UiT The Arctic University of Norway

Children's sensitivity to eyebrow flash as an ostensive stimulus: a pilot study

Emilie Suhr

Candidata psychologiae

Department of psychology, university of Tromsø

Main thesis in psychology, PSY 2901

13/12/2021, Tromsø

Preface

The theme of the final thesis and how to approach it have been a recurrent topic of discussion for me throughout the whole duration of my academic career. After writing and completing smaller assignments during my time at UiT, I realized that the greater topic rather than the specific hypothesis of a paper is the most important factor for me. When I started to think about what I wanted to do my thesis on I only knew that I wanted to do it on developmental psychology. When my mentor Mikołaj Hernik presented his plans to study infants' sensitivity to ostensive signals with contagious enthusiasm, we decided that I could perform this pilot study for him. Mikołaj has throughout the entirety of the project been generous with his time, and has among other things helped with the registration, composition, focus, structure and statistical analysis of the data. Because Mikołaj and I do not share a first language this paper had to be written in English. I am therefore grateful for the help I have received from Aina Hindenes for proofreading and especially my partner Johan Jørgen Losvar for going over the paper multiple times to catch spelling mistakes and suggesting rephrasing. A special thanks is owed to Maria Oppigård Nilssen and Rikke Nordgård for lending me their professional gaze before the deadline. Their help has made me confident in the choices I have made in this paper.

In the first phases of this project a decision was made to do the data collection online due to the Covid 19 pandemic. This proved to be the optimal solution, seeing as Tromsø introduced local restrictions twice during the time I worked on this paper. And it allowed me to recruit a total of 18 infants in the timeframe of 1.5 months. Despite the broad age range in the target demographic and the fact that participation in the study could be completed online, recruiting enough participants was more challenging than I initially imagined. Most of the participants were therefore recruited through family and from friends of friends. I am grateful

for all the support I have received from family and friends who put me in contact with a substantial amount of the children who participated in this study.

Writing my final thesis did not pan out as I had first imagined, I was not planning to write in English and I had plans of writing a review, and I certainly had no plans of working alone. Despite this divergence from my previously intended path, I am happy that the circumstances allowed me to pursue the current topic, because working on this thesis have been both interesting and challenging beyond what I initially imagined. I have had the opportunity to have responsibility for making the stimulus, registration of the study, the advertisement to and recruitment of participants, administration of the trials, coding and analyzing the data, as well as synthesizing it all into this final thesis. In other words, it has been a rewarding process. Therefore, even if I had the chance to choose again, I would not have chosen another assignment form that would have been more familiar to me.

Abstract

Ostensive signals are an integral part of normal everyday communication, a body of evidence indicates that some of them are innately recognizable to infants. The eyebrow flash has been suggested to be an ostensive signal by authors of human ethology studies. This have been shown little attention in the psychological field, this pilot study therefore has as goal of setting the basis for studying young children's sensitivity to the eyebrow flash in a systematic way. And aims to identify if children between the ages 6 to 18 months old prefer to look at an intended ostensive stimulus over a matched distractor stimulus in a preferential looking paradigm. The results revealed a non-significant preference towards the intended ostensive stimulus. One of five exploratory analysis showed a tendency towards a significant preference for the intended ostensive stimulus in the last test trial when compared to chance, but when adjusting with the Bonferroni correction this finding were no longer significant. Six potential reasons for the non-significant results are discussed in the final section of this paper as well as suggestions on how to possibly address those problems in later studies.

Keywords: eyebrow flash, ostensive signals, child psychology

Human communication hinges on language, but language is unnecessary for identifying that someone has an intent to communicate with you. This intent can be inferred by the ostensive signals in the social context. Ostensive signals have been defined by Sperber and Wilson (1995) as signals sent by an addresser to an addressee with the purpose of manifesting communicative intent, followed by elaborations on what the addresser intends to communicate. Ostensive signals are thus signals that lets you know that another person wants to communicate. Extensive research has documented that newborn infants naturally orient towards the ostensive signals eye contact and child directed speech (motherese) (Csibra, 2010). By 13 months of age children respond to their name in the same way (Newman, 2005). In human ethology literature the *eyebrow flash* has been discussed as an ostensive signal for years (Eibil-Elbesfeldt, 2007), but in the field of psychology, studies on the eyebrow flash are scarce. The eyebrow flash can be described as quickly raising the eyebrows in greeting or when talking to another person. This pilot study seeks to start bridging the knowledge gap within our field by using a preferential looking paradigm to identify if there is a preference for face like structures that raises their eyebrows. A positive finding would be consistent with the overall idea that eyebrow flashes is an ostensive signal that is recognizable to young children.

Ostensive signals are used in everyday discourse between people, also between those who can communicate through a common verbal or sign language. As an example, imagine yourself walking down the street and an unfamiliar voice says, "you dropped something". Even if you were to look up and find yourself in a crowd you would know that you were the intended receiver of this message if you saw a woman making eye contact with you. The ostensive context of the situation makes it possible for you to infer that the woman is talking to you and that you need to check your belongings. If language is not a feasible way of communicating with you, the ostensive signals in the context could still make it apparent that the woman is talking to you even if she were to address you in a foreign language. The usefulness of ostensive signals thus becomes apparent when communicating with people who speak a different language and when communicating with nonverbal people like preverbal children. The basic functions of ostensive signals are thus that they attract and holds the addressee’s attention and let them know that they are the intended receiver of a message (Csibra, 2010), independent of their linguistic skills.

The evolution of ostensive signals and what purpose they serve

There are several ideas in the literature concerning the evolutionary history of ostensive signals and their primary functions. Making spontaneously emotionally charged facial expressions in response to different stimuli, as well as at other humans might have been selected for in evolution because of its social benefit. Widening the eyes when scared or smiling when happy are two examples, the first can communicate that there is a need for caution and the latter indicates absence of danger. Because of this social benefit, early humans might have started to use those signals intentionally as a form of nonverbal language (Frith, 2009). They were potentially used in social situations without the person being in the emotional state they were conveying (Egyed et al, 2013). In so making use of the facial expressions that previously were used to reflect a person’s genuine inner emotions. In that way humans might have managed to communicated intent to each other before verbal language became the pillar of human communication, and this might very well also have been the origins of the intentional use of ostensive signals. If this is the case the well documented preferential response children have towards ostensive signals at an early age could be a product of nature rather than nurture.

Csibra (2010) argues that ostensive signals represent a sort of code-system that requires little to no processing by the receiver to be interpreted and create readiness for interpretation of the ensuing information. Language is therefore not necessary to decode and use ostensive signals so preverbal infants can use them readily. Not only do young infants naturally orient towards sources of ostensive signals like eye contact and motherese but they also seem to enjoy them. This is indicated by the fact that infants between the ages of 3 to 6 month smile less if eye contact is broken in interactions with an adult even though the adult still responds to the infant's actions (Hains & Muir, 1996). Likewise, a study showed that when 4.5 to 8-month-old infants were being addressed with motherese they smile and coo more than the control groups who were spoken to with adult directed speech (Werker & McLeod 1989). This indicates that ostensive signal elicits positive emotions, which in turn can prolong the interaction because the child is more responsive (Csibra, 2010).

When talking to preverbal children, most adults naturally use ostensive signals by giving eye contact and smile at the child while speaking in motherese (Lloyd-Fox et al, 2015). This paper will specifically focus on the potentially ostensive nature of eyebrow flash, but beforehand it will give a brief introduction on some of the scientific literature pertaining to young children's sensitivity and orientation towards the ostensive signals eye contact and motherese. As well as how they might interpret them. This context is important because understanding how those signals serve their purpose in development might give insight into how eyebrow flash functions as well if it is found to be an ostensive signal.

