

shown that children with atypical development, where autism or intellectual disability (Embregts & Van Nieuwenhuijzen 2009) or behavioral problems (De Castro et al. 2005) are involved, exhibit characteristic patterns in some SIP steps. These characteristic patterns may in part explain their maladaptive responses to social situations. However, little is known regarding the DHH population.

HEARING STATUS AND EMOTION UNDERSTANDING IN SOCIAL SITUATIONS

Emotion understanding cannot develop without access to the social context in which emotions occur (Crick & Dodge 1994; Rieffe et al. 2015). In a social environment that features spoken communication, DHH children do not access the social world in the same way as their TH peers. This might negatively affect their emotion understanding. During the first years of life, parent–child interactions have already been shown to be different. TH parents with DHH children use more commands, shorter utterances, less mental-state language, and less turn-taking during conversation, compared to TH parents with TH children (Morgan et al. 2014; Dirks et al. 2020). DHH children also miss many cues from the environment when their attention is not directed to the source of the cue (Calderon & Greenberg 2011). Moreover, even when attention is directed to the source, DHH children might still recruit only partial information, for example, due to background noise (Leibold et al. 2013).

To the best of our knowledge, there is only one study to date on DHH individuals' emotion understanding in naturalistic, dynamic social situations, and the study also applied the SIP model (Torres et al. 2016). Torres et al. (2016) interviewed DHH and TH participants (13–21 years old) who had watched videos depicting dynamic social interactions. They found the two groups to differ in all SIP steps: compared to TH participants, DHH participants encoded less relevant cues as defined by the study; made more misinterpretations (i.e., interpreting nonhostile situations as hostile or vice versa); formulated more goals that were either aggressive or not effective (e.g., crying); and generated, decided upon, and enacted more aggressive or ineffective responses. Similarly, in studies that used static stimulus materials (e.g., drawings or photos) and examined children's interpretation of emotions in social situations, more misinterpretations were reported in DHH children than in TH children (2.5 to 8 years old; Gray et al. 2007; Wieferrink et al. 2013). Based on the SIP model, it is reasonable to suspect that DHH participants' difficulties could result from a different encoding pattern that occurs during the first processing step. However, this relation was not examined by Torres et al. (2016).

In any real-life social situation, a considerable amount of information is available. In theory, the encoding stage works as a filter through which people collect the most relevant emotional cues for subsequent processing. Empirical evidence from studies on typical development has shown that the head region of others is an important cue to which people most often direct their attention, when processing social situations. It provides a rich source of information that allows for inferences about other people's attention, intentions, and feelings (Birmingham et al. 2009; End & Gamer 2017). In fact, soon after birth, infants are already more responsive to face patterns than to nonface patterns (Slater & Quinn 2001; Farroni et al. 2005). They also show longer sustained attention to faces than to nonface patterns

when a distractor appeared (Ahtola et al. 2014; Pyykkö et al. 2019), and continued to improve their ability to detect faces or heads among other objects during the first year of life (Frank et al. 2014; Reynolds & Roth, 2018). Even when facial expressions cannot be seen clearly, the head region remains important because its angle (e.g., bowed or raised), orientation, and movement provide information about the emotions and attention of other people (Main et al. 2010; Hess et al. 2015). While cues such as body posture and actions also carry emotional value needed for adequate emotion understanding (Dael et al. 2012), children and adults alike still look at heads for a longer period of time than at other cues, when asked to recognize emotions in dynamic social situations (Nelson & Mondloch 2018).

Yet to date, no research has examined how DHH individuals visually encode emotions in social scenes. Such investigations could be particularly relevant to the DHH population because DHH individuals may collect visual cues in a way that is different from TH individuals, to support daily communication. For example, empirical evidence has shown that DHH adults allocate their visual attention over a wider area in a scene than TH adults when searching for an object, as a strategy to compensate for limited auditory input (e.g., Proksch & Bavelier 2002; Sladen et al. 2005). Similarly, when DHH adults were required to recognize emotions from isolated faces, they distributed their attention equally to the eye and mouth regions, whereas TH adults focused on the eyes (Letourneau & Mitchell 2011). This tendency to allocate visual resources over a wider area may mean that DHH individuals give more weight to body cues than to the head itself, when processing social interactions between people (Rollman & Harrison 1996).

