**Supplementary discussion.**

**Object-interaction videos as control stimulus.**

In the protocol used in this study, we compared the responses toward infant-interaction and object-interaction videos. While the control videos were designed as to closely mimic the hand movements in the infant videos to minimize the motor confounds, and were framed in such a way to minimize other visual-related confounds, there is a possibility that due to the non-social nature of the control videos, a social, non-parenting component remains to be controlled for in our infant stimuli. This means, if we assume that the effect elicited by our infant stimuli contains a visual (scene layout), a motor (hand movement), a social (interaction between two humans), and a parenting (interaction with an infant) components, our object-interaction videos would only control for the former two components. Whether the social and parenting components are distinct or separable has already been argued on the Discussion. However, from a methodological point of view, the choice of these object-interaction videos as the control condition warrant additional discussion.

There were a number of considerations that forced us to use this kind of video as the control condition. First of all, the current study was based on the protocol by Abraham et al. (2014) in order to confirm the parental brain network in human males. In their work, Abraham and colleagues contrasted the participants’ interaction with own infant vs unknown parent interaction with unknown infant, and with parent self-interaction with object. However, it is simply impossible to use the stimuli showing true parent interaction with their own infant since all of the subjects participating in the current study started as either nonfathers or expectant fathers, i.e. they have no infant of their own at the time of the experiment.

Thus, in a sample of non-parents, one could only use stimuli of unknown infant interaction and of object interaction. Because previous studies have suggested that fathers show different neural responses to their own infants than to unknown ones (Kuo et al., 2012; Mascaro, Hackett & Rilling, 2013), it is possible that the stimuli from an unknown infant could have a reduced parenting component, especially in comparison with the social component. Thus, to better address this, we could have expanded on Abraham et al. (2014)’s protocol, and include other kind of control videos to account for this factor. For example, we could have controlled the social component using adult-adult interactions. However, this would have introduced new confounding effects: differences in the *kind* of social interaction between the two stimulus types. While both types of interaction (with infant and with peer adults) are social, the nature of this interaction is fundamentally different. Our infant stimuli were themed on typical infant caregiving actions (S1: display of affection, S2: changing diaper). Situations between two adult strangers (to balance the fact that the infant in the picture is also unknown to the participants) matching this caregiving context may be rare, and much more difficult for the subjects to imagine. Another problem with this approach would be the perspective used in our stimuli (first-person view), designed as such to facilitate the emulation of the caregiving actions as if given to his own child by the subject. It would be difficult to create a social interaction stimulus between two adults in first-person view to which the subject could feel in the role of the actor. Finally, it is important to consider the emotional valence of the stimuli, and how different in may be between subjects. One of our stimuli is the act of changing diaper, a seemingly simple task that may have different emotional valence depending on the subject’s own take on paternity. Participants with little interest in infants may find the video less appealing that those who are eager to take care of their own child. Reliably replicating this valence in a control adult-adult interaction video may not be achievable.

In a similar fashion, other kind of control videos (doll-interaction, pet-interaction) would have introduced their own kind of confounding effects. Thus, the selection of unknown infant-interaction vs object-interaction was the most appropriate for this research. Future research could address these issues. A first big step would be to properly study the response of non-fathers to a variety of infant stimuli, not only infant cry and infant face, data that is severely lacking in the literature (Provenzi et al. [2021]). Likewise, studies contrasting adult-adult and adult-infant interactions, in both fathers and nonfathers, would also elucidate some light into these issues. Finally, advances in technology may someday allow the creation of virtual avatars for unborn infants, modelled after the image of the parents, for more carefully controlled experimental conditions.