Eye contact and motherese as ostensive signals

By the time an infant is 36 hours, they prefer to look at faces that make eye contact (Batki et al, 2000; Farroni et al, 2002). A preferential looking paradigm study performed by Farroni, Menon & Johnson (2006) showed that infants tested between 24 to 120 hours after birth had a significant preference for looking at faces that made eye contact with them compared to faces that looked to the right or left of them. As well as that preference for direct eye contact disappears when the face they were looking at were turned upside down. This shows that infants orient towards the typical eye contact stimulus from an early age.

Unlike eye contact, speech can facilitate contact with the child even though the face is out of view, or at a distance where it is hard for the young infants undeveloped eyesight to make out the eyes. The newborn infant has however limited ways of knowing if a person is talking to them. Because they don't respond to their name until they are older (Newman, 2005) and the speech they are exposed to in early life are most often not directed at them (Csibra, 2010). Despite this, infants form a preference for motherese soon after birth (Cooper & Aslin, 1990). Motherese contains elevation and variance in pitch and usually repetitiveness in the words that are used (Cooper et al, 1997) and can be exemplified in how mothers usually talk to their infants. By 4.5-8-month children respond more positively to motherese compared to adult directed speech (Werker & McLeod, 1989), and by the time a child is 5-month-old they easily differentiate between motherese and other types of speech as well as other background noise (Colombo et al, 1995). The preference children have for motherese is currently one of the most replicated effects in developmental psychology (The MannyBabies Consortium, 2020)

In addition to attaining the infants attention some findings also indicate that ostensive signals affect learning because of the communicative context they facilitate. A study by Wu, Tummeltshammer, Gliga and Kirkham (2014) showed that 8-month-old children could learn to use a nonverbal stimulus to predict where a target stimulus would appear when the stimulus were paired with a face addressing the child beforehand, even though the face did not give indications to which direction the target would appear. Showing that children can learn from ostensive signals even though the signals in themself do not give instructions about the lesson being learned. Rather, they can serve as attention grabbing stimuli that makes the child ready for information processing (Wu et al, 2014), and possibly serve as a code for learning time.

The ostensive nature of the context can also influence the type of learning taking place. A study by Egyed, Király and Gergey (2013) for example showed that 18-month-old learned in a more generalized way when learning about the valence of an object if valance was precented following clear eye contact. In that study, children who were exposed to eye contact before being shown an object pared with a facial expression generalized the object as good or bad. Whereas the children who were only shown the object paired with the facial expression behaved as if they assumed the liking or disliking was a product of personal preference. When the object was introduced in an ostensive context it might have served as a cue to pay attention, because the following was a direct attempt at teaching them a piece of general information. Therefore, the authors of that study discussed the idea that ostensive signals might be an especially important source of information in the beginning stages of children learning about the world around them (Egyed et al, 2013). In a study using the violation of expectation procedure, Hernik & Csibra (2015) showed that when 13.5-month-old children that were taught the function of a tool within a context of motherese in familiarity trials their expectations about the tool endured despite exposure to counterevidence of the tools function. This was in comparison to a group that were taught the function of the same tool in a non-ostensive way. In this study children were shown two new tools in familiarity trials that were used to manipulate an object in different ways as well as a final trial where the manipulation did not precede as expected. The children that were addressed with motherese looked longer at the later test trial that violated the previously set expectations of the tool despite the final familiarity trial. This indicate that motherese helped them form a generalizable understanding of a tool despite evidence suggesting that the tool could function differently. Overall, it seems like children more readily generalize newly acquired information when it is presented within an ostensive context rather than when they are oriented towards material in a non-ostensive way (Yoon, et al, 2008; Csibra, 2010). This idea is described further in the broader theory known as natural pedagogy (Csibra & György, 2009).

As ostensive signals, both eye-contact and motherese attract the child's attention from birth. In addition to this, they seem to prep and aid in learning situations and might therefore be invaluable in healthy development. Because they also elect a positive response from the infants, they can prolong the interaction (Csibra, 2010) and in so doing also prolonging the opportunity for learning. This might be the case for the eyebrow flash as well if it is found to elicit a preferential response from the children in the current study.

Eyebrow flashes as ostensive signals

At 6-month-old, children have shown an ability to follow a grownups gaze more efficiently when it is preceded by an ostensive signal, like motherese or eye contact with raised eyebrows (Senju & Csibra, 2008). In the field of human ethology, the eyebrow flash is discussed in much the same way as defined by Sperber and Wilson (1995). Author Eibil-Elbesfeldt (2007) notes that raised eyebrows can be seen in mother child-interactions in multiple cultures across the whole world. He describes the eyebrows as the exclamation marks of the face and suggests that the eyebrows can be combined with different expressions to change the meaning of the expression. Alone, he suggests that it usually signifies readiness for contact and attention. Eye contact can be used to establish contact from a distance, but it does not on its own indicate commitment to an interaction. The eyebrow flash on the other hand he argues is usually a signal that grants permission for contact. He followingly argues that eyebrow flashes are innate behaviors, and in his anthropological studies he notes that children and adolescents usually don not vary the time their eyebrows are lifted, but adults do. He thus hypothesizes that the youth are closer to the innate eyebrow flash behavior and later learn how eyebrows are used in their own culture.

A study by Grossmanns, Johanson, Lloyd-Fox, Blasi, Deligianni, Elwell & Csibra (2008) showed that compared to adults that are receiving eye contact, infants show similar nerval activation patterns when they are given eye contact or looking at a person with lifted eyebrows. This suggests that eyebrow flash might be processed in the same way as other ostensive signals neurologically. There is a vast diversity in the literature about how the use of eyebrows in communication is interpreted by humans. Some authors suggest that eyebrow flashes translate as positive intent. For example, when presented with a picture of another person, adults rate faces with high eyebrows as more trustworthy than faces with lower eyebrows (Todorov et al 2008). Raising the eyebrows might then facilitate trust in the recipient, and this might play a part when infants chose who to pay attention to. Author Eibl-Eibesfeld (2007) hypotheses that raising the eyebrows in a communicative context stem from early humans attempts at opening the eyes wider. In elaborating on the idea of using eyebrows to give a better view of the eyes, authors Watt, Craven and Quinn (2007) found that raised eyebrows increases the distance from which a person can make out eye contact. In their study they found that lowering the eyebrows decreased the distance a person's gaze could be detected, but not the accuracy of the gaze perception, meaning the direction the eyes were pointing (Watt et al, 2007).

A potential preference towards eyebrows might be based on learning. Humans might be conditioned at an early age to recognize the eyebrow flash as an ostensive signal, because it has through experience been associated with better access to the eyes, which is as previously mentioned an established ostensive signal. This could then make raised eyebrows a type of secondary ostensive signal conditioned by a primary ostensive signal because they usually occur together. This would mean that recognition of eyebrow flash as an ostensive signal is a product of nurture rather than an innate ostensive signal, much like how a child's name becomes an ostensive signal to children after 13 months of age based on learning

(Newman, 2005). This issue can be elaborated on by further studying newborn infants or investigating age effects where an increased tendency to orient towards the eyebrow flash would be expected if it is conditioned as an ostensive signal through learning.

The importance of this pilot study

The importance of this pilot study is rooted in furthering the theoretical knowledge about the understanding and learning of young children. As well as the results may also provide to be useful in diagnostic aid when infants aren't developing like expected. For example, Osterling and Dawson (1994) found that infants who later developed autism showed among other markers of significantly less interest in eye contact and their name when they were younger. If eyebrow flash is interpreted as an ostensive signal by infants, lack of interest or interpretation of them might also be a marker for vulnerability of later autism spectrum disorder diagnosis. Outside of an academic setting, the results could be used for education of vulnerable populations about the social needs of an infant. Furthering the knowledge of what aids healthy development and learning is also in the interest of the greater public, especially in the current technological climate where infants may have to compete with smartphones and other distractors for their parents' attention.

