024; Mazzuca & Santarelli, 2023)—hence possibly hinting at more social interactions and curiosity; finally, while sensorimotor and perceptual components seem to be implicated in the conceptual representations of both kinds of concepts (abstract vs concrete; Kiefer & Harpaintner, 2020; Banks & Connell, 2023), abstract concepts evoke most consistently introspective and mental states aspects, as well as social situations (Barsalou & Wiemer-Hastings, 2005). From a methodological perspective instead, we selected expressions that were found to be most relevant for abstract and concrete sentences in Villani et al. (2022), thereby aiming at increasing the generalisability of previous findings in conversations.

We collected response times and sentence plausibility judgments. In addition, at the end of each experimental session, participants from each experiment were asked to rate on 7-point Likert scales the sentences they had previously seen in terms of how confident they were in their understanding of the sentence (i.e., Confidence; Mazzuca et al., 2022; Fini, Era, et al., 2023); the extent to which they would have relied on experts' knowledge if they could to understand the sentence (i.e., Experts Need; Falcinelli, Fini, Mazzuca, & Borghi, 2024); and the extent to which they thought they needed other people's input to understand the sentence (i.e., Social Metacognition; Borghi, 2022; Villani et al., 2019).

We predicted that participants would find more plausible follow-ups signalling uncertainty, curiosity, and causal-agent questions for abstract sentences. By contrast, they would find more plausible follow-ups signalling certainty, willingness to end the conversation, and spatio-temporal questions for concrete sentences. We also expected consistent differences in RTs as a function of sentence type – follow-up combinations, although we did not predict any specific direction of the effect. Finally, in line with the evidence reviewed in the previous sections, we predicted participants would give lower Confidence scores and higher Experts Need and Social Metacognition scores for abstract sentences compared to concrete sentences.

It is noteworthy that, while this study has its own relevance for the literature on the link between abstractness, conversations, and social interaction, it also constitutes a conceptual replication (Schmidt, 2009) of a previous production study (Villani et al., 2022). In the remainder of the article, we first describe the two norming studies that were implemented to generate stimuli for the three experiments, and then we zoom in on each experiment. Data, analyses scripts, and materials for the norming studies as well as for the three experiments are publicly available on the Open Science Framework project repository at https://osf.io/jf7vk.

2. Norming studies

2.1. Validation of target sentences

To generate the stimuli for the experiments, we conducted two separate norming studies, implemented as online questionnaires on Qualtrics. The first norming study was aimed at selecting suitable sentences for the following experiments. So, we selected 60 sentences from Villani et al. (2022), and asked a sample of 38 participants (27 women, 10 men, 1 "other"; *M* age = 25.81, *SD* = 5.98, Age range = 19–54) to rate them on 7-point Likert scales in terms of concreteness~abstractness (1 = completely concrete; 7 = completely abstract), how much they thought sentences needed a context to be understood (Context Support: 1 = not at all; 7 = very much), how much they thought each sentence was "natural" (Naturalness: 1 = not at all; 7 = very much), and Familiarity (1 = not at all; 7 = very much). In addition to that, we also included 12 filler sentences (e.g., "*I decorated the lion*") for the Naturalness dimension, that were used as a control to determine the reliability of participants' judgments. Sentences were matched for their

morphological complexity and included only subject, verb, and noun (e.g., *"I made a cake", "I made a judgment"*). Each dimension was presented in a separate block, and blocks were randomised across participants. Sentences were also randomised within blocks.

Our first aim was to assess which sentences could be used for the study based on Naturalness ratings, using a score of 4 as a cut-off criterion. Descriptive statistics indicated that all 60 sentences were rated on average > 4 in Naturalness. We used ordinal mixed effects models to model Naturalness rating scores as a function of Category of the sentence (abstract vs concrete) with participants and sentences as random intercepts, using the clmm() function from the "ordinal" R's package (Christensen, 2023). Results show that concrete sentences were perceived as more natural than abstract sentences, $b = 1.007$, SE = 0.307, $z = 3.274$, $p = .001$. To better balance our stimuli and keep only highly sensible sentences, we excluded sentences with the lowest scores in Naturalness, taking care to have the same number of sentences across concrete and abstract subcategories (e.g., animals, foods, tools, foods, philosophical/spiritual, physical/space-time/quantity, and emotional/mental states/social concepts, see Villani et al., 2019). Following this procedure, we removed 3 sentences for each subcategory, resulting in a final dataset of 42 sentences, including 21 abstract sentences and 21 concrete sentences. Within this set of sentences, concrete sentences ($M = 6.50$; $SD = 1.22$) were still perceived as more natural than abstract sentences ($M = 6.29$; $SD = 1.32$), $b = 0.99$, SE = 0.338, $z = 2.931$, $p = .003$.[2]

The second aim of the first norming study was to ensure that the two types of sentences were balanced in terms of abstractness~concreteness, Familiarity, and Context Support. Specifically, we aimed for abstract and concrete eligible sentences differing in abstractness but not in Familiarity and Context Support. To this extent, we modelled rating scores for each dimension with ordinal mixed models as a function of the type of sentence (abstract vs. concrete) with random intercepts for participants and sentences. In the Context Support model, we encountered issues in the estimation of the terms, possibly due to the lack of a sufficient number of observations for each level of the thresholds (i.e., rating points), so we reverted to simple ordinal regressions. We found that abstract ($M = 4.89$; $SD = 2.07$) and concrete sentences ($M = 4.91$; $SD = 2.25$) do not differ for Familiarity, b concrete = 0.22, SE = 0.442, $z = 0.506$, $p = .613$. Importantly, sentences differ in abstractness, with concrete sentences ($M = 1.77$; $SD = 1.80$) being significantly less abstract than abstract sentences ($M = 4.38$; $SD = 1.77$), $b = -3.687$, SE = 0.247, $z = -14.88$, $p < .0001$. However, concrete sentences needed less Context Support ($M = 1.93$; $SD = 1.40$) than abstract sentences ($M = 4.13$; $SD = 2.08$) to be understood, $b = -2.148$, SE = 0.108, $z = -19.82$, $p < .0001$, in line with context-availability literature (Schwanenflugel et al., 1992; Davis, Altmann, & Yee 2020). Nonetheless, we decided to keep this set of sentences as they resulted to be carefully balanced on Familiarity and crucially on abstractness ~ concreteness, but also on the whole perceived as distinctly natural (see Fig. 1).

2.2. Validation of follow-up sentences

The second norming study was aimed at generating follow-up sentences. To do that, we first selected 24 among the most frequently produced follow-ups to abstract and concrete sentences in Villani et al. (2022). Then, we asked a further sample of 22 participants (16 women, 5 men, 1 "other"; M age = 24.18, $SD = 3.64$, Age range = 19–32) to rate on 7-point Likert scales linguistic exchanges composed of abstract and concrete sentences randomly paired with the follow-ups in terms of (a) whether the answer signals uncertainty vs. certainty; (b) whether the answer signals curiosity vs. willingness of ending the conversation. Participants were asked to imagine attending a conversation between two people and hearing a portion of their dialogue in which person A uttered a sentence (e.g., "*I dreamed of heaven*", "*I used the drill*"). They were presented with a list 24 possible answers of B in a randomised order, and were asked to rate on a 7-point Likert scale: (1) whether the answer signals certainty or uncertainty (1 = very certain, 7 = very uncertain); (2) whether the answer signals the willingness to end or continue the conversation, (1 = end, 7 = continue). In order to identify suitable follow-ups to be used in Experiment 1 (certainty vs. uncertainty) and Experiment 2 (end vs. curiosity), we calculated the mean score of each follow-up sentence on the two dimensions. Based on descriptive statistics, we identified 4 follow-ups for each dimension, corresponding to the highest and lowest values of the scale. For Experiment 1, we selected "*I did not understand*" (M Certainty~Uncertainty = 5.45; $SD = 1.84$) and "*What do you mean?*"(M Certainty~Uncertainty = 4.40; $SD = 1.59$) as uncertainty follow-ups, and "*Good job*" (M Certainty~Uncertainty = 2.09; $SD = 1.65$) and "*You did well*" (M Certainty~Uncertainty = 2.27; $SD = 1.51$) as certainty follow-ups. For Experiment 2, we selected "*Ok*" (M End~Curiosity = 1.77; $SD = 1.44$) and "*Thanks*" (M End~Curiosity = 2.36; $SD = 1.50$) as end conversation follow-ups, and "*Describe*" (M End~Curiosity = 6.41; $SD = 1.37$) and "*Tell me more*" (M End~Curiosity = 6.36; $SD = 1.33$) as curiosity follow-ups. For Experiment 3 (type of questions) we did not need to norm possible follow-up sentences, as we could use "*where?*" and "*when?*" as spatiotemporal questions and "*why*" and "*for whom?*" as questions for reasons and agents—in line with Villani et al. (2022). Target sentences are reported in Table 1, Supplementary Materials.

3. Experiment 1: abstractness and uncertainty

Experiment 1 tests whether people feel more uncertain with abstract compared to concrete concepts. We hypothesized that abstract sentences elicit more uncertainity about their possible meaning and evoke a longer monitoring process compared to concrete sentences. So, we expected participants to judge expressions of uncertainty as more plausible follow-ups for abstract sentences compared to expressions signalling that the sentence has been understood, and the opposite for concrete sentences. We also expected a difference in RTs to abstract and concrete sentences as a function of the type of follow-up.