THE PRESENT STUDY

In this study, we wanted to build on the study by Torres et al. (2016). Our first aim was to investigate more closely how DHH children encoded emotional cues in dynamic social situations by using eye-tracking technology. By checking children's gaze patterns, we aimed to understand how long DHH and TH children attended to different nonverbal emotional cues that we preselected. These cues included the head, the body, and actions. Considering the importance of heads (End & Gamer 2017; Nelson & Mondloch 2018), we expected DHH and TH children to show a stronger focus on heads than on the other cues. Yet, given the finding by Torres et al. (2016) that DHH participants encoded less relevant cues than TH participants, we expected the total time spent looking at the preselected emotional cues to be shorter in DHH children than in TH children. We did not make specific hypotheses as to which emotional cues would be viewed for a shorter duration, due to a lack of evidence.

Second, we examined how DHH and TH children interpreted emotions in social situations. Based on Torres et al. (2016), we expected that DHH children would interpret the situations with the emotion predicted by the study less often than TH children. Thus, DHH children were expected to be less accurate than their TH peers when identifying and labeling the emotion triggered in social situations.

Third, for exploratory purposes, we examined how DHH and TH children's encoding patterns were associated with their interpretation of emotions in dynamic social situations. Previously, this relation has only been studied in children with

aggressive or antisocial behaviors who were found to encode less relevant emotional cues, followed by a less thorough exploration of their social environment (Horsley et al. 2010; Troop-Gordon et al. 2018). Due to the lack of previous studies on the DHH population, we explored this association without making specific hypotheses.

Fourth, we expected to find an age effect. We predicted that in both DHH and TH children, an increase in age would be associated with an increase in time spent looking at the preselected emotional cues and in interpretation accuracy.

METHODS

Participants

Sixty DHH and 71 TH children in Taiwan of ages 3 to 10 years old participated in this study. Inclusion criteria for DHH children were mild-to-profound congenital or prelingual bilateral hearing loss; use of hearing aids or cochlear implants; use of spoken language as the primary communication mode; and ages between 3 and 10 years old. Inclusion criteria for TH children were typical bilateral hearing (as reported by parents); use of spoken language as the primary communication mode; and age between 3 and 10 years old. Children with additional disabilities and diagnoses other than hearing loss were excluded from recruitment (e.g., autism, language disorder, or attention deficit hyperactivity disorder, according to parent or teacher report). The two groups did not differ in demographic characteristics (see Table 1).

We approached 153 children in total. For this study, we excluded children who did not finish 50% of the trials ($n = 4$); children who had additional disorders or nonverbal intelligence measured at 2 SD lower or higher than the population mean, or as indicated by teachers or clinicians ($n = 17$); and a child who was tested twice by mistake. The excluded sample did not differ from the included sample in age, gender distribution, or nonverbal intelligence, but had lower parental educational level, $t(141) = 2.31$, $p = 0.022$. See Supplemental Digital Content 1 <http://links.lww.com/EANDH/A743> (Text) for details about sample size justification.

This study was part of a research project that examines social-emotional functioning in children with typical and atypical development. Permission for the study was granted by The Psychology Ethics Committee of Leiden University and Chang-Gung Memorial Hospital Ethics Committee for Human Studies. All parents provided written informed consent. Children between 3 and 10 years old were recruited for this project because previous studies have shown that the period from preschool to middle childhood is crucial for children's development and mastery of skills for understanding the categories and causes of basic emotions expressed by other people (e.g., Kolb et al. 1992; Rieffe et al. 2005; Durand et al. 2007). In addition, recruitment was conducted through a cochlear implant center, and the majority of the DHH children we recruited had profound hearing loss and one cochlear implant. However, we did examine the effects of age and different degrees of hearing loss by including age as a covariate in

TABLE 1. Characteristics of the participants

| Characteristic | DHH | TH | <i>P</i> |
|------------------------------------|-------------|--------------|--|
| N | 60 | 71 | |
| Age, yr (SD) | 6.32 (2.08) | 6.07 (1.75) | $t(129) = -0.74, p = 0.458$ |
| Girls, n (%) | 30 (50%) | 41 (58%) | $\chi^2(1, N = 131) = 0.79, p = 0.375$ |
| Nonverbal intelligence (SD)* | 9.51 (2.67) | 10.31 (2.61) | $t(121) = 1.66, p = 0.099$ |
| Parental education level (SD)† | 3.31 (1.03) | 3.58 (0.96) | $t(121) = 1.47, p = 0.143$ |
| Age at amplification, yr (SD) | 2.51 (1.29) | | |
| Duration of amplification, yr (SD) | 3.81 (1.87) | | |
| Degree of hearing loss, n (%) | | | |
| Mild (26–40 dB) | 2 (3%) | | |
| Moderate (41–60 dB) | 2 (3%) | | |
| Severe (60–80 dB) | 1 (2%) | | |
| Profound (> 80 dB) | 55 (92%) | | |
| Type of amplification, n (%) | | | |
| Unilateral cochlear implant | 40 (67%) | | |
| Bilateral cochlear implants | 15 (25%) | | |
| Hearing aid only | 5 (8%) | | |
| Etiology, n (%) | | | |
| Congenital | 29 (48%) | | |
| Inner ear anomaly | 14 (23%) | | |
| Waardenburg syndrome | 1 (2%) | | |
| Auditory neuropathy | 2 (3%) | | |
| Unknown | 14 (23%) | | |
| Type of education, n (%) | | | |
| Regular schools | 57 (95%) | | |
| Special schools | 3 (5%) | | |