The implementation and goal of this pilot study

Ideally the children in this study would be tested in person, but because of practical uncertainties and restrictions due to the global Covid 19 pandemic, this pilot study will be conducted online. Recent studies have shown that it is possible to study preferential looking online and still maintain infants' attention as well as produce effects that are comparable to laboratory results (Tran et al, 2017; Scott et al, 2017). Caution should be taken when giving instructions to the caregivers in the administration to prevent inflated exclusion rates and therefore lead to a underpower study (Tran et al, 2017). In contrast to the above cited studies, the current study will be performed with me on video during the trial. This might make it feasible to bypass some of the possible standardization problems an online study might rise. For example, correcting problems with setup or misunderstandings of the instructions that could interfere with validity of the results. In this way this pilot study will also add on to the exploration of the lucrative possibilities of testing infants online.

The goal of the current study is to set a basis for studying young children's sensitivity to eyebrow raise in a systematic way. The aim is to identify if children between the ages 6 to 18 month show a significant preference for looking at a face-like figure that mimics an eyebrow flash over a matched distractor stimulus. In addition to this main analysis, three secondary analyses will be performed respectively: Whether age is a confounding variable in time spent looking at the target stimulus, whether time spent looking at the screen changes significantly during the administration, and lastly whether a preference emerges for either of the two shapes or direction on the screen (left/right).

If eyebrow flash is an ostensive signal it should elicit an orientation towards the stimulus like other well-established ostensive signals do. For a fundament-establishing pilot study such as this, the preferential looking paradigm is a natural choice. This method has traditionally been used to study nonverbal individuals such as infants, the principle of the method hinges on the assumption that if the individuals show a preference between two presented stimuli, then those stimuli can be distinguished between (American psychological association, 2020). Author Spelke writes in her chapter about the preferential looking method that looking behavior is an exploratory behavior that develops early, and that like other exploratory behavior the way in which exploration is preformed is associated with the before-held knowledge of that individual (Spelke, 1985, chapter 13). The preferential looking paradigm allows for a basic dichotomous choice. Seeing as this pilot study is intended as an introduction to the study of eyebrow flash as an ostensive stimulus in the psychological field it is appropriate, as it allows for investigation of the eyebrow flash at its most basic level. If a preference for the intended ostensive stimulus is found, that could be interpreted as the children expecting social interaction from that stimulus and indicate children in the current population is sensitive to the eyebrow flash as an ostensive stimulus. The results of the current study would not lead to an exclusive conclusion that the eyebrow flash is an ostensive signal however, but orientation towards the interpreted ostensive signal is a necessity for the stimulus to be considered ostensive. Whether the eyebrow flash also give rise to positive emotions and influences learning remains a question to be answered by later studies. In the meantime, this pilot study constitutes as an important steppingstone.

Method

Participants

A G*power 3.1 analysis was conducted with an effect size set to medium ($d = 0.5$) and power set to .8, this analysis showed that the minimum number of participants should be $N = 34$. Assuming the normal distribution in the population is short tailed, then 20-30 participants is enough for the central limit theorem to apply and thereby satisfy the assumptions for parametric tests (Field, 2016). Although the sample size goal was 34, an additional stopping rule defined in the pre-registration was that data collection had to be finalized by the middle to end of November to allow for analysis of the results before the deadline.

The data collection cumulated by the end of November without knowing the pending results of the analysis and with a sample size just about half of the ideal sample size. Most studies of infants using looking time as a method use a sample including 11-24 infants per cell (Oakes, 2017). Making the sample size at hand average for this study method. Some of the participants were recruited through the Facebook posts made throughout the collection period, but the greater majority was recruited by directly approaching parents via family and friends. Children were assigned to different stimulus videos according to the order they were tested and none of the participants was compensated for their participation.

In total 18 children were tested, 10 girls and 8 boys. Per parental report, all of them were neurotypical and had no known sensory defects. One of the participants had to be cut from the analysis because of faulty interintromission during administration. The final sample included 9 girls and 8 boys. They had an age range between 6 -18-months-old (*M* = 331 days, *SD* = 103 days), 94% of the children were form Norwegian speaking homes.

Design

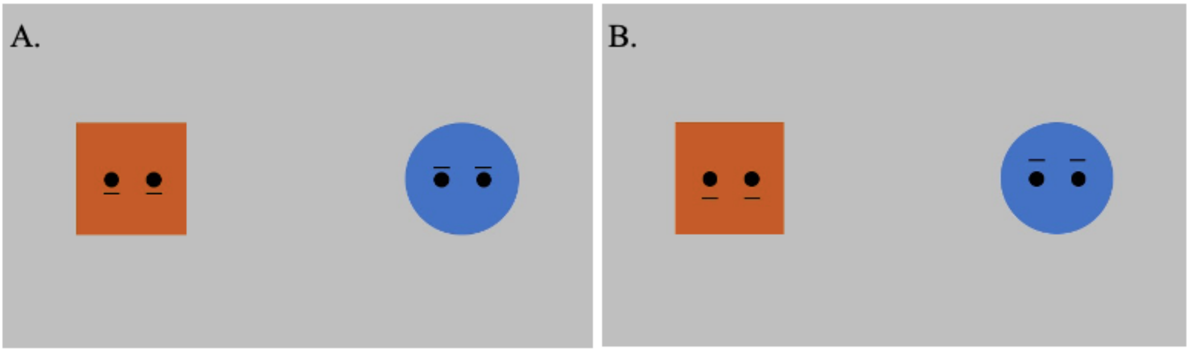
This experiment was conducted with a within-subject design. Using a preferential looking paradigm. Two figures were presented on the screen at the same time, where only one of them was the target stimuli. The independent variable was the dynamic versus static state of the target stimulus, when the figure was in movement in the test trials it constituted as the intended ostensive stimulus. The dependent variable was the preference scores that was calculated using the percentage of time spent looking at the target stimulus.

Materials

The participants were provided with one of four stimulus videos, that were made using iMovie and the figures themselves were made using PowerPoint. The same two shapes appeared in all the videos, these were a blue disc and a red square. The figures each contained two dots and two lines directly placed above (on the target stimulus) or below (on the distractor stimulus) their respective dots. The distance between the features on the figures and the size of the figures themselves were adjusted so that the only difference between the frames was the locations of the lines on the figures. During the test trials the target stimulus were intended to mimic the eyebrow flash, and thus constitute as the ostensive stimulus. The figures appeared together in all the frames and changed position in between trials. Within each of the four video the target stimulus was always the same to allow for learning effect. Figure 1 shows both the distractor and target stimuli of one of the videos.

Figure 1

Frames depicting distractor and target stimulus

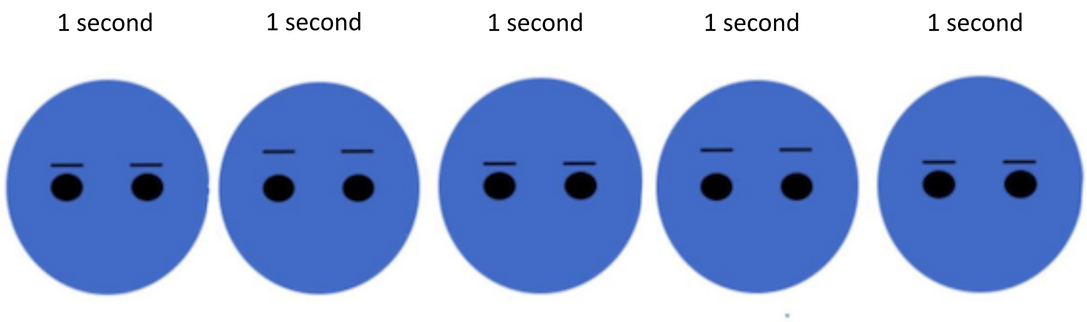


Note. Frame A depicts both the baseline trial as well as the starting possession of the lines in the test trials for one of the videos with the disc being the target. During the test trial this image was shown three times, each time for 1 second. Frame B depicts the frame that made the test trial lines dynamic, here the target stimulus lines are moved up and the distractor stimulus lines are moved down in relation to frame A. This frame lasted for 1 second and was shown two times in a 5 second consecutive test trial.