3.1. Methods

3.1.1. Participants

Twenty-eight Italian speakers were recruited at the University of Bologna (24 females; M age = 19.32; $SD = 1.22$, age range = 18–24).[3]

3.1.2. Materials, design, and procedure

The target sentences were 21 abstract and 21 concrete sentences resulting as eligible from the first norming study (see ¶2.1), while follow-up sentences were two sentences with the highest scores on the Certainty~Uncertainty rating of the second norming study (i.e., "I did not understand"; "What do you mean"), and two sentences with the lowest scores on the same dimension (i.e., "Good job"; "Well done", see ¶2.2).

After providing their consent to participation and demographic

[2] Rating data from the norming study were originally analysed using linear mixed models (for the complete set of analyses please refer to the OSF preregistration page), that detected no significant differences in terms of Naturalness for the selected ($N = 42$) abstract and concrete sentences. However, ordinal mixed models are better suited for this kind of data, as one of the three independent reviewers suggested, so we re-analysed the data using ordinal mixed models.

[3] Sample size estimation was based on an erroneous power analysis (please refer to the OSF preregistration page of the project for details) that suggested a sample size of 26 participants for each experiment. Given that we already performed the studies, we revert to confidence intervals of the predicted interactions to get a glimpse of the credibility of our experiments (Giner-Sorolla et al., 2024).

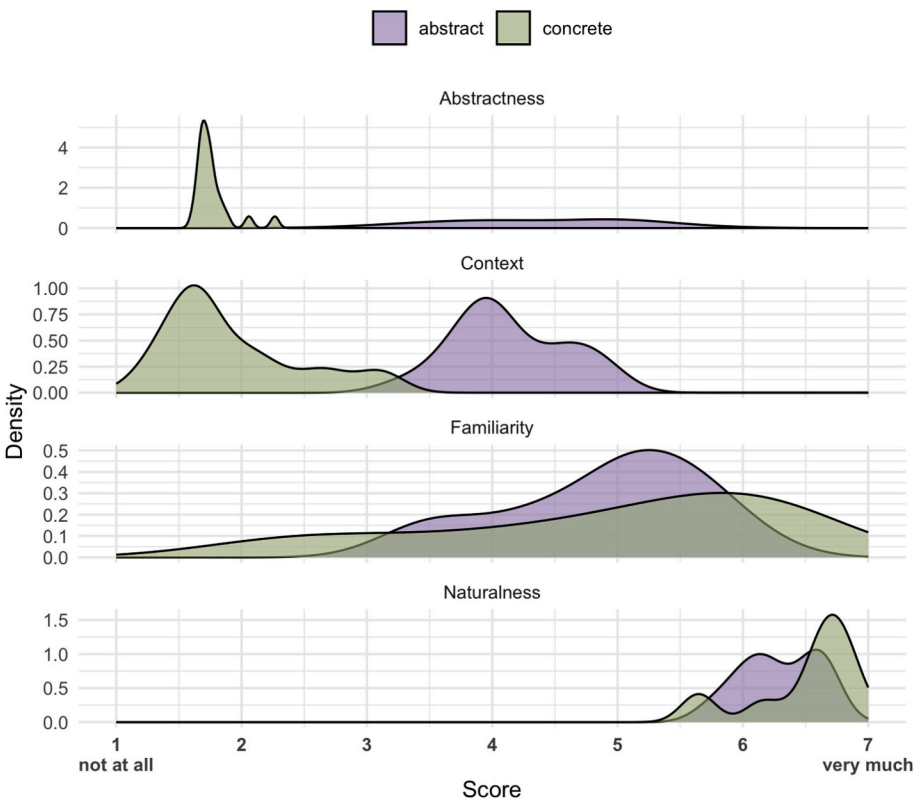

Fig. 1. Distribution of rating scores for Abstractness, Context Support, Familiarity, and Naturalness of abstract and concrete sentences.

information, participants started the experimental session. They were tested in a dimly lit room and presented with the experiment implemented in *E*-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA). In the first session, participants were presented with sentences–follow-up pairings on a computer screen in two blocks separated by a short break, and were asked to decide whether the combinations could plausibly occur in a real conversation (yes vs. no). The order of sentences was randomised across participants, ensuring that each sentence was repeated only twice within a single block. Each trial started with a central black fixation cross of 500 *ms*, followed by a sentence lasting for 1500 *ms*, then the follow-up appeared, remaining on the screen until the response was given by participants up to a maximum of 3000 *ms*. RTs were recorded from follow-up onset. Participants provided their response by pressing two keys on the keyboard (i.e., "x" and "m"), the mapping of which was counterbalanced between participants (see Fig. 2). In the second session, participants were presented with the rating task, in which they were asked to rate on 7-point Likert scales the sentences they were presented with on three semantic dimensions (Confidence, Experts Need, and Social Metacognition, see ¶ 6, Ratings).

3.1.3. Data analysis

Data analysis and visualization for all the experiments were carried out with R (R Core Team, 2019) and RStudio (v4.0.3). In keeping with the preregistration plan, data points exceeding $+/-$ 2*SD* from the average RTs of plausible and not plausible answers were excluded from the analyses as they were considered outliers. This criterion was chosen hypothesising that our RTs would approximate a normal distribution, and therefore that—following the 68–95-99.7 rule of thumb—around

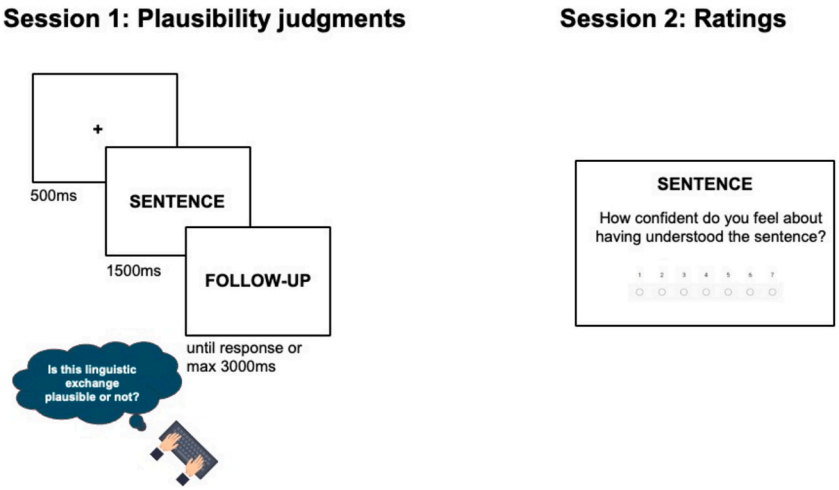

Fig. 2. Schematic depiction of the experimental procedure.

95 % of observations would fall within 2 *SD*s from the mean (see also Winter, 2019). Plausibility judgments for each experiment (1 = yes; 0 = no) were modelled with a mixed effects binomial logistic regression using glmer() function from "lme4" R's package (Bates, Mächler, Bolker, & Walker, 2015). The model featured Category (abstract vs. concrete), Type of Follow-up (Experiment 1: uncertainty vs. certainty; Experiment 2: curiosity vs. end; Experiment 3: causal vs. spatio-temporal), and their interaction as fixed effects, with participants and sentences as random intercepts. Significance of the main effects and interactions was assessed with Wald Chi-Squared tests using the Anova() function from "car" R's package (Fox & Weisberg, 2019). RTs for each experiment were first visually inspected for skewness, log-transformed and then analysed separately for "plausible" and "not plausible" responses with linear mixed models performed with the lmer() function from "lme4" R's package (Bates et al., 2015). Plausible and implausible responses were analysed separately as negation is known to produce longer RTs (see Banks et al., 2023 for a brief review). The structure of the models was identical to the one used for plausibility judgments. Post-hoc contrasts were carried out with the "emmeans" R's package (Lenth, 2021) using Tukey's adjustment for multiple comparisons. Models' assessment was performed with the "performance" R's package (Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski, 2021). All linear models resulted to comply with most of the more important assumptions for linear mixed models (i.e., homoskedasticity of residuals, normality of residuals, and linearity, see Meteyard & Davies, 2020), except for normality of residuals that was detected only in the model on plausible RTs of Experiment 1 (but see Schielzeth et al., 2020 for a demonstration of the robustness of linear mixed models to violations of distributional assumptions).

3.2. Results

In accordance with the data exclusion criterion described above, we excluded 2.8 % of datapoints ($n = 135$) from further analyses.

3.2.1. Plausibility judgments

We found a significant main effect of Category, χ^2 (1) = 4.26, p = .038, Type of Follow-up, χ^2 (1) = 972.91, $p < .001$, and a significant interaction between Category and Type of Follow-up, χ^2 (1) = 102.07, $p < .001$, $b = -1.76$, 95 % CI [−2.10, −1.42]. Post-hoc contrasts showed that for both abstract, $z = 19.687$, $p < .0001$, and concrete sentences, $z = 26.763$, $p < .0001$, participants judged certainty follow-ups as more plausible compared to uncertainty follow-ups. However, we also found that uncertainty follow-ups were considered as more plausible for abstract than for concrete sentences, $z = 6.066$, $p < .0001$, whereas there was no difference within certainty follow-ups, $p = .66$ (see Fig. 3).