DHH indicates deaf and hard-of-hearing; TH, typically hearing.

*For nonverbal intelligence, age-corrected norm scores are presented. The grand population mean is 10 and the SD is 3. Children 3 to 5 years old were tested with Block Design and Matrix subscales of the Wechsler Preschool and Primary Scale of Intelligence Revised Edition (Wechsler 1989). Children 6 to 10 years of age were tested with Block Design and Picture Arrangement subscales of the Wechsler Intelligence Scale for Children Third Edition (Wechsler 1991). These tests were used because the experimenter had access to them and had received training to administer these versions.

†Parental education level: 1 = no/primary education; 2 = lower general secondary education; 3 = higher general secondary education; and 4 = college/university.

all analyses, and by checking how results differed when analyses were conducted separately for children with a cochlear implant instead of for the entire DHH group.

MATERIALS AND PROCEDURE

Emotion Understanding Task

This task was designed specifically for this study [see Supplemental Digital Content 2 <http://links.lww.com/EANDH/A743> (Text and Table) for information about video validation]. Children conducted the task in a quiet room at a cochlear implant center or at participants' schools. A Tobii X3-120 eye tracker (Tobii Technology, Sweden) was placed 65 cm away from the participant's eyes and mounted below a computer screen. After a centrally presented fixation cross (1000 ms), children watched eight silent videos. Each was 10 to 15 seconds long. Each video showed a prototypical social situation where an interaction partner triggered either a positive emotion (happiness, excitement, or pleasant surprise) or a negative emotion (anger, sadness, or fear) in a target person. After each video, the participant was presented with six different still photographs of the face of the target person, each expressing a different emotion. The child was then asked to give a nonverbal response (by pointing to the face showing the emotion that the child thought the target person would express, at the end of each video) and a verbal response (by labelling the emotion).

Before the experiment, a five-point calibration was conducted, and then children completed two practice trials. The structure of these practice trials was similar to the testing trials, but with pictures that allowed for step-by-step instruction. All children gave nonverbal and verbal answers within the expected valence in the practice trials, so they all continued to the testing trials.

Stimuli

All videos were silent, displayed at a size of 640 × 480 pixels and centrally presented on a computer screen of 1024 × 768 pixels. The videos started with a contextual scene: A target person showed an initial emotional state on the face, and an interaction partner entered. Next, in a key-action scene, the partner carried out an emotion-eliciting action on the target person. In the key-action scene, the target person was viewed from the side by the camera, so that about three quarters of the face was not visible. A red arrow pointing to the target person was presented at the end of each video, to ensure that the child understood which person was the target person.

Each video had a counterpart that showed a parallel contextual scene and key-action scene that ended in an opposite emotional valence. For example, one video showed a man with a broken leg walking with sticks and a woman passing by (the contextual scene), and then the woman laughing at the man (the key-action scene). Its counterpart video showed the same contextual scene, but it ended with the woman giving a cake to the man. The counterpart videos were put into two different sets of tests and watched by different children, who were randomly assigned to one of the sets.

We only considered the key-action scene in our analyses because this was where the emotion was triggered in the target person. The key-action scenes were 5 to 9 seconds long

(mean = 6838.13 ms, SD = 1038.84 ms). See Supplemental Digital Content 2 <http://links.lww.com/EANDH/A743> (Text and Table) for more details about stimuli.

Encoding of Emotional Cues

We preselected four areas of interest (AOIs) that corresponded with four emotional cues: the target person's head, the target person's body, the partner's head, and the partner's action (e.g., an arm engaged in the action of pushing or pointing at someone; see Fig. 1). Children's eye movements were measured by the eye tracker at a sampling rate of 120 Hz using corneal reflection techniques, and processed by Tobii Studio 3.2.1. Using a Dynamic AOI tool provided by the software, we predefined AOIs by drawing freeform shapes, frame by frame. Tobii I-VT fixation filter was applied to define fixations. Only the sampling points where both eyes were detected by the eye tracker were included in later calculations. The total fixation duration within each AOI was the sum of the duration times for all fixations within an AOI. Fixation ratios were calculated by dividing total fixation durations within each AOI by the total fixation duration within the entire screen. Moreover, we calculated the fixation ratio within the video frame, to examine attention to the videos.