Prior to the first presented test trial, and in between all the following test trials an attention getter was displayed. The attention getter contained shapes in vivid colors that spun on its own axis, as well as the sound of a bell ringing twice. The first attention getter lasted for 5 seconds, and the intermediate attention getters were shown for 4 seconds each. The administration began with two baseline trials, where the static frame of the target and distractor stimulus played for 5 seconds to record a possible preference for either of them. During the test trials the lines on the target stimulus sequenced the movement pattern lines down, up, down, up, down, each position lasting for 1 second. For the distractor stimulus the pattern was inverted with the pattern up, down, up, down, up. The sequence of the blue disc target stimulus is depicted in figure 2. Here the target stimulus is depicted by itself to show how it moved and the timing of the frames in the stimulus video.

Figure 2

An illustration of the timing and movement on the target stimulus within a test trial



Note. Here the target stimulus is presented by itself with its movement within a single trial that lasted a total of 5 seconds.

The four videos were intended to counterbalance target shapes (disc/square) and orientation of the shapes (left/right) across participants. In two of the videos the orientation of the target (T) and distractor (D) followed the sequencing (left/right) D/T, T/D, D/T, T/D, T/D, D/T, T/D, D/T. In the remaining two videos this sequence was exactly inverse for both shapes, thereby also providing a counter for the sequence across participants. The videos ended with a duck appearing on screen with sounds of quacking. The duck appeared in every corner so that it was possible to identify what it looked like when the child looked at different parts of the screen as opposed to off screen. To code the data resulting from the administration VideoProc Vlogger was used by browsing the recordings frame by frame (25 fps) to identify the number of frames spent looking to the left or to the right.

Procedure

An application for Ethical approval was sent to and granted by the Department of psychology's internal research ethics committee at UiT. Permission to gather and store information about the participants was granted by NSD – Norsk senter for forskningsdata AS (form number 313871). A preregistration of this study was also published in embryo state to Open Science Framework (OSF) before the data collection begun (Suhr & Hernik, 2021).

The participants caregivers were sent an information and consent form by e-mail and a time where the data could be collected was scheduled. They were asked about the child's date of birth and gender. In preparation for the Zoom call, the caregivers were sent an information letter on the practicalities of the Zoom call as well as information about how to act during administration, as well as a link to the video they would be watching (see Appendix A). Because of potential standardization problems associated with testing online, I was present on Zoom when the video was administrated to ensure that problems with setup, or misunderstandings of the instructions could be corrected immediately. A pre-prepared script (see appendix B) were read to the caregivers before the administration proceeded. To control for unintentional help from the parents, such as the parents moving their upper body or turning their head during the trial the parents were asked to close their eyes during the administration. Potential problems with reliability were thus controlled for as best as possible. The whole session was filmed for later coding. The recordings were stored on an encrypted USB.

As detailed in the preregistration, this study was liberal with its inclusion. The children had to provide at least one valid baseline and one test trial during the administration. This resulted in only one of the 18 recruited children being excluded from the final analysis.

To code the data the shared screen was used to identify when the first test trial began, from there the onsets and offsets of the consecutive trials and attention getters were calculated according to the lengths of the stimulus video. Each of the frames during the total eight trials were viewed and coded as looking either left or right with respect to where the target and distractor were oriented in that stimulus video. The frame number of looking off screen were also coded. The number of frames was transferred to an excel sheet where it was counted and divided by the appropriate number of frames to calculate the number of seconds looking at either the target or distractor stimulus.

The amount of time the children looked at either the target stimulus or the distractor stimulus was then summarized as percentages of preference scores (PS). This allowed for comparison between children that spent different amounts of time looking at the screen. For a test trial to be analyzed and included in the dataset the children had to look at the screen for at least 2.5 seconds (50%) of the duration of a specific trial. If not the data for the trial was coded as missing. Administrations in which children were exposed to obvious distractors or help in the environment were not analyzed. Four trials were dismissed in total, one due to interaction from the parent and three because of distraction in the environment. During coding it was discovered that the amount of lag on the recordings increased throughout the playtime. This was made apparent by the change in the light on the participants faces, and because the participants eyes were being drawn to the middle of the screen as opposed to the left or right. This revealed that the attention getter had started sooner than the calculated intervals predicted. Inclusion of the last test trials would thus lead to invalid coding and for this reason, test trials 5 and 6 were cut from the analysis.

Deviations from the preregistered procedure and analyses

The following paragraph is meant for transparency and will summarize the discrepancies between the preregistration and the resulting procedure and analysis in the current study. These changes were made before the data were analyzed, and the results were known. Firstly, the last two test trials were cut from analysis as previously explained. In addition to this the preregistration stated that half of the recordings would be coded by a second person that would be blind to the hypothesis to assess reliability of the scoring. This was planned for half of the ideal sample size. Since the sample size was not attained the double coding was not preformed. A paired sample t-test were used instead of the pre-planned ANOVA because the last two test trials were cut, resulting in two means being compared instead of tree. This would have no effect on the results. In the preregistration there were made plans to do the exploratory analysis, one analysis on differences between genders, one analysis on the course of preferential looking towards the target stimulus within single trials, and one analysis of changes in latency during the administration. In addition to this exploratory analysis was planned if a preference for figure or location was identified. These were intended to investigate if preference was influenced by age and if preference changed during the administration. None of the pre-planned exploratory analyses were preformed, in favour of five other non-planned exploratory analysis. The pre-planned exploratory analysis was not prioritized to keep the numbers of tests low, to safeguarding against inflated probability of spurious significant findings.

Results

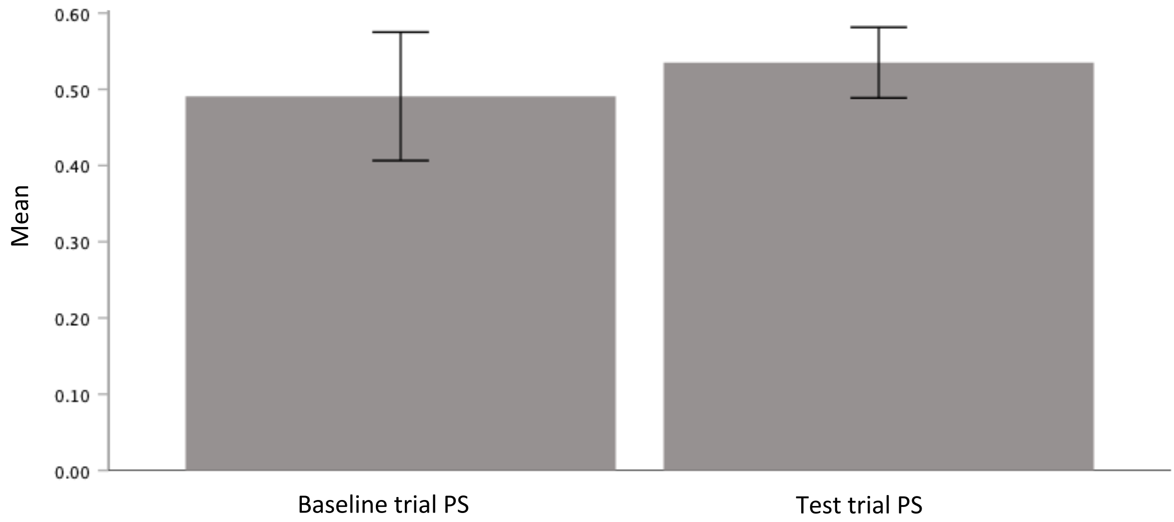
The looking time in both baseline and test trials was converted to mean percentage preference scores (PS) to ensure that the looking time could be compared across participants despite the children having different amounts of looking time during the administration. In the following sections the score for the mean percentage looking time at the target stimulus during baseline is referred to as *baseline trial PS*. The mean percentage time spent looking at the target stimulus during test trials will be referred to as *test trial PS*. These are thus two average percentage scores of the time looking at the screen, meaning that the time during the trials where the children looked off screen is excluded. These scores will be presented as decimals that are rounded up to two. All the statistical tests in this paper were two tailed-tests.