3.2.2. Response times

Plausible. We found a significant main effect of Type of Follow-up, F (1, 1059.93) = 155.04, $p < .001$, showing that participants were overall slower with follow-ups signalling uncertainty (M uncertainty = 1264.09 *ms*; SD = 424.38; M certainty = 958.91 *ms*; SD = 388.75). No other main effect or interaction reached significance, all $p_s > 0.062$.

Implausible. We found a significant main effect of Type of Follow-up, F(1, 1061.05) = 27.40, $p < .0001$, and a significant two-way interaction between Category and Type of Follow-up, F(1, 1105.80) = 5.77, p = .016, $b = -0.07$, 95 % CI [−0.13, −0.01]. No other main effect reached significance (Category $p = .11$).

Post-hoc contrasts showed that participants were faster to judge as not plausible abstract sentences when they were paired with uncertainty follow-ups compared to when they were paired with certainty follow-ups, t(984) = −2.020, $p = .043$. Likewise, they were faster to judge as not plausible concrete sentences when they were paired with uncertainty follow-ups compared to when they were paired with certainty follow-ups, t(1508) = −5.390, $p < .001$. We also found that participants responded slower to abstract–uncertainty pairings than concrete–uncertainty pairings, t(67.1) = 3.748, $p < .001$. There was instead no difference within certainty follow-ups, t(187.5) = −0.5364, $p = .716$ (see Fig. 4).

3.3. Discussion

Overall, plausibility judgments show participants found follow-up sentences indicating uncertainty less plausible than those indicating certainty, regardless of the concreteness ~ abstractness of sentences. This was also reflected in longer RTs for plausible answers featuring uncertainty follow-ups compared to those featuring certainty follow-ups. However, pairings with uncertainty follow-ups preceded by abstract sentences were judged as more plausible than those preceded by concrete sentences—partially supporting our hypotheses. Indeed, participants were also slower in judging abstract–uncertainty pairings as not plausible compared to concrete–uncertainty pairings, whereas there was no difference in response times to abstract and concrete sentences followed by certainty expressions.

4. Experiment 2: abstractness and curiosity

Experiment 2 asks whether abstractness elicits more curiosity than concreteness. We hypothesized that with abstract sentences participants rely more on others to understand better their meaning or tend to ask for

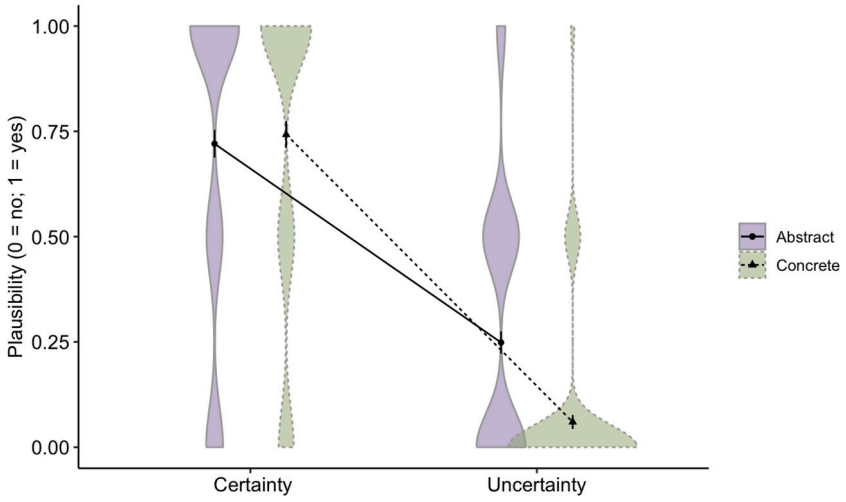

Fig. 3. Predicted probabilities of plausibility judgments for abstract and concrete sentences as a function of follow-ups.

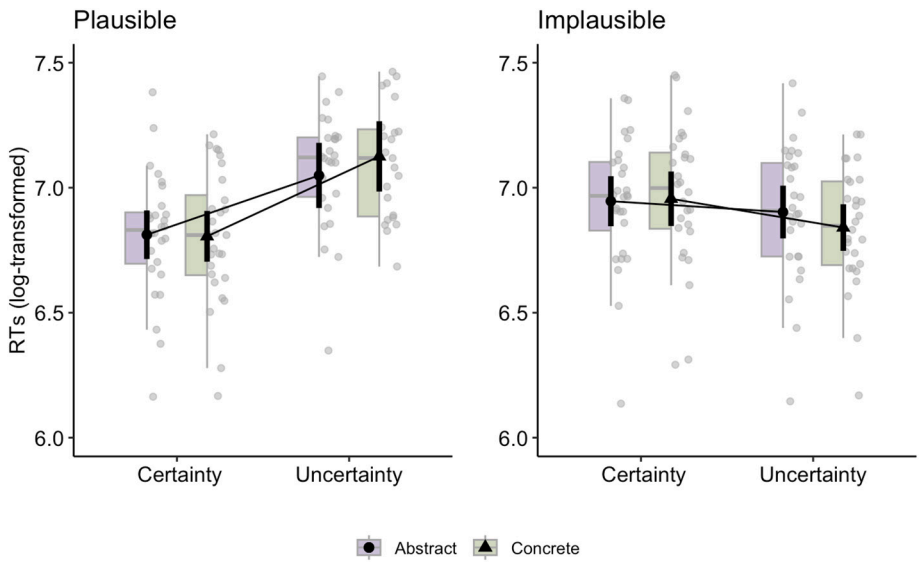

Fig. 4. Log-transformed Response Times for abstract and concrete sentences as a function of follow-ups. Raw data in the background are aggregated over participants.

more information (i.e., they are more curious) compared to concrete sentences. So, we expected participants to judge more plausible for abstract sentences expressions signalling curiosity, or the need for more specifications compared to expressions signalling the willingness to end the conversation, and the opposite for concrete sentences. We also expected a difference in RTs to abstract and concrete sentences as a function of the type of follow-up.

4.1. Methods

4.1.1. Participants

Twenty-eight Italian speakers were recruited at the University of Bologna (21 females; *M* age = 20.96; *SD* = 2.52, age range = 18–29).

4.1.2. Materials, design, and procedure

The procedure, design, and data analysis were identical to those of Experiment 1. Materials were also identical, except for the follow-up sentences that were two sentences with the highest scores on the End~Curiosity rating of the second norming study (i.e., "Describe"; "Tell me more"), and two sentences with the lowest scores on the same dimension (i.e., "Ok"; "Thanks", see ¶2.2).

4.2. Results

We removed 2.84 % of datapoints ($n = 134$) for further analyses.

4.2.1. Plausibility judgments

We found a significant main effect of Category $\chi^2(1) = 19.20$, $p < .0001$, and a significant interaction between Category and Type of Follow-up, $\chi^2(1) = 718.88$, $p < .0001$, $b = 3.84$, 95 % CI [3.56, 4.12]. No other main effect reached significance, Category $p = .15$. Post-hoc contrasts showed that participants judged as more plausible abstract sentences when they were paired with curiosity follow-ups compared to when they were paired with end follow-ups, $z = 19.949$, $p < .0001$. Conversely, concrete sentences were judged as more plausible when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, $z = 18.268$, $p < .0001$. We also found that abstract–curiosity pairings were judged as more plausible than concrete–curiosity pairings, $z = 17.459$, $p < .0001$, and abstract–end pairings were judged as less plausible than concrete–end pairings, $z = -5.569$, $p < .0001$ (see Fig. 5).

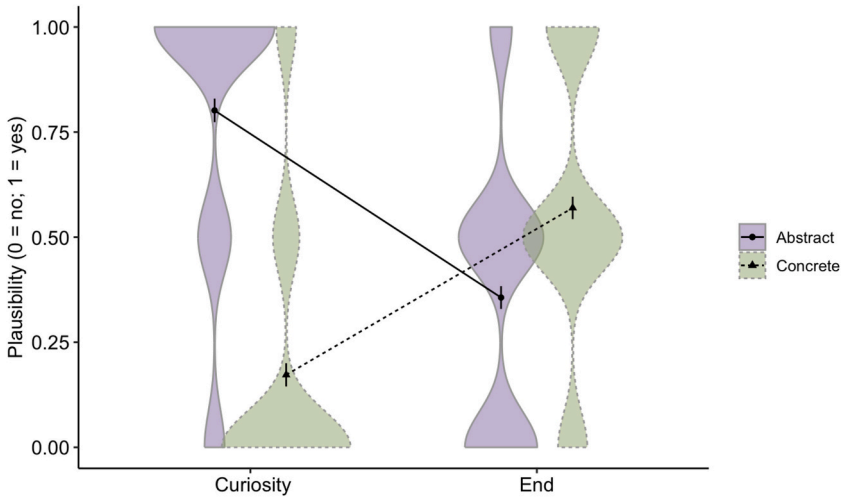

Fig. 5. Predicted probabilities of plausibility judgments for abstract and concrete sentences as a function of follow-ups.