Interpretation of Emotions

Children's ability to identify an emotion nonverbally (by pointing at a picture of an emotional expression) and verbally (by labelling an emotion) was scored on a three-point scale: 2 = the emotion predicted by the study; 1 = other emotions within the valence predicted by the study; and 0 = not within the valence intended by the study. For example, in a video predicted to trigger anger (e.g., someone falls from a bike and is laughed at by a passer-by), a score of 2 was given to children who stated anger or its synonyms (e.g., rage); a score of 1 for answers that fell within the negative valence (e.g., sad/depressed, fearful/scared, unhappy, and upset); and a score of 0 for any positive emotions or answers that referred to actions rather than emotions (e.g., crying and shouting).

Note that people's responses toward an emotion-evoking situation are affected by their action tendencies, which reflect the goal they aim to achieve in that particular situation. In turn, this action tendency defines which emotion is expressed (Frijda 1986; Rieffe et al. 2005). Yet, the same emotion-evoking situation can evoke different action tendencies in different children. Consequently, individual differences toward the same emotion-evoking situation are possible, which can result in different emotion expressions. In the example situation given above, a person can feel angry because of the passer-by's socially inappropriate behavior; or, a person may feel sad because he/she thinks his/her riding skill is poor. Therefore, different emotions within the same valence (positive or negative) can be a correct or appropriate emotional response toward the same emotion-evoking situation. However, some emotions are considered more prototypical, that is, in line with common knowledge and expectations, such as feeling angry when your bike is stolen (focusing on the aggressor) or sad when your cat died (focusing on the loss).



Fig. 1. An example last scene of video presentation. The areas within the white lines denote the four areas of interest used in the analyses (solid lines denoting the target person; dashed lines denoting the interaction partner). The white lines were not presented to the participants during the experiment.

ANALYSIS AND RESULTS

Because the data had a two-level structure of trials (level 1) nested in participants (level 2), we used generalized linear mixed models (GLMMs) for analyses. The fixed factors included in each model are specified below, and all models included a random intercept for each participant. Centering was used on all independent variables. Age was added as a covariate in all models. Following a standard model selection procedure, nonsignificant factors were removed one by one from the full model, starting with higher-order interactions, to derive a final model. Factors with a trend toward significance could stay in the final model if removing them worsened model fit. A lower -2

log likelihood indicates a better model fit. Normal probability plots were used to inspect the normality of the residuals. The residuals of the fixation ratios within the AOIs and of the interpretation scores were close to a normal distribution, while the residuals of the fixation ratios within the video frame were not (data were negatively skewed). Therefore, GLMMs with a normal distribution were used to model the fixation ratios within the AOIs and the interpretation scores. The fixation ratios within the video frame were modeled with a GLMM where a binomial distribution (link function = logit) was selected. Level of significance was set at $P < 0.05$. Standardized effect sizes (δ) were calculated based on the study by Raudenbush and Liu (2000). The authors extended Cohen's approach to fit multilevel