The dataset was explored for violation of the parametric assumptions required to be met by the tests that was preplanned in the preregistration. A Shapiro-Wilk test showed that the difference between the baseline trail PS and test trail PS was not significantly different from a typical normal distribution with a $W(17) = .92$, $p = .20$. The accompanying skew of the baseline trial PS was –0.97 ($SE = 0.55$), and the test trial PS was –0.03 ($SE = 0.55$). The results indicated that the data fit the criteria for normal distribution. Linearity and homogeneity was explored with scatterplots which revealed no obvious pattern, and Levene's test showed that variance was equal for both the baseline trail PS $F(3, 13) = 1$ $p = .43$ and test trial PS $F(3, 13) = 0.56$, $p = .65$. The dataset therefore satisfied the criteria for the parametric tests.

For the primary analysis a paired sample t-test was used. At a group level the infants showed a slight mean preference for looking at the ostensive stimulus with a test trial PS of .54 ($SE = .02$). When compared to the baseline which had a baseline target PS of .49 ($SE = .04$), the difference −.04, 95% CI [−.12, .03] was non-significant $t(16) = -1.32$, $p = .21$. It was accompanied by a small effect size $d = 0.14$. Figure 3 depicts the baseline and test trial PS below. When baseline and test trial PS was compared to chance using a one sample t-test, the difference −.01, 95% CI [−.93, .08], and .04, CI [−.01, .08] was still not significant $t(16) = -0.23$, $p = .82$, $d = 0.16$, and $t(16) = 1.60$, $p = .13$, $d = 0.09$, respectably.

Figure 3

A graphic illustration of the baseline and test trial PS



Note. This figure illustrates 95% confidence intervals around the mean with more than a moderate overlap.

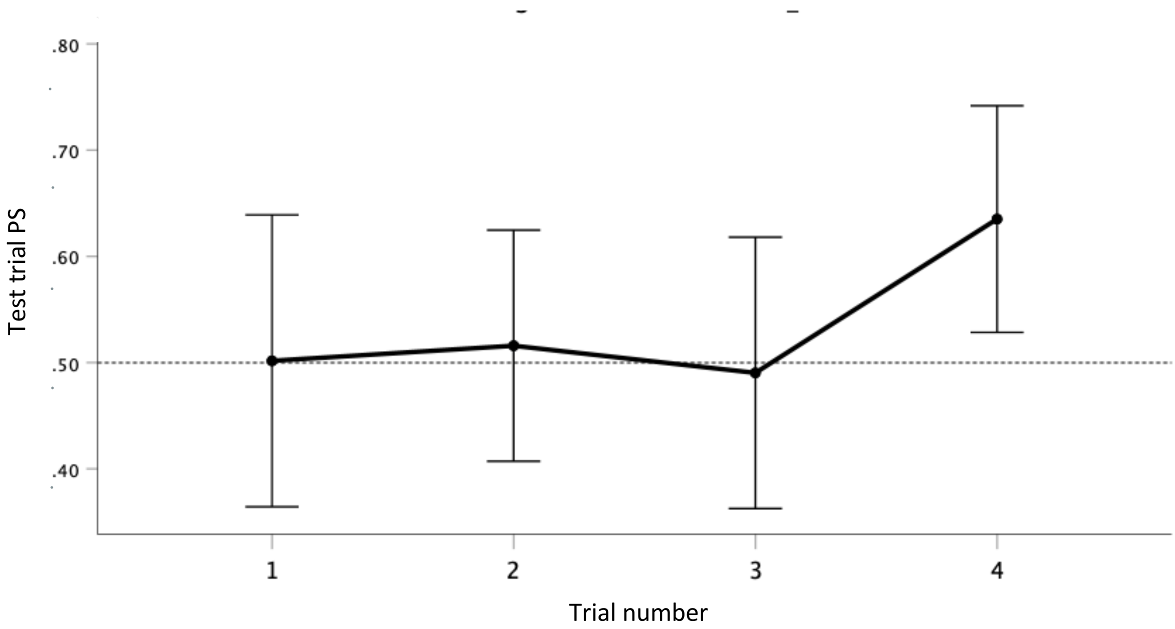
To analyze a possible relationship between age and looking time a bivariate correlational analysis was used with a resulting non-significant correlation of the baseline trial PS $r = .15$, 95% CI [−.35, .59], $p = .56$. This analysis was also non-significant for the test trial PS $r = -.07$, CI [−.53, .43], $p = .79$, To analyze whether there was a significant change in looking time during the consecutive viewing time, mean PS for looking at the target stimulus during test trial 1 + trial 2 and trial 3 + trial 4 was calculated and checked for violation of assumptions (see appendix C). A paired t-test was used instead of the pre-planned ANOVA listed in the preregistration because the final two trials were cut from the dataset. On average, the participants looked longer at the target stimuli in the later trials with a test trial PS of .56, ($SE = .03$) versus .51 ($SE = .03$), this t-test showed that the difference −.05, 95% CI [−.13, .02] was non-significant $t(15) = -1.45$, $p = .17$, $d = 0.15$.

For the four final secondary analysis that were planned in the preregistration, t-tests were used to identify a possible preference or difference in preference due to orientation and shape. Preference scores were made for looking at the target when the target was presented to either left or right the first time it was presented, looking to the right or the left regardless of target, looking at the disc or the square depending on which was the target and finally looking at one of the shapes regardless of which was the target. The PS values all met the assumptions for parametric tests (see appendix C). A paired sample t-test was used to compare PS for both baseline and test trials between the participants where the target was first presented to the left or right, thereby comparing the difference between the two sequences used in the videos. The children in the group that was presented with the target to the left in the first baseline trial had a higher PS (M = .54, SE = .04) than those who were presented with the first target to the right (M = 0.53, SE = .02). This difference .01, 95% CI [−.06, .08], were not significant $t(7) = 0.43$, $p = .677$, $d = .08$. A one sample t-test was used to identify that children looked longer to the left of the screen regardless of orientation of the target stimulus (M = .53, SE = .03), this difference .02, CI [−.05, .10] was however non-significant $t(16) = 0.73$, $p = .48$, $d = 0.14$. An independent t-test showed that the children that were shown the disc as target during the test trials had a higher preference for the target stimuli (M = .55, SE = .03) across test trials versus the preference shown for the target stimuli (M = .50, SE = .03) by the children who had the square as target. This difference .05, CI [−.04, .14] was non-significant $t(15) = 1.23$, $p = .23$, $d = 0.09$. A one sample t-test showed that the children on average looked longer at the disc than the square (M = .53, SE = .02) compared to chance but the t-test showed that this difference, .03, CI [−.01, .08] was also non-significant $t(16) = 1.52$, $p = .15$, $d = 0.09$.

In addition to the pre-planned analysis that was described in the preregistration, an exploratory repeated measures ANOVA and four one sample t-test was conducted. The PS variables used in these tests all met the assumptions for parametric tests (see Appendix C). The ANOVA was used to investigate if there was a significant change in preference for looking time as the stimulus video proceeded. The ANOVA was conducted with the use of all the four test trial PS variables as opposed to the paired t-test described above that used two groups of trials. Because in the planning phase it was expected that a substantial number of trials would have to be cut because of lack of interest in the video. However, only four trials had to be dismissed, an analysis using individual trials was thus possible. Pollai's statistic showed V = 0.27, $F(3,13) = 1.62$, $p = .23$, indicating no significant change in preference during the administration. A graphic illustration of the trend in the test trial PS is shown in figure 4. The final set of t-tests that were conducted is summarized in table 1 and show the results of comparing the test trial PS against chance levels, trial 4 had a PS that was significantly higher than chance levels if an alpha level of .05 were used. When correcting with the Bonferroni correlation the alpha level that is suggested is $p < .001$ if all the tests that were ran on the individual test trial PS is included in the adjustments. When this alpha level is used none of the t-tests in the table should be considered significant.