4.2.2. Response times

Plausible. We found a significant main effect of Type of Follow-up, $F(1, 1472.40) = 5.09$, $p = .024$, and a significant interaction between Category and Type of Follow-up, $F(1, 1446.66) = 20.50$, $p <. 0001$, $b = -0.14$, 95 % CI [−0.21, −0.08]. No other main effect reached significance, Category $p = .11$. Post-hoc contrasts showed that participants were faster to judge as plausible concrete sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, $t(1122) = -4.153$, $p < .0001$, whereas there was no difference within abstract sentences across types of follow-ups, $t(2036) = 1.858$, $p =. 063$. We also found that participants responded faster to abstract–curiosity pairings than concrete–curiosity pairings, $t(161) = -3.791$, $p = .0002$, whereas there was no difference between abstract–end pairings and concrete–end pairings, $t(113) = 1.922$, $p = .057$.

Implausible. We found a significant main effect of Category, $F(1, 45.16) = 17.84$, $p < .0001$, Type of Follow-up, $F(1, 2161.16) = 41.40$, $p < .0001$, and a significant interaction between Category and Type of Follow-up, $F(1, 2180.16) = 4.54$, $p = .033$, $b = 0.06$, 95 % CI [0.004, 0.11]. Post-hoc contrasts showed that participants were faster to judge as not plausible abstract sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, $t(2070) = 5.279$, $p < .0001$. Likewise, they were faster to judge as not plausible concrete sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, $t(2332) = 3.743$, $p = .0002$. We also found that abstract–curiosity pairings were responded to slower than concrete–curiosity pairings, $t(126.1) = 4.344$, $p < .0001$. There was also a difference within end follow-ups, where abstract–end pairings were also responded to slower than concrete–end pairings, $t(74.3) = 2.504$, $p = .014$. (see Fig. 6).

4.3. Discussion

In line with our predictions, participants judged abstract–curiosity pairings as more plausible than both concrete–curiosity pairings and abstract–end pairings. This was partly confirmed by RTs, where abstract–curiosity pairings were judged as plausible faster than concrete–curiosity pairings. Response times of implausible answers instead revealed a general advantage of pairings featuring end follow-ups: both abstract and concrete sentences were judged as implausible faster when followed by end expressions. There was also a general advantage in the processing of pairings preceded by concrete over abstract sentences, regardless of the follow-up. Finally, consistent with our expectations, participants were slower to judge implausible abstract–curiosity pairings than concrete–curiosity pairings, but the same pattern emerged for abstract–end pairings and concrete–end pairings.

5. Experiment 3: abstractness and types of questions

Experiment 3 investigates whether, compared to concrete sentences, abstract sentences evoke more interest in internal processes and causal mechanisms than spatial and temporal specifications. We hypothesized that participants judge as more plausible follow-ups for abstract sentences "why" and "who" questions (i.e., causal questions) compared to "where" and "when" questions (i.e., spatio-temporal questions), and the opposite for concrete sentences. We also expected a difference in RTs for abstract and concrete sentences as a function of the type of follow-up.

5.1. Methods

5.1.1. Participants

Twenty-seven Italian speakers were recruited at the University of Bologna (22 females, 1 "other"; *M* age = 21.81; *SD* = 2.77, age range = 18–28).

5.1.2. Materials, design, and procedure

The procedure, design, and data analysis were identical to those of Experiment 1 and 2. Materials were also identical, except for the follow-up sentences for which we selected the most frequently produced questions related to spatio-temporal and causal–agent issues from Villani et al. (2022), i.e., "where?/when?" and "why/for whom?", respectively.

5.2. Results

We removed 2.75 % of datapoints ($n = 125$) for further analyses.

5.2.1. Plausibility judgments

We found a significant main effect of Category, $\chi^2(1) = 14.79$, $p < .0001$, Type of Follow-up, $\chi^2(1) = 123.05$, $p < .0001$, and a significant interaction between Category and Type of Follow-up, $\chi^2(1) = 60.78$, $p < .0001$, $b = 1.06$, 95 % CI [0.79, 1.33]. Post-hoc contrasts showed that participants judged as more plausible abstract sentences when they were paired with spatio-temporal follow-ups compared to when they were

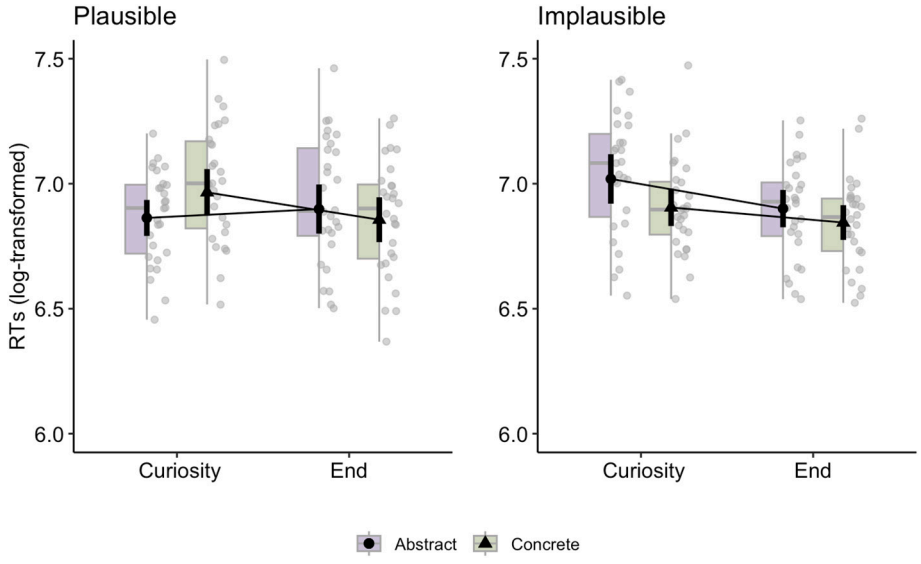

Fig. 6. Log-transformed Response Times for abstract and concrete sentences as a function of follow-ups. Raw data in the background are aggregated over participants.

paired with causal follow-ups, $z = 3.180$, $p = .001$, and the same was true for concrete sentences, $z = 13.189$, $p < .0001$. We also found that abstract–spatiotemporal pairings were judged as less plausible than concrete–spatiotemporal pairings, $z = -6.686$, $p < .0001$. There was instead no difference within causal follow-ups across types of sentences, $z = -1.088$, $p = .27$ (see Fig. 7).

5.2.2. Response times

Plausible. We found a significant main effect of Category, $F(1, 39.28) = 13.72$, $p = .006$, showing that concrete sentences were responded to faster than abstract sentences (M concrete = 956.25 *ms*; $SD = 340.13$; M abstract = 994.64 *ms*; $SD = 376.94$), and a significant main effect of Type of Follow-up, $F(1, 2579.06) = 32.55$, $p <. 0001$, showing that overall participants were faster with spatio-temporal follow-ups ($M = 949.24$ *ms*; $SD = 341.86$) compared to causal follow-ups ($M = 1004.37$ *ms*; $SD = 374.23$). The two-way interaction did not reach significance, $p = .095$.

Implausible. We found a significant main effect of Type of Follow-up, $F(1, 1675.18) = 4.27$, $p = .038$, and a significant interaction between Category and Type of Follow-up, $F(1, 1669.74) = 9.01$, $p = .002$, $b = 0.09$, 95 % CI [0.03, 0.15]. No other main effect reached significance, Category $p = .173$. Post-hoc contrasts showed that participants were faster to judge as not plausible concrete sentences when they were paired with causal follow-ups compared to when they were paired with spatio-temporal follow-ups, $t(1591) = 3.176$, $p = .001$. There was instead no difference within abstract sentences, $t(1730) = 0.738$, $p = .460$. We also found that spatio-temporal follow-ups were responded to faster when they were paired with abstract sentences compared to when they were paired with concrete sentences, $t(148.6) = -2.660$, $p = .008$. There was instead no difference within causal follow-ups across types of sentences, $t(69.5) = 0.825$, $p = .412$. (see Fig. 8).

5.3. Discussion

Overall, contrary to our predictions, both abstract–spatio-temporal pairings and concrete–spatio-temporal pairings were always judged as being more plausible than abstract–causal and concrete–causal pairings—suggesting spatio-temporal follow-ups are generally considered more plausible linguistic continuers than causal/agent follow-ups, regardless of the sentence abstractness. This was also confirmed by faster RTs for plausible answers to spatio-temporal follow-ups. However, while there was no difference in plausibility judgments within causal/agent follow-ups across abstract and concrete sentences, spatio-temporal follow-ups were instead judged as more plausible for concrete sentences than for abstract sentences. Consistently, participants were faster in judging as not plausible concrete sentences when they were followed by causal expressions (vs spatio-temporal expressions), and faster in judging abstract sentences as being not plausible when they were followed by spatio-temporal follow-ups compared with concrete sentences followed by spatio-temporal expressions.

6. Ratings

To probe more explicitly whether abstract sentences elicit more uncertainty and need of others compared to concrete sentences, at the end of each experimental session we asked participants to provide ratings on three semantic dimensions targeting these constructs. Specifically, each sentence was rated by participants on 7-point Likert scales in terms of Confidence, Experts Need, and Social Metacognition. Confidence is defined as the degree of confidence participants have in understanding the meaning of the concept (e.g., Mazzuca et al., 2022), and was operationalised as follows: "How confident are you in your understanding of the sentence?" (1 = not at all; 7 = very much); Experts Need quantifies how much participants perceive the need to rely on experts for understanding the meaning of a concept (see also Falcinelli, Fini, Mazzuca, & Borghi, 2024 where this is called Expert Social Metacognition), and was operationalised as follows: "How much would you have relied on experts' knowledge to understand the sentence?" (1 = not at all; 7 = very much); Social Metacognition is defined as the extent to which people rely on other people's input to understand or co-construct the meaning (e.g., Borghi, 2022; Villani et al., 2019), and was operationalised as follows: "How much do you think you need other people's input to understand the sentence?" (1 = not at all; 7 = very much).