TABLE 2. Fixed and random effects in the generalized linear mixed models

| Fixed and random effect | Fixation ratio within AOIs | | Interpretation | | Effect of encoding | | | |
|-------------------------|----------------------------|----------------|-----------------------|----------------|-----------------------|----------------|-----------------------|----------------|
| | <i>b</i> (δ) | 95% CI | <i>b</i> (δ) | 95% CI | Nonverbal | | Verbal | |
| | <i>b</i> (δ) | 95% CI | <i>b</i> (δ) | 95% CI | <i>b</i> (δ) | 95% CI | <i>b</i> (δ) | 95% CI |
| Intercept | 0.18 | (0.16, 0.19) | 1.13 | (1.06, 1.20) | 1.18 | (1.11, 1.25) | 1.11 | (1.05, 1.18) |
| Age | 0.00 (0.01) | (0.00, 0.00) | 0.01 (0.01) | (0.01, 0.01) | 0.01 (0.02) | (0.01, 0.01) | 0.01 (0.02) | (0.01, 0.01) |
| Group | -0.03 (0.16) | (-0.05, -0.00) | -0.13 (0.20) | (-0.22, -0.04) | -0.06 (0.09) | (-0.17, 0.04) | -0.11 (0.18) | (-0.21, -0.01) |
| Valence | <i>ns</i> | | <i>ns</i> | | <i>ns</i> | | <i>ns</i> | |
| Task | — | | 0.08 (0.12) | (0.02, 0.13) | — | | — | |
| Group × Valence | <i>ns</i> | | <i>ns</i> | | <i>ns</i> | | <i>ns</i> | |
| Group × Valence × Task | — | | <i>ns</i> | | — | | — | |
| Target Head | <i>ref</i> | | — | | 0.14 (0.22) | (-0.20, 0.48) | 0.37 (0.61) | (0.10, 0.64) |
| Target Body | -0.07 (0.44) | (-0.09, -0.05) | — | | -0.32 (0.50) | (-0.61, -0.03) | 0.05 (0.08) | (-0.32, 0.42) |
| Partner Head | 0.05 (0.31) | (0.03, 0.07) | — | | -0.27 (0.41) | (-0.51, -0.02) | <i>ns</i> | |
| Partner Action | -0.07 (0.44) | (-0.09, -0.05) | — | | <i>ns</i> | | -0.35 (0.58) | (-0.65, -0.06) |
| Group × Target Head | <i>ref</i> | | — | | 0.61 (0.94) | (.06, 1.15) | <i>ns</i> | |
| Group × Target Body | 0.04 (0.25) | (0.01, 0.07) | — | | <i>ns</i> | | -0.54 (0.89) | (-1.05, -0.03) |
| Group × Partner Head | 0.03 (0.18) | (0.00, 0.06) | — | | <i>ns</i> | | <i>ns</i> | |
| Group × Partner Action | 0.03 (0.18) | (-0.00, 0.05) | — | | <i>ns</i> | | <i>ns</i> | |
| Variance – Intercept | 0.00 | (0.00, 0.00) | 0.03 | (0.02, 0.06) | 0.03 | (0.02, 0.06) | 0.02 | (0.01, 0.05) |
| Residual | 0.03 | (0.02, 0.03) | 0.41 | (0.37, 0.45) | 0.39 | (0.35, 0.43) | 0.35 | (0.31, 0.38) |

Group was coded as $-1 = DHH$, $1 = TH$. Valence was coded as $-1 = negative$, $1 = positive$. Task was coded as $-1 = nonverbal$ interpretation, $1 = verbal$ interpretation. AOI was coded as $-2 = interaction$ partner's head, $-1 = interaction$ partner's action, $1 = target$ person's body, $2 = target$ person's head. The last category was used as the reference ("ref"). An "ns" indicates that the variable was removed from the final model due to insignificance. An "—" indicates that the variable was not included in the full model. Significant fixed effects ($P < 0.05$) are bolded. CI indicates confidence interval. δ = standardized effect size, calculated by dividing fixed coefficient by the square root of the sum of Level 1 (residual) and Level 2 (intercept) variances (formula suggested by Raudenbush and Liu 2000).

research, and suggested effect sizes of 0.20, 0.50, and 0.80 to be small, medium, and large, respectively. Statistical analyses were carried out using SPSS 23.0 (IBM Corp., Armonk, NY), except for the GLMM with a binomial distribution, which was conducted using the *glmmTMB* package in R version 3.6.3.

Complete models are specified in Tables 2 and Supplemental Digital Content 3 (Tables) <http://links.lww.com/EANDH/A743>. We also conducted analyses with the children with a cochlear implant separately ($n = 55$), instead of with the entire DHH group. However, the directions of the results did not change. We report the results where children with a cochlear implant and children with a hearing aid were analyzed as one group. See Supplemental Digital Content 4 <http://links.lww.com/EANDH/A743> (Text and Table) for separate analyses on the children with a cochlear implant. In addition, Supplemental Digital Content 5 <http://links.lww.com/EANDH/A743> (Table) shows the correlations between study variables and hearing-related variables within the DHH group.

MISSING DATA

Nonverbal intelligence scores were missing for eight children (7 DHH and 1 TH). Eight parents (6 DHH and 2 TH) did not provide information about their own educational level. Six children (3 DHH and 3 TH) did not have eye-tracking data due to device failure. Little's MCAR test (Little, 1988) showed that the data were missing completely at random, $\chi^2 = 597.98$, $df = 726$, $p = 1.00$.