Figure 4

Profile plot illustrating the change in test trial PS during administration



Note. This profile plot illustrates 95% confidence intervals around the mean. PS values are provided as decimals. The chance level is marked with a dotted line.

Table 1

Exploratory t-test results

Test trial PS	*t*	*df*	95% CI	Cohen's *d*	*p*
Trial 1	0.03	16	–0.13, 0.13	0.25	.98
Trial 2	0.31	15	–0.09, 0.12	0.20	.76
Trial 3	-0.16	15	–0.14, 0.12	0.24	.88
Trial 4	2.7	15	0.03, 0.24	0.20	.02

Discussion

This study used a preferential looking paradigm to identify if children between the ages 6 to 18 months old preferred the potential ostensive stimulus eyebrow flash over a distractor stimulus. The main hypothesis was not supported by statistically significant results. One of the exploratory tests were significant before adjusting for the number of tests used and should be interpreted as such. This t-test showed a significant difference between the PS of test trial 4 versus chance. When considering this finding it is important to keep in mind that this is one of many tests, making the likeliness of spurious significant findings more likely (Field, 2016). When using the Bonferroni correction and adjusting using only the tests that used all of the test trial PS the difference was no longer significant.

The non-significant results can theoretically be explained in a multitude of ways. Six of them being: Movement in the peripheral might have triggered an orientation response, the distractor stimulus might have triggered a violation of expectancy response, noise in the data might have suppressed the effect, the stimuli might have been misinterpreted, and lastly, the study is underpowered or there is no effect. Some of these explanations could by themselves or together account for the results and will be explored in the section below, with accompanying suggestions on how to mediate the problems in the following study, or other studies of similar nature.

Movement in the peripheral might have triggered an orientation response

Choice is the basic principle of a preferential looking time paradigm, but because of how the stimuli was presented, the children might have missed out on the intended ostensive stimulus initially. A preference might therefore have unveiled itself if the administration time was longer or if the last two trials that was cut was included in the final datafile. The trend shown in the profile plots in figure 4 show that the test trial PS increases during the trial even though the ANOVA showed that this increase was non-significant. The exploratory t-tests presented in table 1 show a tendency towards statistically significance in test trial 4, which is in line with the argument that a trend might present itself in later test trials. This might be because the intended ostensive signal was missed in the first three test trials. A solution to this problem might be to present both figures by themself with their respective lines moving in a familiarity trial before the administration trials begins.

At a higher level it could be argued that when the children looked at one shape there would still be visible and potentially distracting movement in their peripheral vision. The human peripheral vision uses almost exclusively rods compared to our central vision that is packed with cons. Rods are especially sensitive to light and movement (Statped, 2021). Because of how the eye functions, the perceived motion in their peripheral vision might have prompted an orientation towards the dynamic object. When using an infant study method called fixation shift paradigm, authors Kulke, Atkinson & Braddick (2015) shows that when children between the age of 1 to 9 months are allowed to fixate on a stimulus in their central vision and an additional stimulus is presented in their peripheral vison competing with the initial stimulus, children show an ability to shift attention towards the new stimulus. The ability to shift attention during this method increases in efficiency with age. Even though the latency in attention shifting was longer in a competing scenario, which is a situation the current study might resemble. The effect was still present in the study when the shifting involved looking away from a face-like stimulus to lines appearing in the periphery. Problems associated with attention shifts in fixation shift paradigm is often seen in clinical populations (Kulke et al, 2017). Therefore, the normal and healthy attention shifting in response to movement in the periphery might in the case of the current study present as a cofounding variable.

If the problem lies with distracting movement in the peripheral, a possible solution might be to make additional changes to the stimulus videos. In a later study both the target and distractor stimulus might be presented on screen alone with their respective lines moving before a static image of both stimulus at the same time is presented on screen. In this setup PS would be made from the static frames, potentially representing preference for the figure the children are expecting further interaction from. This would also mediate the potential problem of missing out on the intended ostensive stimuli presented before, because the children would have time to get familiar with bot stimulus before the test trials.

The distractor stimulus might have triggered a violation of expectancy response

In the current dataset there was a tendency towards looking at the target stimulus for a longer time when the stimulus was moving versus when the stimulus was static. This is in accordance with the main hypothesis even though the difference were non-significant. The lack of a statistically significant results may not be a consequence of lack of an underlying preference for the target stimulus, it may be because of an unexpected preference for the distractor stimulus that interfered with the presumed sensitivity to the intended ostensive signal.

If we assume that both the distractor and target stimulus were interpreted as faces, the distractor stimulus might have skewed the looking preference away from the target stimulus and therefore overshadowed an underlying preference for the target stimulus as an effect of violation of expectancy (VOE). This response entails a surprised reaction to an unexpected scenario or behavior that in turn reveals an expectation about that scenario (Stathl & Feigenson, 2017). The VOE response has traditionally been studied with preferential looking, where infants typically spend significantly more time looking at stimuli that violates their understanding of the world. This effect is not restricted to impossible scenarios, but also entails improbable scenarios (Sim & Xu, 2013) possibly like a face that has moving lines under the eyes instead of over them. If this was the only effect in this study the results would reflect a preference for the distractor stimulus, and thereby lower PS values for the test trials compared to the baseline trials. This is not the case, therefore if the VOE response is to blame for the non-significant results in this study it hinges on the assumption that the target stimulus also was preferred to some degree, in so canceling out a preference for the distractor stimulus.

Stimulus that triggers a VOE response might be given priority when young children choose what to pay attention to, because the tendency to look at unexpected scenarios have been linked to learning. The VOE response has been shown to significantly increase the efficiency of infants learning (Stathl & Feigenson, 2015), and a study by Stathl & Feigenson (2017) found that this effect is still present for children between the ages of 3 and 6 years old in a naturalistic learning environment. A similar response to VOE is prediction error in animal studies. According to Schultz & Dickinson (2000) an animal will adjust its expectations and thereby increase its prediction success in a future scenario to a great extent if the prediction error is large. In animals the need for correcting and adjusting it understandings can easily be interpreted in the context of adaptation for survival. The VOE response in infants and children might share a similar origin and be prioritized above other enticing stimulus, like ostensive signals. Therefore, the current study's failure to produce supporting evidence for the main hypothesis can be a biproduct of the interest children have towards learning about a face that is behaving in an unexpected way.

When this study was planned the VOE response was considered as a possible problem, but that does not mean that the choice of stimuli was ill-considered. There is after all a robust body of evidence suggesting that infants prefer face-like structures over stimulus that should in theory give rise to the VOE response. For example, infants prefer to look at oval shapes with dark shapes in the top, rather than in the bottom (Macchi et al, 2004), they prefer to look at faces that are oriented correctly rather than upside down (Di Giorgio et al, 2012), and they preferer face-like stimulus with darker features (eyes and mouth) over face like stimulus with lighter features than its background (Farroni et al 2005). As well as by the time children are 3 months old, they prefer to look at normal faces, compared to faces with exaggerated features (Bhatt et al, 2005). Together this shows that infants typically orient towards the stimulus they recognize as a face even though the alternative stimulus could evoke the VOE response.

Even though an explanation of the current results might partly be made with the VOE response it is important to clarify that this is just an alternative explanation for the lack of outcome in favor of the hypothesis. If this explanation is assumed, then a prerequisite would be that both the stimuli were interpreted as faces at the same time as that a preference for the target stimulus cancelled out a preference for the stimulus that triggered the VOE response. This makes for a complicated explanation. This would also present a flaw in these kinds of preferential looking studies. Because if the VOE response was used as an explanation for all non-significant results, this method of study would be in violation of the scientific method, because their hypothesis would be unfalsifiable. Lastly if the results were a product of a VOE response the results would show greater variations in response to the dynamic stimuli rather than the baseline. The response variety is not statistically different as proclaimed by Levene's test. In fact, the variance in response were smaller in test trials compared to the baseline trials, as can be seen in figure 2. All things considered, the VOE response is probably not the reason for the non-significant results, even though it deserves conidiation.