We hypothesized that if abstract sentences elicit more social validation than concrete concepts, then across the three experiments participants should provide lower Confidence rating scores but higher Experts Need and Social Metacognition rating scores for abstract sentences compared to concrete sentences. Alternatively, ratings for the three dimensions should not differ across abstract and concrete sentences.

6.1. Methods

All participants from Experiments 1–3 ($N = 83$) provided their ratings for all the target sentences ($N = 42$, $n = 21$ abstract; $n = 21$ concrete) on Confidence, Experts Need, and Social Metacognition. The order of both semantic dimensions and sentences was randomised across participants within each experiment. Sentences were presented on a computer screen in a separate session implemented on Qualtrics.

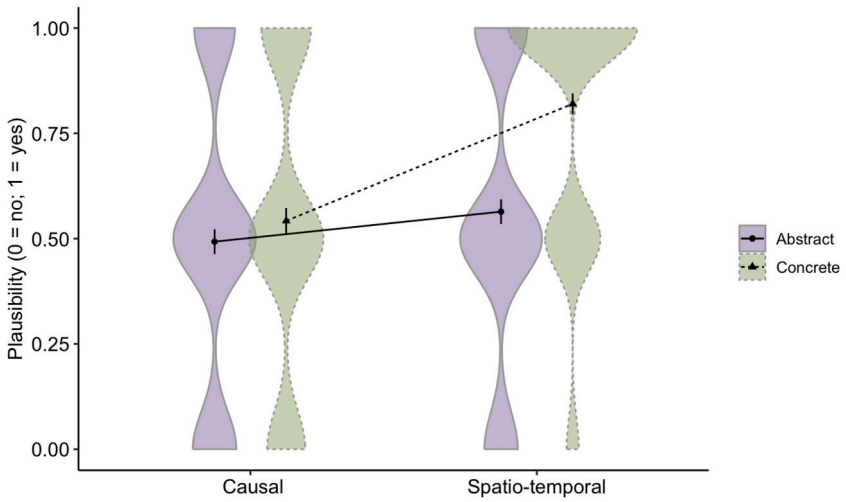

Fig. 7. Predicted probabilities of plausibility judgments for abstract and concrete sentences as a function of follow-ups.

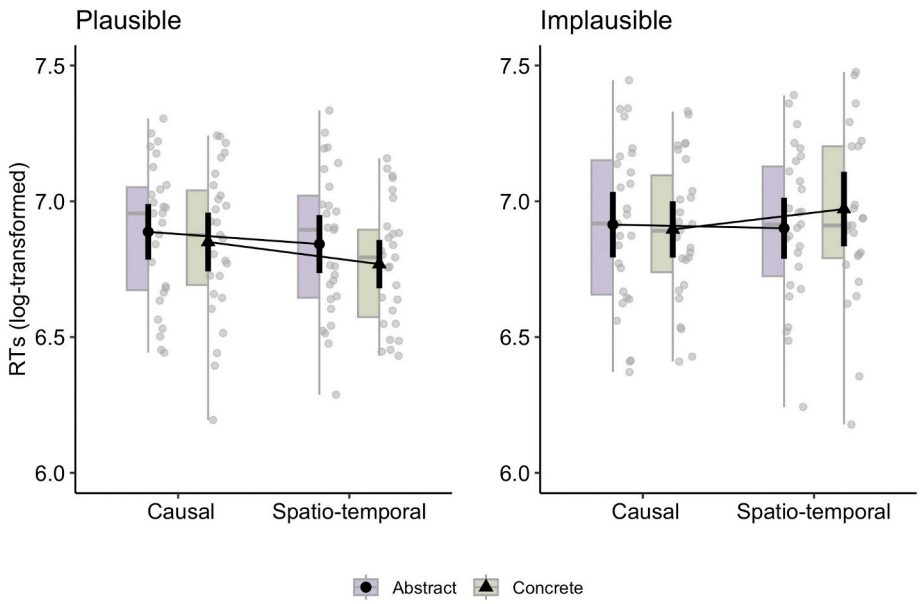

Fig. 8. Log-transformed Response Times for abstract and concrete sentences as a function of follow-ups. Raw data in the background are aggregated over participants.

6.1.1. Data analysis

Data were analysed fitting ordinal mixed effects models for each semantic dimension using the clmm() function from the "ordinal" R's package (Christensen, 2023). Each model featured rating score as dependent variable, Category of the sentence (abstract vs. concrete) as fixed factor, and participants and sentences as random intercepts.

6.2. Results

We first report descriptive statistics, and then we turn to the description and discussion of each model.

Overall, participants provided higher Confidence scores for concrete compared to abstract sentences (*M* concrete = 5.72; *SD* = 0.76; *M* abstract = 4.92; *SD* = 1.48), and higher Experts Need (*M* abstract = 3.21; *SD* = 2.38; *M* concrete = 1.77; *SD* = 1.84) and Social Metacognition scores (*M* abstract = 3.30; *SD* = 2.40; *M* concrete = 1.79; *SD* = 1.87) for abstract than concrete sentences (see Fig. 9).

Results from the models confirmed this trend. In fact, concrete sentences scored significantly higher on Confidence, $b = 1.82$, SE = 0.175, $z = 10.44$, $p < .001$, and lower on both Experts Need, $b = -2.10$, SE = 0.174, $z = -12.06$, $p < .001$ and Social Metacognition, $b = -2.13$, SE = 0.192, $z = -11.07$, $p < .001$ compared to abstract sentences. This

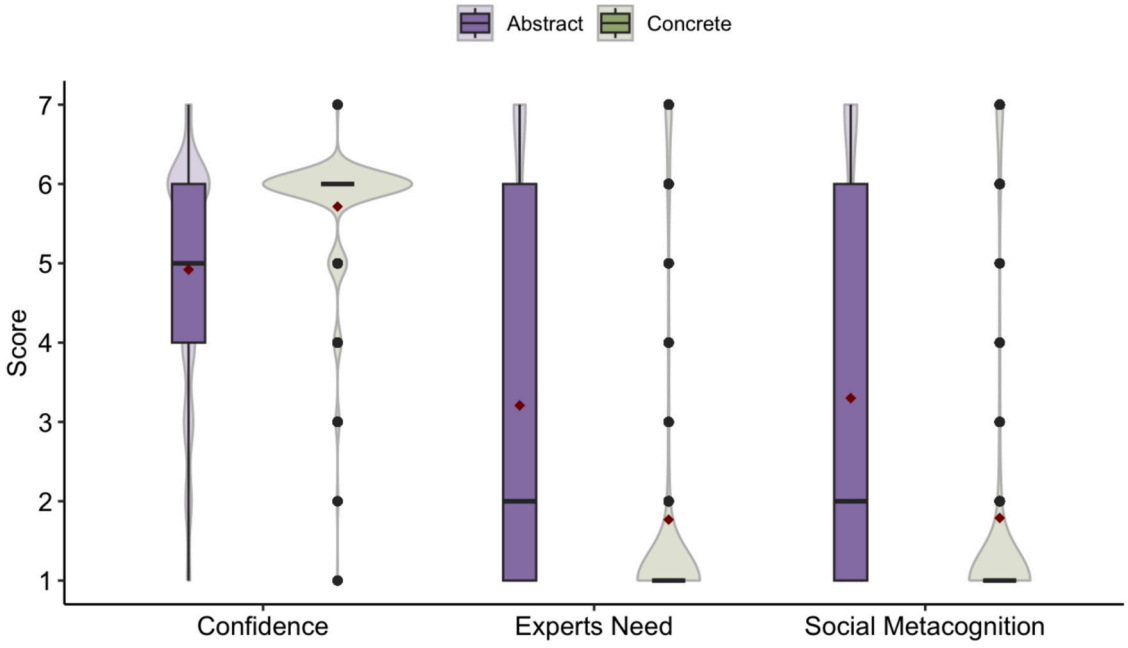

Fig. 9. Rating scores for abstract and concrete sentences on Confidence, Experts Need, and Social Metacognition. In the boxplots, red dots indicate mean values, black bold horizontal lines the median, and vertical extremes of the boxplots represent minimum and maximum values in the data. The boxes' length shows the interquartile range, with the upper side indicating the 75th percentile and the lower side indicating the 25th percentile. Violins represent the distribution of the data, with black dots showing extreme values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

suggests, in line with our hypotheses, that abstract sentences elicit more need for social validation and interaction compared to concrete sentences—for the meaning of which participants feel more confident. Importantly, these results complement and extend analogous previous findings obtained with single-word ratings (Falcinelli, Fini, Mazzuca, Alessandri, et al., 2024; Falcinelli, Fini, Mazzuca, & Borghi, 2024; Fini, Era, et al., 2023; Mazzuca et al., 2022; Villani et al., 2019).