ENCODING

First, we checked the fixation ratios within the video frame. Fixed effects included Age, Group (2: DHH and TH), Valence (2: Positive and Negative), and Group \times Valence. Since no effects related to Group were observed, we verified that the two groups paid similar attention to the videos (i.e., 97% of the screen time. See Supplemental Digital Content 3 <http://links.lww.com/EANDH/A743>). Age was observed to have an effect: An increase in Age was associated with an increase in the fixation ratio within the video frame, $b = 0.04$, $p = 0.001$. Last, an effect for Valence was found, $b = -0.46$, $p = 0.003$. Videos that featured positive emotions were looked at for a longer time than videos that featured negative emotions. The interaction of Group \times Valence was not observed.

Second, we modelled the fixation ratio within the AOIs. Fixed effects included Age, Group, Valence, AOI (4: Target Head, Target Body, Partner Head, Partner Action), and interactions with Group. We observed an effect for Group, which indicated a group difference in the reference category, i.e., the Target Head: Compared to the TH children, the DHH children spent less time looking at a Target Head, $b = -0.025$, $p = 0.030$ (Table 2 and Fig. 2A). We also found interactions for Group \times Target Body and for Group \times Partner Head. Compared to the TH children, the DHH children spent more time looking at a Target Body, $b = 0.04$, $p = 0.006$, and at Partner Head, $b = 0.029$, $p = 0.048$. The two groups generally showed the same pattern: participants looked at the Partner Head for a longer time than the Target Head, and participants looked at the Target Head for a longer time than the Target Body and Partner Action (Fig. 2A). An effect for Age was observed: An increase in Age was associated with an increase in the fixation ratio within the AOIs

in the videos, $b = 0.001$, $P < 0.001$. There was no effect for Valence. See Supplemental Digital Content 3 <http://links.lww.com/EANDH/A743> for more details about the exact duration children spent looking at the screen versus off the screen and at each AOI.

INTERPRETATION

A model was developed for nonverbal and verbal interpretation. Fixed effects included Age, Group, Valence, Task (2: nonverbal and verbal), Group \times Valence, Group \times Task, and Group \times Valence \times Task. An effect for Group was observed: The DHH children scored lower than the TH children in both nonverbal and verbal tasks, $b = -0.13$, $p = 0.005$ (Table 2 and Fig. 2B). An effect for Task indicated that both groups of children performed better on nonverbal interpretation than on verbal interpretation, $b = 0.08$, $p = 0.005$. Last, there was an effect for Age: An increase in Age was associated with an increase in interpretation scores, $b = 0.01$, $P < 0.001$. No other effects were observed.

EFFECT OF ENCODING ON INTERPRETATION

In the exploratory analysis, we developed two models respectively for nonverbal and verbal interpretation. Fixed effects included Age, Group, Valence, and looking times at Target Head, Target Body, Partner Head, and Partner Action, as well as their interactions with Group.

For nonverbal interpretation, we observed an interaction for Target Head \times Group, $b = 0.61$, $p = 0.029$ (Table 2). The increase in nonverbal scores associated with longer looking times for the Target Head was greater in the DHH children than in the TH children (Fig. 2C). It was also observed that lower nonverbal scores were associated with longer looking times for the Target Body, $b = -0.32$, $p = 0.029$, and for the Partner Head, $b = -0.27$, $p = 0.033$, in both groups.

For verbal interpretation, we observed an interaction for Target Body \times Group, $b = -0.54$, $p = 0.037$ (Table 2). Longer looking times at the Target Body were associated with a decrease in verbal scores in the DHH children, but we found no effect for the TH children (Fig. 2D). Moreover, in both groups, the verbal scores increased with longer looking times at the Target Head, $b = 0.37$, $p = 0.006$, but with shorter looking times at the Partner Action, $b = -0.35$, $p = 0.018$.

DISCUSSION

In this study, we examined the initial two steps of the SIP model, emotion encoding and emotion interpretation. We investigated how DHH and TH children, respectively, understood nonverbal emotional cues in dynamic social situations, where an interaction partner elicited an emotion in a target person. The results showed characteristic patterns at both SIP steps in the DHH children. The DHH and TH children both paid more attention to the heads of the target person and of the partner than to their bodies or actions when encoding emotional cues. However, the DHH children spent less time looking at the target person's head and more time looking at the target person's body and at the partner's head than the TH children. When interpreting emotions, the DHH children scored lower than their TH peers and their lower scores were associated with their distinctive encoding pattern of spending less time looking at the target