Noise in the data due to error might have suppresses the effect

Because of local restrictions due to the Covid 19 pandemic, this study was conducted online. The decision to proceed online were made with knowledge of other successful studies on infants conducted online as sited in the introduction (Tran et al, 2017; Scott et al, 2017). Komarov, Reinecke and Gajos (2013) reported results from tree time-sensitive experiments that were conducted both online and in the lab. Their results showed no significant difference between the online versus lab setting. Showing that studies that uses time as a measurement such as this can be conducted successfully online. One of the greatest problems in this study is that there was an unforeseen amount of lag on the shared screen on some of the recordings. This was possibly due to varying stability of internet connection at the individual participant's homes. A potential solution to this problem would be to ask the participant's caregivers to hold up a mirror behind the children so the onsets and offsets the test trials could be coded more accurately and result in less invalid frames. If this had been done in the current study there might not have been necessary to cut the two last trials, and this might in turn have produced different results.

Even if the problem with the lag was handled, there might still have been discrepancies in the environment. This might have been because of misunderstanding of the information letters. Or the circumstances parents of young children might find themselves in, one of them not being alone in the house. Additional problems that might have interfered with the results are slight differences in the procedure due to for example the relative distance between the screen and the participants faces, the brightness, color output and size of the different screens that were used by the participants. Because the margins in this study was small, an underlying effect might have been suppressed due to any of these sited problems with the testing scenario. Problems that arise due to the changing testing scenario in online testing might not be preventable but seeing as online testing has been done successfully in the past, a solution could be to increase the sample size to ensure that the potential underlying effects are uncovered in later studies.

The stimuli might have been misinterpreted

The lack of preference towards the target stimulus might be because the stimuli were to abstract. The shapes were designed to be simplistic to allow for investigation of the eyebrow flash at its most basic level, but the degree of simplicity might have removed the recognizability of the eyebrow flash. The raising of eyebrows will after all naturalistically appear in relation to other facial features. Making the stimuli as simple and geometrical as they were in the current study might have been pushing the limit for what is recognizable as face templates for young children. For example, when the baseline trials were compared to random chance, there were no significant preference present. If the target stimulus was interpreted as a face, logic predicts a preference for the target stimulus would be present, seeing as young children have a well-documented preference towards faces (Bhatt et al, 2005). The baseline trial was intended to control for a preference towards either stimulus (shape/color). Triggering the facial preference were never an intention, but such a preference was not unexpected.

When comparing the stimuli used in this study to other studies where a face preference was found (Goren et al, 1975; Kleiner 1987; Johnson et al, 1991; Macchi et al 2004; Farroni et al 2005) the stimuli used in the current study is quite different. The most simplistic stimuli used in other studies have usually been oval shaped or contained more facial details like mouths and noses. Or have had orientation of the features to the top part of the figures in contrast to in center of them like in this study. When preference towards no stimuli reveals itself in a preferential looking paradigm one can make the argument that the participants perceived the difference between the stimuli as unsignificant. This might indicate that there is no effect, or that the stimuli did not manage to trigger said effect because it was not perceived as intended. Later studies should therefore consider using stimuli that resemble faces, but control for a general facial preference with baseline trials.

The study is underpowered or there might not be an effect

Because of the typical issues with recruitment that researchers in the field of child psychology often faces, this study is underpowered. The goal was to recruit 34 children, an additional stopping rule was defined in the pre-registration and proclaimed stopping by the end of November to allow for time to analyze and interpret the results. The analysis was therefore conducted with 17 participants which is half of the sample size the G*power analysis deemed as appropriate. This is unfortunate because underpowered studies are associated with problems with interpreting both significant and non-significant outcomes (Fraley & Vazire, 2014). As well as makes for a more fluctuating p-value than in a sufficiently powered study (Halsey, 2015). Big differences can be overlooked in small sample sizes (Field, 2016). Therefore, following studies with less time constraint should recruit more participants.

Most studies on infants using looking time as a method use's 11-24 infants per cell. This includes many of the frequently cited and discussed studies in the child developmental field (Oakes, 2017) Therefore, a decision was made during the initial planning phase of this project that a sample size of about 10 participants would still be acceptable if the recruitment goal were not reached. The target sample size was seen as an ideal goal that would most likely not be attainable given the time restrictions of one semester and was not regarded as a reason to not proceed with this study. Compared to similar studies that have been conducted in the past, the sample size at hand is in the average range. It is important to not dismiss the findings of the study on the sole basis of its small sample size. The results of this study might very well reflect a failure of recognizing the eyebrow flash as an ostensive stimulus.

The results of this pilot concluded in favor of the null hypothesis, reflecting the possibility that eyebrow flash is not an ostensive signal. In a study that compared the findings of looking time studies conducted using infants under the age of 18 months, Oakes (2017) found that most looking time studies that resulted in favor of the hypothesis had medium effect sizes. Because about 80% of the 70 studies included in that study had less than 24 children per cell, an argument can be made that if an effect were to exist it the current population it would be detectable in a sample of 17 participants. If there is at a medium effect in the target population the test trial PS should in theory be higher than a little above chance levels and if an effect is large a small sample size should be able to identify it (Field, 2016). It might also be the case that the effect size associated with a sensitivity to eyebrow flash is so small that its existence is of little significance to the understandings and learning of young children.

A sensitivity towards eyebrow flash might not be reflected in the population at hand. Despite negative results this does not negate the possibility that the eyebrow flash is an ostensive signal for another population. As discussed in the introduction, the eyebrow flash might be interpreted as an ostensive stimulus by older individuals because it has been associated with eye contact. In this study there were no significant correlations with age. If the findings of this study are taken at face value, this could be because the eyebrow flash, unlike eye contact and motherese, is not an innate ostensive signal that is present in the current age group. An effect might appear after 18 months of age. Later studies could test older children and perhaps adults to make this clear.

A final summary

This pilot study concluded with a trend in the results that is in line with what was predicted by the hypothesis, even though the findings were non-significant. It also identifies problems with the present design concerning the stimuli as well as testing scenario. In this way this study has provided a basis for the future study of children's sensitivity to the potential ostensive stimulus eyebrow flash. If the trend is investigated further, consideration should be made to adjust stimuli as suggested and more time should be scheduled for the recruitment period to allow for a more appropriate sample size, as well as potentially testing other age ranges. When considering all the limitations put on this study, as well as the study being well planned out with an average samples size compared to other studies of this kind, support for the null hypothesis stands for the time being.

References

American psychological association. (2020). *Preferential looking technique*. From https://dictionary.apa.org/preferential-looking-technique

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior & Development, 23*(2), 223–229. https://doi.org/10.1016/S0163-6383(01)00037-6

Bhatt, R. S., Bertin, E., Hayden, A. & Reed, A. (2005). Face processing in infancy: developmental changes in the use of different kinds of relational information. *Child development*, *76*(1), 169-181, Doi: https://doi.org/10.1111/j.1467-8624.2005.00837.x

Colombo, J., Frick, J. E., Ryther, J. S., Coldren, J. T. & Mitchell, W. (1995). Infants' detection of analogs of "motherese" in noise. *Merrill-palmer quarterly*, *41*(1), 104-133. From: http://www.jstor.org/stable/23087457

Cooper, R. P, Abraham, J., Berman, S. & Staska, M. (1997). The development of infants' preference for motherese. *Infant behaviour and development, 20*(4), 477-488. Doi: https://doi.org/10.1016/S0163-6383(97)90037-0

Cooper, R. P. & Aslin, R. N. (1990) Preference for infant directed speech in the first month after birth. *Child development, 61*(5), 1584-1595. Doi: https://doi.org/10.2307/1130766

Csibra, G. (2010). Recognizing Communicative intention in infancy. *Mind & language*, *25*(1). 141-168. Doi: https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G. & György, G. (2009). Neural pedagogy. *Trends in cognitive science*, *13*(4). 148-153. Doi: http://doi.org/10.1016/j.tics.2009.01.005

Di Giorgio, E., Pascallis, L. I., Simion, F. (2012). Is the face-perception system human specific at birth? *Developmental psychology*, *48*(4), 1083-1090, Doi: https://doi.org/10.1037/a0026521

Egyed, K., Király, L. & Gergely, G. (2013). Communication shared knowledge in infancy. *Psychological science*, *24*(7), 1348-1353. Doi: https://doi.org/ 10.1177/0956797612471952

Eibl-Eibesfeld, I. (2007). *Human ethology*. Routledge.