7. General discussion

In three preregistered experiments we asked participants to evaluate the plausibility of linguistic exchanges referring to concrete and abstract concepts. Specifically, we predicted that abstract sentences would be evaluated as more plausible with follow-up expressions indicating uncertainty, curiosity, and questions related to causal/agent aspects, while concrete sentences would be judged as more plausible with follow-ups indicating certainty, willingness to end the conversation, and questions related to spatio-temoporal aspects.

In Experiment 1 we found that abstract concepts are more associated with uncertainty than concrete ones. Participants considered uncertainty expressions as more plausible when matched with abstract than with concrete sentences, and data on RTs consistently showed that participants were slower to judge as implausible abstract–uncertainty pairings than concrete–uncertainty pairings. Results from Experiment 2 show that, with pairings involving abstract sentences, people seem to think of an interaction that continues. Indeed, with abstract sentences, pairings with curiosity follow-ups are evaluated as more plausible and processed faster than with concrete sentences, while the opposite is true for pairings with end follow-ups. Experiment 3 showed that spatio-temporal follow-ups are considered as more plausible with concrete than abstract sentences and require more time to be judged as implausible when paired with concrete sentences. This indicates that people tend to locate concrete concepts spatiotemporally, consistent with the fact that their referents are spatially bounded objects or entities. Contrary to our expectations, we did not find that abstract sentences were judged as more plausible with causal follow-ups, suggesting that causality is relevant for both kinds of sentences. We speculate that the discrepancy with Villani et al.'s (2022) results on types of questions might be due to the difference between the two tasks, with the production task likely capturing more immediate associations than the plausibility task. A further alternative explanation might concern the stimuli selection. It is possible that, overall, causal/agent follow-up questions were particularly unlikely for both types of sentences, especially those relating to agency (e.g., concrete: "*I pet the dog – for whom?*"; abstract: "*I understood the reason – for whom?*", see Table 1, Supplementary Materials). In Experiment 1 we also found that participants judged certainty expressions equally plausible as follow-ups for abstract and concrete sentences, and that overall they were faster to be processed compared to uncertainty expressions. While we did not predict such a result, this finding is consistent with multiple lines of evidence. First, research in conversation analysis suggests that in short and contiguous conversational turns like the ones we devised there is a preference for 'agreement' responses (e.g., Sacks, 2020). Second, insights from studies concerned with conversation incoherencies consistently show that in "casual" conversations (e.g., conversations about cartoons or celebrities) people typically do not notice communication problems, both in online and in face-to-face interactions (Galantucci, Roberts, & Langstein, 2018; Roberts, Langstein, & Galantucci, 2016). More to the point, Galantucci, Langstein, Spivack, and Paley (2020) showed that participants are more likely to avoid initiating repairs (e.g., asking for clarifications) after hearing a non-word when the communication problem is less likely to lead to overt consequences (i.e., failing the experimental task). This further suggests the emergence of conversation devices can be modulated by contextual—or task—constraints. Finally, some accounts of language comprehension propose interlocutors create representations that are "good enough" for the task to be performed, rather than accurate and detailed (Ferreira & Patson, 2007). Typically, the task entails the production of a satisfying or appropriate follow-up, rather than a proof of accuracy of refinement of understanding. On the whole, these observations align with our results, indicating an overall preference for responses that potentially allow the flow of conversation to run smoothly compared to those signalling communication disruptions. Interestingly, this tendency seems to be unaffected by sentence abstractness.

Finally, data from the plausibility judgment task are complemented by a more explicit measure i.e., ratings. We found that when explicitly asked to estimate their confidence about their understanding of sentences, participants feel less confident with abstract sentences than with concrete sentences. Accordingly, they also think they would rely more on other people and on experts to grasp the meaning of abstract sentences compared to concrete sentences.

From a methodological perspective, our study has some characteristics that are worth mentioning—as they align well with the recent calls for a paradigmatic shift advanced by scholars interested in abstractness and social interaction (Banks et al., 2023; Borghi et al., 2023). Rather than using concrete and abstract words embedded in a single sentence, we employed a simulated verbal exchange, albeit only composed of two turns. In addition, both target sentences and follow-up expressions were derived from a prior linguistic production study (Villani et al., 2022). So, while the linguistic exchanges were "simulated" and not drawn from interactive corpora, we examined concepts embedded in expressions taken from natural linguistic exchanges—thereby enhancing the ecological validity of our findings. The natural continuation of this pursuit would necessarily aim to strengthen and replicate these findings using multimodal interactive corpora of social interactions.

Taken together, our findings are in line with proposals linking abstract concepts processing with inner monitoring mechanisms and social interactions (Andrade-Lotero, Ortiz-Duque, Velasco-García, & Goldstone, 2023; Binder et al., 2005; Fernyhough & Borghi, 2023; Olsen & Tylén, 2023; Rączaszek-Leonardi & Zubek, 2023; Shea, 2018). Indeed, findings from both Experiment 1 and the sentences rating task are in line with single words rating results that show people typically feel less confident, and more uncertain about the meaning of abstract words (Fini, Era, et al., 2023; Mazzuca et al., 2022). This possibly leads speakers to cling to social interactions of various kinds, spanning from the reliance on others to the deference to experts in the case of more complex abstract meanings (Falcinelli, Fini, Mazzuca, & Borghi, 2024; Mazzuca et al., 2022; Villani et al., 2019). Results from Experiment 2 shed a different light on the relationship between abstractness and sociality. In fact, our findings indicate that, in conversations involving abstract concepts, people are more willing to engage in longer and more in-depth interactions. This is consistent with proposals suggesting that abstract concepts might be more variable across individuals and cultures (Borghi & Mazzuca, 2023; Wang & Bi, 2021), and their meaning is more likely to be the remit of negotiations and be co-constructed in online interactions by speakers (Borghi, 2022; Fini, Era, et al., 2023; Mazzuca & Santarelli, 2023). So, people might need longer and more thorough interactions to reach a conversational common ground when dealing with abstract topics (see also Gandolfi et al., 2023). It is also worth noting that, while the literature on context availability consistently showed that abstract concepts typically require more contextual information than concrete concepts (Schwanenflugel & Stowe, 1989; Schwanenflugel et al., 1992)—hence possibly eliciting more questions—here we focus on a different aspect, i.e., the type of questions and follow-up expressions associated with more abstract sentences. We found that, all things being equal, abstract sentences are linked with uncertainty and spark people's interest more than concrete sentences.

Our results are also relevant for research on interactive repairs, as they extend its scope to the relation between conversational devices and abstractness. In fact, even though our operationalisations are not completely overlapping with those proposed in the literature (e.g., Dingemanse & Enfield, 2024), they might still offer interesting insights

for this research domain. Roughly speaking, our data suggest more abstract sentences are associated with open requests (e.g., "What do you mean?"; "Tell me more"), whereas our experimental design did not allow us disentangle the impact of abstractness on restricted requests—as both abstract and concrete sentences were paired with these conversational devices, which only differed for the type of semantic information they conveyed (causal/agent vs. spatio-temporal). As for backchannels, the implementation of our study prevents us from drawing any meaningful conclusions. Here, expressions often identified as backchannels like "ok" were used as a single conversational turn in a simulated two-turn conversation rather than being embedded in a real, interactive linguistic exchange. This aspect intrinsically obscured the feedback and monitoring function that backchannels are thought to have in interactions (Dideriksen, Fusaroli, Tylén, Dingemanse, & Christiansen, 2019). Importantly, these follow-ups were previously selected from a norming study as they clearly manifested the willingness to end a conversation. While this was beyond the aims of our study, we believe future research may deepen the intersection between abstractness and conversational devices with a more balanced and tailored design.

On top of all that, these findings conceptually replicate and extend with a novel task the results of a production task (Villani et al., 2022). In line with the increasingly growing awareness of reproducibility issues in psychological science (Open Science Collaboration, 2015; Youyou, Yang, & Uzzi, 2023), we sought to increase the confirmatory power of our replication by introducing differences in materials that would have made it possible to generalise our underlying hypothesis to a different domain (i.e., conversations). In other words, we aimed at implementing a conceptual replication that, in Schmidt's (2009) words "[…] reaches further than a direct one. The successful replication of a hypothesis validates this hypothesis but it also corroborates the theory behind it. […] While a direct replication is able to produce *facts* a conceptual replication may produce *understanding*. […]. The function of a conceptual replication is not only to confirm facts but also to assist in developing models and theories of the world" (p. 18).

Nevertheless, this study is not without limitations. For instance, we focused on differences between concrete and abstract concepts, yet many scholars emphasise the importance of examining various types of concrete and abstract concepts (Conca et al., 2021; Desai, Reilly, & van Dam, 2018; Persichetti, Shao, Denning, Gotts, & Martin, 2024). Although we controlled this aspect in the stimuli selection using different kinds of abstract and concrete concepts, we did not take this into account for the present analyses. Thus, future studies might be planned to specifically investigate fine-grained differences in plausibility judgments using different types of concrete and abstract concepts. Another potential limitation is that we did not consider the hierarchical level of concepts (i.e., their specificity, Bolognesi & Caselli, 2023; Villani et al., 2024), although an ongoing study is currently addressing this aspect.

