KU LEUVEN

INFANTS USE KNOWLEDGE TO AUGMENT FACE PERCEPTION-EVIDENCE OF TOP-DOWN MODULATION OF PERCEPTION EARLY IN LIFE

STUDENT NAME	**Reema Qaiser Khan**
STUDENT EMAIL	**reemaqaiser.khan@student.kuleuven.be**
STUDENT NUMBER	**0775319**
STUDENT OPTION	**MASTER OF ARTIFICIAL INTELLIGENCE [ECS TRACK]**
DATE	**25th December 2019**

MANUSCRIPTS/REFERENCES:

1. Nielsen Norman Group (2018), *Designing for Kids: Cognitive Considerations,* Received from: https://www.nngroup.com/articles/kids-cognition/
2. Grossmann, Tobias. "The development of emotion perception in face and voice during infancy." *Restorative neurology and neuroscience* 28.2 (2010): 219-236.
3. Kahana-Kalman, Ronit, and Arlene S. Walker-Andrews. "The role of person familiarity in young infants' perception of emotional expressions." *Child development* 72.2 (2001): 352-369.
4. Nawrot, Elizabeth S. "The perception of emotional expression in music: Evidence from infants, children and adults." *Psychology of Music* 31.1 (2003): 75-92.
5. Xiao, Naiqi G., and Lauren L. Emberson. "Infants use knowledge of emotions to augment face perception: Evidence of top-down modulation of perception early in life." *Cognition* 193 (2019): 104019.
6. Sugase, Yasuko, et al. "Global and fine information coded by single neurons in the temporal visual cortex." *Nature* 400.6747 (1999): 869.

Top down processing helps understand what our senses are perceiving in our daily lives. An example is let's say you receive a letter, but it gets drenched because of the rain. Some words have letters that are unreadable. However, you are still able to understand what is written in the letter, this is because of context. Top-down modulation (top down processing) plays an important role in assisting perception of adults. Numerous studies can be found when trying to explore top-down modulation in adult perception. However, a lot of work is being done to probe how infant's perception is being developed using top-down modulation.

In the current paper, the authors use emotional face perception to understand the top-down modulation of infant's perception. The target group used for the experiments comprised of infants who were 9-month-old. The main objective of the researchers was to find evidence of whether it is possible to increase the face perception in the target group based on visual input of faces, and congruent and incongruent sounds.

When sensory input is weak or unclear, the adults use top-down signals to understand their surroundings. Similarly, an infant's perception can be strengthened at an early phase in their life using top-down signals. Top-down signals can enhance the development phase of an infant's perceptual abilities when understanding the world.

According to Piaget's Development Theory, kids' cognitive abilities are still developing. Cognitive abilities include reasoning, inference, and trying to logically connect things. There are two stages when the development happens: 1- The preoperational stage: Language skills are still developing. Children between the ages 2 and 7 think in terms of symbols. 2- The concrete operational stage: Children can logically connect things through inference and reasoning. Children can do this between the ages 7 and 11. Some other considerations where cognitive capacities are immature: Theory of mind- The understanding of emotions and intensions is immature. Cognitive flexibility: The ability to switch perspectives is immature. Executive Function: They are incapable to do planning and observing their own behavior. It is therefore evident, that children between the ages of 2 and 7 have cognitive abilities are still developing. When we compare this to infants, then it becomes more challenging. We can conclude from the theory that if infant's perspective abilities are strengthened, they can perform better as they age in terms of cognitive abilities. [1]

An important social skill of humans is interacting with other people. When they tend to do this, they analyze, read emotional expressions. This ability helps humans to interpret emotional expressions. How this ability forms for infants and how their brain interprets perception of emotion are question dealt with in a recent study. The study found out that when an infants' brain is developing, the faces and voices sensory processing are enhanced by emotions. Infants' react more to emotional negative information provided in the vocal domain as compared to facial domain. Hence, this ability gets developed faster. It is not until 7 months that infants can fully recognize emotional information with respect to face and voice. [2]

A similar study was conducted where it was investigated whether an infant who is 3.5-month-old is able to identify emotional expressions more quickly if the target object of recognition is someone familiar, i.e. mother. The researchers had 72 infant participants. They displayed two facial expressions on film to the infants, the happy facial expression and the sad facial expression. This was accompanied with a vocal expression which could match any one of the two facial expressions, i.e. happy or sad. It was observed that the infant tends to significantly look longer at the emotional expressions when it was a picture of their Mother, accompanied by the

matching vocal expression. Infants who were presented with unfamiliar faces of emotional expression information, did not respond well. It was observed that there may be another factor present apart from the temporal synchrony that infants preferred when responding positively to their Mother's expressions and voice information. The infants were more interactive and positive when they were with their Mothers. The infants' responded with a smile when were made to hear sound specified emotion 'happy' of their Mothers. Negative affect was observed longer among infants who were shown facial expressions of unfamiliar faces and unfamiliar vocal expressions. The study was concluded with the fact that in the early stages of development, contextual information plays an important role to infants as it assists in understanding emotional expression. [3]