Farroni, T., Csibra, G., Simion, F. & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the national academy of science of the united states of America*, *99*(14), 9602-9605. Doi: https://doi.org/10.1073/pnas.152159999

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D. & Csibra, G. (2005). Newborn's preference for face-relevant stimuli: effects of contrast polarity. *Proceedings of natural academy of science of the USA*, *102*(47),17245-17250, Doi: https://doi.org/10.1073/pnas.0502205102

Farroni, T., Menon, E. & Johnson, M. H. (2006). Factors influencing newborns' preference for faces with eye contact. *J Exp Child Psychol, 95*(4), 298-308. Doi: http:// 10.1016/j.jecp.2006.08.001

Fraley, R. C. & Vazire, S. (2014). The n-pact factor: evaluating the quality of empirical journals with respect to sample size and statistical power. *Plos one, 9*(10), 1-12. Doi: https://doi.org/10.1371/journal.pone.0109019

Field, A. (2016). *Discovering statistics using IBM SPSS statistics* (4 ed). Sage.

Frith, C. (2009). Role of facial expressions in social interaction, *Philosophical transactions B, 364*(1535), 3453-3458. Doi: https://doi.org/ 10.1098/rstb.2009.0142

Goren, C. C., Sarty, M. & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544-549.

Grossmann, T., Johanson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C. & Csibra, G. (2008). Early cortical specialization for face-to-face communication in human infants. *Proc Biol Sci*, *275*(1653), 2803-2811. Doi: https://doi.org/10.1098/rspb.2008.0986

Hains, S. M. J. & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child development, 65*(5), 1940-1951. Doi: https://doi.org/10.2307/1131602

Halsey, L. G., Curran-Everett, D., Vowler, S. L. & Drummond, G. B. (2015). The fickle P value generates irreproducible results. *Nature methods, 12*(3), 179-185. Doi: https://doi.org/10.1038/nmeth.3288

Hernik, M. & Csibra, G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of experimental child psychology*, *130*, 176-192. Doi: https://doi.org/10.1016/j.jecp.2014.10.004

Johnson, M. H., Dzirawiec, S., Ellis, H. & Marton, J. (1991). *Cognition, 40*(1-2), 1-19. Doi: http://doi/10.1016/0010-0277(91)90045-6

Kleiner, K. A. (1987). Amplitude and phase spectra as indices of infants' pattern preferences. *Infant behavior and development*, *10*(1), 49-59. Doi: https://doi.org/10.1016/0163-6383(87)90006-3

Komarov, S., Reinecke, K., Gajos, K. Z. (2013). Crowdsourcing performance evaluations of fuser interfaces. *Proceedings of the SIGCHI conference of human factors in computing systems*, 207-216. Doi: https://doi.org/10.1145/2470654.2470684

Kulke, L., Atkinson, J. & Braddick, O. (2015). Automatic detection of attention shifts in infancy: eye tracking in the fixation shift paradigm. *Plus One, 10*(12). Doi: https://doi.org//:10.1371/journal.pone.0142505

Kulke, L., Soranzo A., Christopher, W. & Bertamini, M. (2017). The effect of stimulus size and eccentricity on attention shift latencies. *Vision*, *1*(25), 1-9. Doi: https://doi.org// 10.3390/vision1040025

Lloyd-Fox, S., Széplaki-Köllód, B., Yin, J. & Csibra, G. (2015). Are you talking to me? Neural activation in 6-month-old infants in response to being addressed during natural interactions. *Cortex*, *70*, 35-48. Doi: https://doi.org/10.1016/j.cortex.2015.02.005

Macchi, C. V., Turati, C. & Simion, F. (2004). Can a nonspecific bias toward top- heavy patterns explain newbons' face preference?. *Psychological Science*, *15*(6), 379-383. Doi: https://doi.org/10.1111/j.0956-7976.2004.00688.x

Newman, R. S. (2005) Cocktail party effect in infants revisited: listening to one's name in noise. *Developmental psychology*, *41*(2), 352-362. Doi: https:// doi.org/10.1037/0012-1649.41.2.352

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *The official journal on the international congress of infant studies, 22*(4), 426-469. Doi: https://doi.org/10.1111/infa.12186

Osterling, J. & Dawson G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of autism and developmental disorder, 24*(3), 247-257. Doi: https://doi.org/10.1007/BF02172225

Tran, M., Cabral L., Patel, R. & Cusack, R. (2017). Online recruitment and testing of infants with Mechanical Turk. *Journal of experimental child psychology,* 156(x), 168-178. Doi: https://doi.org/10.1016/j.jecp.2016.12.003

Senju, A. & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current biology*, *18*(9), 668-671. Doi: https://doi.org/10.1016/j.cub.2008.03.059

Schultz, W. & Dickinson, A. (2000). Natural coding of prediction errors. *Animal review of neuroscience, 23*, 473-500, Doi: https://doi.org/ 10.1146/annurev.neuro.23.1.473

Scott, K., Chu J. & Schulz, L. (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies online developmental case studies. *Open mind*, 1(1), 1-15. Doi: https://doi.org/10.1162/OPMI_a_00001

Sim, Z. L. & Xu, F. (2013). Infants' early understanding of coincidences. *Cognitive science, 35*(35), 3402-3407. Permalink: https://escholarship.org/uc/item/6399q853

Spelke, E. S. (1985). *Measurement of audition and vision in the first year of postnatal life: a methodological overview*. Ablex publishing corporation.

Sperber, D. & Wilson, D. (1995). *Relevance: Communication and cognition* (utg.2). Wiley-Blackwell

Stathl, A. E. & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration, *Science*, *348*(6230), 91-94. Doi: http://doi.org/ DOI: 10.1126/science.aaa3799

Stathl, A. E. & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Author manuscript*, *163*(x), 1-4. Doi: http://10.1016/j.cognition.2017.02.008

Statped (26.04.2021). Staver og tapper. From: https://www.statped.no/ferdig-migrert-innhold/fagomrade/syn/undertekster/staver-og-tapper/

Suhr, E., & Hernik, M. (30, 09, 2021). Infant sensitivity to the ostensive signal of eyebrow raise. Retrieved from osf.io/tmkj4

The MannyBabies Consortium (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in methods and practices in psychological science, 3*(1), 24-52. Doi: https://doi.org/10.1177/2515245919900809

Todorov, A., Baron, S. G. & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social cognitive and affective neuroscience,* 3(2), 119-127. Doi: https://doi.org/10.1093/scan/nsn009

Watt, R., Craven, B. & Quinn, S. (2007). A role for eyebrows in regulating the visibility of eye gaze direction. *The quarterly journal of experimental psychology, 60*(9). Doi: https://doi.org/10.1080/17470210701396798

Werker, J. F., McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can J psychol, 43*(2), 230-246. Doi: https://doi.org/10.1037/h0084224

Wu, R., Tummeltshammer, K. S., Gliga, T. & Kirkham, N. Z. (2014). Ostensive signals support learning from novel attention cues during infancy. *Frontiers in psychology, 5*(251), 1-9. Doi: https://doi.org/10.3389/fpsyg.2014.00251

Yoon, J. M. D., Johnson, M. H. & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the national academy of sciences of the United States of America, 105*(36), 13690-13695. Doi: https://doi.org/10.1073/pnas.0804388105