Finally, while we leveraged on two short simulated conversation turns, future studies should expand the investigation of abstractness and conversation dynamics using real and longer conversations. In keeping with research underlining the entwinement between abstractness and social interaction and negotiation (Borghi, 2022, 2023; Falcinelli, Fini, Mazzuca, & Borghi, 2024; Mazzuca & Santarelli, 2023), one might hypothesise that abstract words in conversations are more often followed by open requests compared to concrete words, which on the other hand would be more often followed by restricted requests. This would reflect the relative difficulty of establishing a common ground with abstract concepts stemming from their indeterminate and variable character. However, based on the literature addressing repair avoidance (e.g., Galantucci et al., 2020) one could also expect that in face-to-face interactions people might avoid asking for clarifications to get the conversation going. This might be especially true for abstract concepts embedded in conversations that are not goal-oriented—i.e., where agreeing on what nuance of "freedom" the speakers are talking about is not essential. Last, naturally occurring conversations would be the perfect setting to test predictions on the impact of abstractness on linguistic alignment (Pickering & Garrod, 2004). We might expect syntactic and semantic structures to align more across interlocutors in more abstract conversations compared to more concrete ones (Gandolfi et al., 2023), as speakers strain to reach a common ground. Conversely, some studies suggest linguistic diversity fosters abstract problem solving (Tylén, Fusaroli, Østergaard, Smith, & Arnoldi, 2023). So, speakers might employ different syntactic structures or lexical items in disagreement or miscommunication situations, to aid the intelligibility of the conversation, thereby reducing linguistic alignment.

To sum up, the implications of our results are multilayered, spanning from research on concepts in general—and abstractness in particular—to studies on dialogue (Pickering & Garrod, 2021) and conversational devices (Dingemanse & Enfield, 2024). Indeed, overall, we found that when conversational topics are more abstract people judge them as requiring more effort to align with the interlocutor compared with concrete conversational topics. Operationally, this effort in reaching alignment is revealed in the higher uncertainty of follow-ups associated with abstract sentences and increased willingness to continue the conversation (curiosity) with pairings including abstract sentences. Hence, our results contribute to the existing literature on dialogue and conversational devices, showing that the concreteness ~ abstractness of the topic might differentially modulate the conversational dynamics (Gandolfi et al., 2023).

CRediT authorship contribution statement

Claudia Mazzuca: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Caterina Villani:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization. **Tommaso Lamarra:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Marianna Marcella Bolognesi:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Anna M. Borghi:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Acknowledgments

The authors would like to thank all the Body, Action, and Language Lab (BalLab, https://sites.google.com/view/annaborghilab) for insightful remarks on early versions of this study. In particular, we thank Chiara Fini, Luca Tummolini, Angelo Mattia Gervasi, Chiara De Livio, Ilenia Falcinelli, and Valentina Rossi for their stimulating thoughts. The authors would also like to thank all members of ABSTRACTION research group (GRANT AGREEMENT: ERC-2021-STG-101039777) for comments and discussions. They are, in alphabetical order, Adele Loia, Andrea Amelio Ravelli, Claudia Collacciani, and Giulia Rambelli. Data presented in the present contribution have been collected at the Experimental Lab of the Department of Modern Languages, Literatures and Cultures of the University of Bologna, Italy (https://site.unibo.it/laboratorio-sperimentale/en). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the Experimental Lab.

Funding

Claudia Mazzuca was funded by Research Projects of National Relevance PRIN 2022 – PROJECT ASSO (Abstract conceptS and Social InteractiOn). Caterina Villani, Tommaso Lamarra, and Marianna Bolognesi were funded by ABSTRACTION European Union (GRANT AGREEMENT: ERC-2021-STG-101039777). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council

Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Anna Borghi was funded by PNRR funds, PE8 - AGE-IT - Spoke 4: Healthy aging. Project line led by AMB: "Older adults' relationship with the natural environment to enhance their cognitive performances and promote mental health".

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare no generative AI tool was used for the preparation of the present manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2025.106084.

Data availability

All the links to the preregistration and to open data, code, and materials are referenced in the manuscript.

References

Andrade-Lotero, E. J., Ortiz-Duque, J. M., Velasco-García, J. A., & Goldstone, R. L. (2023). The division of linguistic labour for offloading conceptual understanding. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210360.

Banks, B., Borghi, A. M., Fargier, R., Fini, C., Jonauskaite, D., Mazzuca, C., … Woodin, G. (2023). Consensus paper: Current perspectives on abstract concepts and future research directions. *Journal of Cognition, 6*(1).

Banks, B., & Connell, L. (2023). Multi-dimensional sensorimotor grounding of concrete and abstract categories. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210366.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher, & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129–163). New York: Cambridge University Press.

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

Bellagamba, F., Borghi, A. M., Mazzuca, C., Pecora, G., Ferrara, F., & Fogel, A. (2022). Abstractness emerges progressively over the second year of life. *Scientific Reports, 12*(1), Article 20940.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience, 17*(6), 905–917.

Bolognesi, M., & Steen, G. (2018). Editors' introduction: Abstract concepts: Structure, processing, and modeling. *Topics in Cognitive Science, 10*(3), 490–500.

Bolognesi, M. M., & Caselli, T. (2023). Specificity ratings for Italian data. *Behavior Research Methods, 55*(7), 3531–3548.

Borghi, A. M. (2022). Concepts for which we need others more: The case of abstract concepts. *Current Directions in Psychological Science, 31*(3), 238–246.

Borghi, A. M. (2023). *The freedom of words: Abstractness and the power of language*. Cambridge University Press.

Borghi, A. M., Barca, L., Binkofski, F., & Tummolini, L. (2018). Abstract concepts, language and sociality: From acquisition to inner speech. *Philosophical Transactions of the Royal Society B, 373*(1752), Article 20170134.

Borghi, A. M., & Mazzuca, C. (2023). Grounded cognition, linguistic relativity, and abstract concepts. *Topics in Cognitive Science, 15*(4), 662–667.

Borghi, A. M., Osińska, A., Roepstorff, A., & Raczaszek-Leonardi, J. (2023). Concepts in interaction: Social engagement and inner experiences. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210351.

Christensen, R. (2023). Ordinal—Regression models for ordinal data. R package version 2023.12-4. https://CRAN.R-project.org/package=ordinal.

Clark, H. H. (1996). Communities, commonalities, and communication. *Rethinking Linguistic Relativity, 17*, 324–355.

Conca, F., Borsa, V. M., Cappa, S. F., & Catricalà, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neuroscience & Biobehavioral Reviews, 127*, 474–491.

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: The forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 373*(1752), Article 20170143.

Davis, C. P., Altmann, G. T., & Yee, E. (2020). Situational systematicity: A role for schema in understanding the differences between abstract and concrete concepts. *Cognitive Neuropsychology, 37*(1–2), 142–153.

Della Rosa, Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract—concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods, 42*, 1042–1048.

Desai, R. H., Reilly, M., & van Dam, W. (2018). The multifaceted abstract brain. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 373*(1752), Article 20170122.

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General, 152*(3), 864.

Dideriksen, C. R., Fusaroli, R., Tylén, K., Dingemanse, M., & Christiansen, M. H. (2019). Contextualizing conversational strategies: Backchannel, repair and linguistics alignment in spontaneous and task-oriented conversations. In *The 41st annual meeting of the cognitive science society* (pp. 261–267). Cognitive Science Society.

Dingemanse, M. (2017). Brain-to-brain interfaces and the role of language in distributing agency. In N. J. Enfield, & P. Kockelman (Eds.), *Distributed agency* (pp. 59–66). Oxford University Press.

Dingemanse, M., & Enfield, N. J. (2015). Other-initiated repair across languages: Towards a typology of conversational structures. *Open Linguistics, 1*(1).

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., … Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS ONE, 10*(9), Article e0136100.

Dingemanse, M., & Enfield, N. J. (2024). Interactive repair and the foundations of language. *Trends in Cognitive Sciences, 28*(1), 30–42.

Diveica, V., Pexman, P. M., & Binney, R. J. (2023). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods, 55*(2), 461–473.

Dove, G. (2022). *Abstract concepts and the embodied mind: Rethinking grounded cognition*. Oxford University Press.

Enfield, N. J. (2017). *How we talk: The inner workings of conversation*. Basic Books.

Falcinelli, I., Fini, C., Mazzuca, C., Alessandri, G., Alivernini, F., Baiocco, R., … Borghi, A. M. (2024). What does "internet" mean to us as we age? *A multi-task investigation on the conceptualization of the technological domain across generations. Computers in Human Behavior Reports, 16*(10–51), 100531.

Falcinelli, I., Fini, C., Mazzuca, C., & Borghi, A. (2024). The TECo database: Insights on the semantic Organization of the Ecological Domain. In *, vol. 46. Proceedings of the annual meeting of the Cognitive Science Society*.

Fay, N., Walker, B., Swoboda, N., Umata, I., Fukaya, T., Katagiri, Y., & Garrod, S. (2018). Universal principles of human communication: Preliminary evidence from a cross-cultural communication game. *Cognitive Science, 42*(7), 2397–2413.

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. (2024). Language is primarily a tool for communication rather than thought. *Nature, 630*(8017), 575–586.

Fernyhough, C., & Borghi, A. M. (2023). Inner speech as language process and cognitive tool. *Trends in Cognitive Sciences, 27*(12), 1180–1193.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Lang & Ling Compass, 1*(1–2), 71–83.