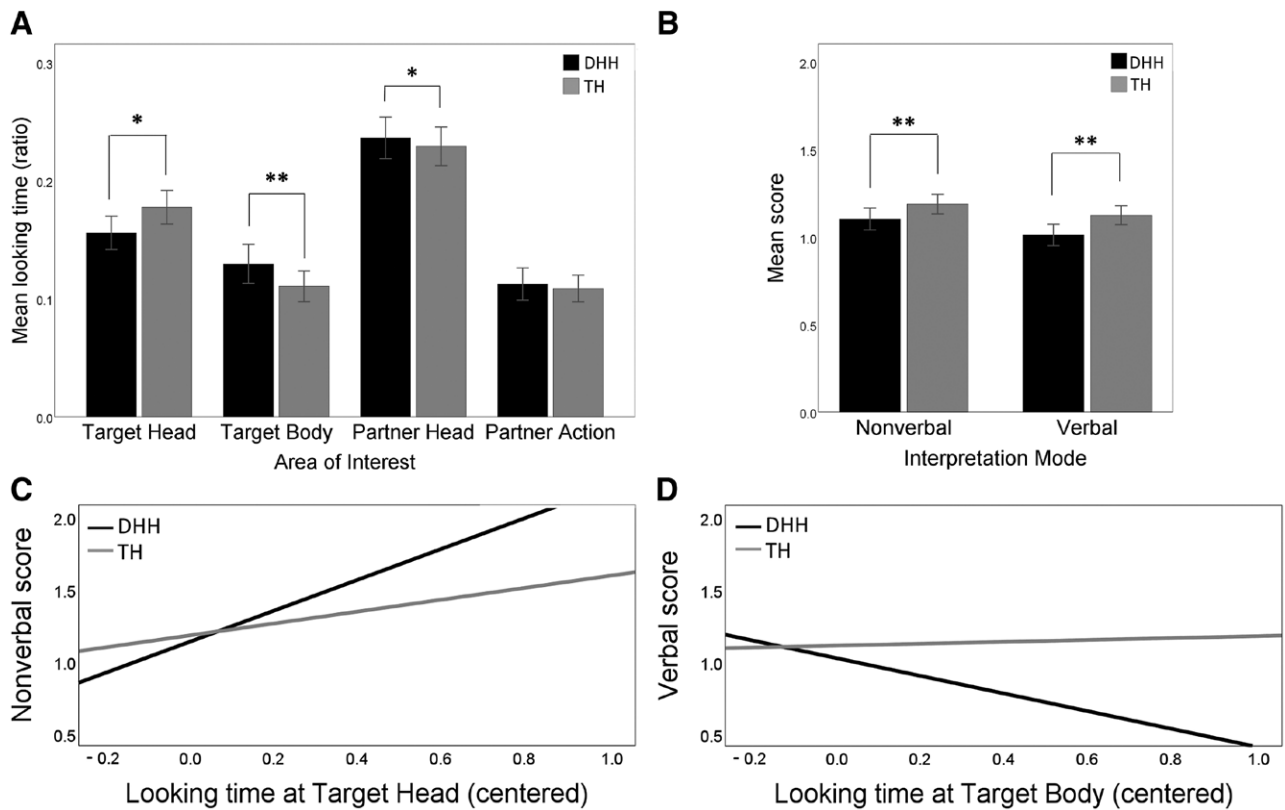


Fig. 2. Group differences during the emotion understanding task. (A) DHH children (in black) looked for a shorter time at the target person's head, and for a longer time at target person's body and trigger person's head, than TH children. (B) DHH children were less accurate (on a scale from 0 to 2) than TH children when nonverbally and verbally interpreted the emotion in the situations. (C) A larger increase in nonverbal interpretation scores with longer looking times at target person's head was observed in DHH children than in TH children. (D) A decrease in verbal interpretation scores with longer looking times at target person's body was observed in DHH children, but not in TH children. DHH children were represented with black bars/lines; TH children were represented with gray bars/lines. The error bars indicate 95% confidence interval. * $P < 0.05$; ** $P < 0.01$. DHH indicates deaf or hard-of-hearing; TH, typically hearing.

person's head and more time looking at the target person's body. With increased age, children attended to the relevant emotional cues preselected by this study longer and interpreted situations with the emotion intended by the study more often. The implications of these outcomes will be discussed below.

Outcomes showed that the DHH children diverted their attention from the target person's head to the target person's body and partner's head, whereas their TH peers exhibited a clear focus on the heads of the two protagonists. Note that the facial information of the target person was not visible in our videos, as the target person was not facing the camera. Despite the missing facial expressions, the TH children collected information mainly from the heads. This tendency is congruent with past studies that indicated a preference among typically developing individuals for attending to heads over other body regions (End & Gamer 2017; Nelson & Mondloch 2018).