A study was conducted in the Minnesota State University Moorhead by Elizabeth S. Nawrot. The study investigated the development of emotion in music among infants, children, and adults. The study was divided into two parts. In study 1, preschool children and adults were used as participants. The objective that was given to this group was to match nine pieces of music to 5 different photographs containing facial expressions. Happy, sad, anger, fear, and neutral were the five facial expressions that were used. Even though the results were not similar between the children and adults when choosing most pieces, however, the children's pattern of responding to music was similar to the adults. In study 2, an experiment of intermodal matching was used on 5 to 9-month-old infants. In this intermodal matching experiment, a dynamic visual display along with happy and sad music were used. It was observed that the infants preferred the concordant happy display. They looked at it for a longer time. The infants did not respond well to the concordant sad display. The study was summed up at the fact that the emotional perception from music is because of innate perceptual predispositions combined with associations that develop in childhood which are learned. [4]

In the current paper, the authors investigated whether 9-month-olds without any familiarization procedure or training can increase face perception by only using predictive and auditory cues. The infants were made to hear a 3 second emotional sound which was either happy or sad, whilst their gaze was positioned in the center of the screen. After that infants were shown neutral face images with no audio being played. The face images were of the size 4.70^0 x 5.80^0. These images were randomly displayed at different locations which was not the center of focus. The authors measured the latency to shift gaze of the infants when they were shown the congruent emotional face. It was observed that the face perception of the infants was increased when met emotional cues. The infants were faster at reacting to emotional faces when they heard an emotional congruent sound with it. The same was not observed when compared to incongruent ones. It therefore does provide evidence that top-down modulation in terms of perception is present early in development.

For Experiment 1, the participants, and materials and procedure were as follows: Participants were 9-month-old infants', 8 females and 10 males. All participants were Asian. The study was conducted in Asia. For the materials and procedure: It starts with participants hearing a 1 sec vocal sound. This sound (laughing and grumbling) is repeated 3 times. While this is happening, on a blank screen a light grey ball is shown on the center. The ball's size changes to get the attention of the infant. The ball was replaced with colorful rings before the sound ends. By doing this, the participant's attention is grabbed in the center where the face images are then displayed. Two face images were displayed. At this point, the audio has been stopped. All images

were of color that were displayed on the grey background. The face images were taken from the NimStim face set. All faces were of female actresses.

Trials followed a 2x2 design. In one trial, the face image was shown of the same person but with two face emotions (happy and angry) and a neutral face expression. There were two vocal sounds with two emotional faces. There were four conditions: Congruent happy face (happy vocal audio and a happy face), Congruent angry face (angry vocal audio and angry face), Incongruent happy face (angry vocal audio and happy face), and Incongruent angry face (happy vocal audio and angry face). For each condition there were 4 trials.

For Experiment 1, the following result were obtained: The initial latency pointed out to the fact that infants reacted more quickly to faces when they were preceded by Congruent emotional sounds. This was not the case when preceded by incongruent vocal audios. The result shows evidence that when infants hear an emotional sound and shown a corresponding face, it boosts the infant's perceptual ability. In the case of 'looking preference'; the results pointed out that the vocal congruent sound could guide the infant to the Congruent images. However, it was not driven by preference as we observed in the Incongruent conditions.

In the case of 'looking preference, which is motivated by initial looking behavior'; the infants preferred the emotional face in the case of congruent conditions. Also, the infants preferred the neutral face in the incongruent conditions.

In Experiment 2, the difference between cross-modal matching and top-down modulation in perception is clarified. Everything is kept same only that the audio preceding visual stimulation has been changed to simultaneous.

For Experiment 2, the participants were 11 females and 7 males. All participants were Asian. These participants were new, they did not previously participate in Experiment 1. The methods and procedure are identical to Experiment 1. The only difference is that the images and audio were presented concurrently for 3 secs.

For Experiment 2, the results were as follows: For the simultaneous presentation of vocal sounds, the result points out that the infants initial looking mechanism differs in both the experimentation. In Experiment 1 it reflects on top-down modulation of perception, instead of cross modal matching.

In the case of 'proportional looking time for simultaneous presentation'; by playing simultaneous sounds it did not affect the infant's visual preference. In the 'time-course of infants' preference is unaffected by simultaneous vocal cues'; the results did not match to Experiment 1. Experiment 2 did not have any effect on face perception due to congruency. The infants did not look longer at the Congruent face.

In Experiment 3, the same experiment was repeated as Experiment 1, only that the infants were of the age of 6.5 months. This experiment's objective was to prove that such young infants will not be able to show signs of top-down modulation in face perception.

In Experiment 3, the participants, and methods and procedure were as follows: There were 18 infants, 7 females and 11 males. All participants were Asian. None of these participants previously participated in Experiment 1 or in Experiment 2.