Fini, C., Era, V., Cuomo, G., Falcinelli, I., Gervasi, A., Candidi, M., … Borghi, A. M. (2023). *Digital connection, real bonding: brief online chats boost interpersonal closeness regardless of the conversational topic*. https://doi.org/10.31219/osf.io/b8c2d

Fini, C., Era, V., Da Rold, F., Candidi, M., & Borghi, A. M. (2021). Abstract concepts in interaction: The need of others when guessing abstract concepts, smooths dyadic motor interactions. *Royal Society Open Science, 8*(7), Article 201205. https://doi.org/10.1098/rsos.201205

Fini, C., Falcinelli, I., Cuomo, G., Era, V., Candidi, M., Tummolini, L., … Borghi, A. M. (2023). Breaking the ice in a conversation: Abstract words prompt dialogs more easily than concrete ones. *Language and Cognition*, 1–22.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Thousand Oaks, CA: Sage.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology, 32*, 147–157.

Galantucci, B., Langstein, B., Spivack, E., & Paley, N. (2020). Repair avoidance: When faithful informational exchanges don't matter that much. *Cognitive Science, 44*(10), Article e12882.

Galantucci, B., & Roberts, G. (2014). Do we notice when communication goes awry? An investigation of people's sensitivity to coherence in spontaneous conversation. *PLoS ONE, 9*(7), Article e103182.

Galantucci, B., Roberts, G., & Langstein, B. (2018). Content deafness: When coherent talk just doesn't matter. *Language & Communication, 61*, 29–34. https://doi.org/10.1016/j.langcom.2018.01.001

Gandolfi, G., Pickering, M. J., & Garrod, S. (2023). Mechanisms of alignment: Shared control, social cognition and metacognition. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210362.

Gardner, R. (2001). When listeners talk. *Benjamins*. https://doi.org/10.1075/pbns.92

Gilhooly, K. J., & Logie, R. H. (1981). Word age-of-acquisition, reading latencies and auditory recognition. *Current Psychology, 1*, 251–262.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language learning and development, 1*(1), 23–64.

Healey, P. G., De Ruiter, J. P., & Mills, G. J. (2018). Editors' introduction: Miscommunication. *Topics in Cognitive Science, 10*(2), 264–278.

Henningsen-Schomers, M. R., & Pulvermüller, F. (2022). Modelling concrete and abstract concepts using brain-constrained deep neural networks. *Psychological Research, 86*(8), 2533–2559.

Hoffman, P. (2016). The meaning of "life" and other abstract words: Insights from neuropsychology. *Journal of Neuropsychology, 10*(2), 317–343. https://doi.org/10.1111/jnp.12065

Kiefer, M., & Harpaintner, M. (2020). Varieties of abstract concepts and their grounding in perception or action. *Open Psychology, 2*(1), 119–137.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(1), 92–107. https://doi.org/10.1037/0278-7393.12.1.92

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Putting everything in context. *Cognitive Science, 39*(8), 1987–1995.
Lenth, R. V. (2021). emmeans: Estimated marginal means, aka least-squares means. R package version 1.7.1-1. https://CRAN.R-project.org/package=emmeans.
Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139
Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society, B: Biological Sciences, 373*(1752), Article 20170137.
Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster sensorimotor norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods, 52*, 1271–1291.
Manrique, E., & Enfield, N. J. (2015). Suspending the next turn as a form of repair initiation: evidence from Argentine Sign Language. *Frontiers in Psychology, 6*(1326), 1–21.
Mazzuca, C., Borghi, A. M., van Putten, S., Lugli, L., Nicoletti, R., & Majid, A. (2024). Gender is conceptualized in different ways across cultures. *Language and Cognition, 16*(2), 353–379.
Mazzuca, C., Falcinelli, I., Michalland, A. H., Tummolini, L., & Borghi, A. M. (2022). Bodily, emotional, and public sphere at the time of COVID-19. An investigation on concrete and abstract concepts. *Psychological Research, 86*(7), 2266–2277.
Mazzuca, C., Fini, C., Michalland, A. H., Falcinelli, I., Da Rold, F., Tummolini, L., & Borghi, A. M. (2021). From affordances to abstract words: The flexibility of sensorimotor grounding. *Brain Sciences, 11*(10), 1304.
Mazzuca, C., Majid, A., Lugli, L., Nicoletti, R., & Borghi, A. M. (2020). Gender is a multifaceted concept: Evidence that specific life experiences differentially shape the concept of gender. *Language and Cognition, 12*(4), 649–678.
Mazzuca, C., & Santarelli, M. (2023). Making it abstract, making it contestable: Politicization at the intersection of political and cognitive science. *Review of Philosophy and Psychology*, 1–22.
McRae, K., Nedjadrasul, D., Pau, R., Lo, B. P. H., & King, L. (2018). Abstract concepts and pictures of real-world situations activate one another. *Topics in Cognitive Science, 10*(3), 518–532.
Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112*, Article 104092.
Olsen, K., & Tylén, K. (2023). On the social nature of abstraction: Cognitive implications of interaction and diversity. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210361.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), Article aac4716.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie, 45*(3), 255.
Persichetti, A. S., Shao, J., Denning, J. M., Gotts, S. J., & Martin, A. (2024). Taxonomic structure in a set of abstract concepts. *Frontiers in Psychology, 14*, Article 1278744. https://doi.org/10.3389/fpsyg.2023.1278744
Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(02), 1162–1168.
Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press.
Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental Science, 21*(2), Article e12549.
R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
Rączaszek-Leonardi, J., & Zubek, J. (2023). Is love an abstract concept? A view of concepts from an interaction-based perspective. *Philosophical Transactions of the Royal Society B, 378*(1870), Article 20210356.
Reilly, J., Shain, C., Borghesani, V., Kuhnke, P., Vigliocco, G., Peelle, J. E., … Vinson, D. (2024). What we mean when we say semantic: Toward a multidisciplinary semantic glossary. *Psychonomic Bulletin & Review*, 1–38.
Roberts, G., Langstein, B., & Galantucci, B. (2016). (in)sensitivity to incoherence in human communication. *Language & Communication, 47*, 15–22. https://doi.org/10.1016/j.langcom.2015.11.001
Sacks, H. (2020). On the preferences for agreement and contiguity in sequences in conversation. In *Language in use* (pp. 8–22). Routledge.
Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh'and other things that come between sentences. *Analyzing Discourse: Text and Talk, 71*, Article 93.
Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., … Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141–1152.
Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*(2), 90–100.
Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Erlbaum.
Schwanenflugel, P. J., Akin, C., & Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition, 20*(1), 96–104.
Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27*(5), 499–520. https://doi.org/10.1016/0749-596X(88)90022-8
Schwanenflugel, P. J., & Stowe, R. W. (1989). *Context availability and the processing of abstract and concrete words in sentences* (pp. 114–126). Reading Research Quarterly.
Shea, N. (2018). Metacognition and abstract concepts. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 373*(1752), Article 20170133.
Skedsmo, K. (2020). Multiple other-initiations of repair in Norwegian Sign Language. *Open Linguistics, 6*(1), 532–566.
Templeton, E., Chang, L., & Wheatley, T. (2024). Conversational launch pads: Strangers start their conversations with topics that lead to many other topics. In *, vol. 46. Proceedings of the Annual Meeting of the Cognitive Science Society*.
Tillotson, S. M., Siakaluk, P. D., & Pexman, P. M. (2008). Body—Object interaction ratings for 1,618 monosyllabic nouns. *Behavior Research Methods, 40*(4), 1075–1078.
Tylén, K., Fusaroli, R., Østergaard, S. M., Smith, P., & Arnoldi, J. (2023). The social route to abstraction: Interaction and diversity enhance performance and transfer in a rule-based categorization task. *Cognitive Science, 47*(9), Article e13338.
Villani, C., Loia, A., & Bolognesi, M. M. (2024). The semantic content of concrete, abstract, specific, and generic concepts. *Language and Cognition, 2024*, 1–28. https://doi.org/10.1017/langcog.2023.64
Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition, 11*(3), 403–430.
Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., & Borghi, A. M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *Journal of Memory and Language, 116*, Article 104173.
Villani, C., Orsoni, M., Lugli, L., Benassi, M., & Borghi, A. M. (2022). Abstract and concrete concepts in conversation. *Scientific Reports, 12*(1), Article 17572.
Wang, X., & Bi, Y. (2021). Idiosyncratic tower of babel: Individual differences in word-meaning representation increase as word abstractness increases. *Psychological Science, 32*(10), 1617–1635.
Wauters, L. N., Tellings, A. E., Van Bon, & Van Haaften, A. W. (2003). Mode of acquisition of word meanings: The viability of a theoretical construct. *Applied Psycholinguistics, 24*(3), 385–406.
Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.
Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review, 23*, 1015–1027.
Yngve, V. (1970). On getting a word in edgewise. In *Chicago Linguistics Society, 6th meeting* (pp. 567–577).
Youyou, W., Yang, Y., & Uzzi, B. (2023). A discipline-wide investigation of the replicability of Psychology papers over the past two decades. *Proceedings of the National Academy of Sciences, 120*(6), Article e2208863120.
Zdrazilova, L., Sidhu, D. M., & Pexman, P. M. (2018). Communicating abstract meaning: Concepts revealed in words and gestures. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 373*(1752), Article 20170138.