The procedure and methods were identical to Experiment 1. The results were as follows: In the 'Initial looking latency not affected by vocal cues'; the 6.5-month-old visual preference was not affected when tested between neutral face and emotional face. In the case of 'looking preference is not affected by initial looking behavior'; the 6.5-month-old were not able to use emotional

vocal sound to increase their perception of facial expression because of the absence of congruency effect. Therefore, the study does provide evidence that top-down modulation is present early in development. [5]

There are some questions that I would like to highlight regarding the current research. In the current research, if the participants belonged to another ethnicity, would the results differ? Is it possible that the babies were only reacting to images rather than to vocal audios? According to picture superiority effect, images and pictures are more likely to be remembered than words and audios. This is based on the symbolic modality of presentation of information that the human brain is sensitive to. Is it possible that the babies develop the picture superiority effect in their younger stages of life? Looking at the dataset of images, the researchers used the images of female actresses. If the dataset images were replaced with images of any one of the following: man, older man, older woman, man of another ethnicity, female of another ethnicity; will the results differ? If audio sounds of different categories of people were used, will the results be the same? Were the babies medically examined before the experiment to ensure that their hearing and visual capability is strong, and they aren't malnourished? Another important aspect that arises regarding the participants of the experiment is the type of environment the baby is being raised. Is the baby being raised around grandparents or not, or a single parent? There are numerous studies that suggest that when a baby is raised among grandparents, it builds resilience in them and makes them less ageist. What results would be achieved if the participants were pair of twins? Will both the siblings have the same visual and vocal preference? Another important question is, what was the proximity of the mothers and their babies during the experiment? Was the mother of each baby inside the room when the experiment was being conducted? The drawback to this is that it may influence the result and may lead to unauthentic results. Also, was the baby agitated if the mother was not near? These things will affect the result of the experiments.

The Cognitive Science book talks in detail about the concept of top-down processing (modulation). It begins by first explaining about inverse-problem. For example, there are several triangles of different sizes, if these triangles are each placed at a certain angle, they can form an equilateral triangle. A person will be able to see this triangle when standing at a certain angle. This information is hard to interpret from the physical world which leads to information that is lost by the eye, but the brain needs to recover this information. This is known as inverse problem. In order to tackle this problem, the concept of flow of information is introduced, where two types of principles are analyzed: Bottom-up and Top-Down Processing. Bottom-Up processing is the notion that the sensory information that is entering continuously is being processed continuously. For example, if a picture of a car is flashed on the screen, the eyes detect the features, the brain connects the information, and one will perceive it as a car. Top-Down Processing is a type of processing where perception is driven by cognition. For example, as explained in the Cognitive Science class, consider an object that looks like a number 13 but at the same time may be perceived as the letter B. Brain is finding it hard to identify. There is no Top-Down modulation here. However, to be certain we need context. If this object is placed between the letters A and C, Top-Down processing tells us that this is the letter B. Similarly, If the object

is placed between the numbers 2 and 5, Top-Down processing tells us that these are the numbers 1 and 3, giving us 2135. In the current research the authors have provided perception evidence of Top-Down processing in infants. In the experiment, the context for the babies was the audios and the pictures. The babies reacted well when experimented with congruent examples. In the case of incongruent examples, the context was confusing that is why the babies did not react well to it.

The book highlights on emotions and perception, which is another aspect to the current research. The book mentions a research paper Sugase et al.,1999 [6], in this paper the authors provide evidence that facial expression information is the fastest information that is understood by the brain than any other information. Emotions are associated with facial expressions. One such theory, the James-Lange theory which seems to be present in the research that is being evaluated for this essay. The theory states that due to physiological reactions to events, this will result in emotions. For example, in the current research when babies were being experimented using congruent sounds and pictures, they reacted very well. When a happy audio or angry audio was played followed by a happy or angry picture in the congruent experiment; the babies reacted well mainly because of the facial expressions supported by congruent sounds which activates the physiological reaction in turn creating emotions. This shows evidence of James-Lange theory. It is still hard to understand why James-Lange theory seem to fail in the incongruent condition. Is it because of the confusing context?

Top-Down modulation is an impressive theory belonging to the field of Cognitive Science. Looking at the amount of research that is being conducted around this theory is monumental. The current research being one of them. The research has been conducted in a very meticulous manner. The way the researchers designed the experiment is indeed worth appreciating. The way they used colored rings to grab the attention of an infant is a unique approach. However, for future work, the authors should consider answering the questions highlighted in this essay. The role of perception is also worth appreciating, the engrossing knowledge of perception, how it is one of the most important proficiency in developing the intelligent characteristics of a human being. Perception is an active process. Emotions is a worthy endeavor. The way it constitutes in structuring human behavior is astounding. Infant's showing such signs of intelligence in early life is very perplexing. Signs of identifying familiar faces, congruency, understanding context, and many more is simply amazing. In conclusion, after thoroughly examining how Top-Down modulation is present in the early stages of a human's life, just proves how powerful, robust, and complex the brain is.

Word Count:3080

END