Although the DHH children also exhibited a preference for looking at heads, the head region seems to have been less informative to them than to the TH children, when the facial information was missing. It is worth noting that in real life, the head region carries not only visual cues but auditory cues. While TH individuals do not need specific facial information such as lip movements to fully understand speech in real-time, this may be the case for DHH individuals (Letourneau & Mitchell 2011; Tye-Murray et al. 2014; Wang et al. 2017; Schreitmüller et al. 2018; Worster et al. 2018). This stronger dependence on facial

information during daily communication may explain why the DHH children found the head region less informative when facial cues were missing. We speculate that the DHH children reduced their attention to the target person's head and increased their attention to other cues that could give more information about the situation, including the target person's body (which informs emotions, as in moving forward in anger or backward in fear, shows where the emotion is directed to, and reveals physical conditions such as falling) and the interaction partner's head (i.e., the emotion expressed by the partner, whose face was shown). Taken together, these outcomes suggest that DHH children try to make use of more explicit, visually observable cues to compensate for the ambiguous information (Kret & De Gelder, 2013; Kret et al. 2017). As the saying goes, "Let the evidence speak for itself." DHH children may use this visual cue-based strategy to minimize misinterpretations or fill in the blanks during their daily social interactions and observations, given that they lack full auditory access to the social world around them that their TH peers have (Rieffe & Terwogt 2000; Rieffe et al. 2003).

In line with the SIP framework, the DHH children's encoding pattern of diverting attention from ambiguous to explicit cues was associated with their interpretation of emotions in social situations. Indeed, a target person's body can provide essential information when the situation results in a clear physical condition such as falling down. However, it could also be misleading

when more ambiguous physical outcomes are to be observed, such as being laughed at. Without adequate social-emotional knowledge, an encoding pattern that depends on explicit cues only may lead to misinterpretations. We did observe small effect sizes for group differences when analyzing emotion encoding and emotion interpreting separately. However, large effect sizes were observed when we analyzed the effect of encoding on interpretation. This suggests that an encoding strategy that is not supported by adequate social-emotional knowledge could potentially lead to serious misinterpretations. Such findings may carry important rehabilitative implications. As the SIP model proposes that children constantly refer to their past experiences when processing social information at different steps (Crick & Dodge 1994), providing children with an accessible social environment to allow for easier acquisition of social-emotional knowledge may be essential for supporting DHH children's emotion understanding.

LIMITATIONS AND FUTURE DIRECTIONS

Through the use of eye tracking technology, this study is among the first to show that DHH and TH children encoded nonverbal emotional cues differently in a dynamic social situation, and that furthermore, these patterns were linked to their interpretations of these situations. However, some limitations must be considered.

First, in this study we included only basic emotions and prototypical situations, and recruited children three to ten years old. Also, out of consideration for the complexity of our models and for the likelihood that participants' response toward a certain situation might be affected by their past experiences (Frijda 1986; Rieffe et al. 2005), we examined the valence of emotions rather than discrete emotional categories. The relatively easy task and older age might explain the small effect sizes for the group differences observed in the encoding stage and in the interpretation stage. However, it should be noted that difficulties in emotion understanding could be even more evident when complex emotional categories and novel situations are involved.

Second, in the present study we did not include any auditory signals in our video stimuli. However, in daily life people usually process emotional information through multiple channels. According to previous studies that examined how DHH children looked at the eyes, nose, and mouth when looking at faces, DHH children only increased their attention to the mouth region when presented with a face along with auditory linguistic information (Wang et al. 2017; Worster et al. 2018). This indicates that DHH children could encode cues differently in the presence of auditory information. Further investigations are needed to establish whether such an adaptation can be observed when DHH children encode emotions in dynamic social situations that include auditory information.

Third, the heterogeneity of the DHH population is an aspect that we did not tackle. Instead, we included a group of DHH children where the majority had profound hearing loss and used cochlear implants. All had TH parents and used spoken language. Our results showed that hearing loss could affect emotion understanding, despite partially restored hearing and spoken language mode. Yet, future studies are needed to confirm whether this pattern also emerges in DHH individuals with other hearing, family, or language backgrounds. Moreover, given that the SIP model centers around children's past experiences and

social-emotional knowledge, examining the SIP patterns in DHH children with different backgrounds would provide further insight into the model. This may include DHH children in their early years before amplification, and DHH children with DHH parents, who may receive more nonverbal language input from their parents than DHH children with TH parents (e.g., Loots et al. 2005).

Last, we suggest that future studies take into account DHH children's visual cue-based compensatory mechanism when interpreting the social interaction patterns of these children. On the one hand, DHH children may use this strategy to maximize their understanding of what is happening in a situation and to facilitate their communication with others. On the other hand, DHH children may easily misunderstand a situation when they do not have adequate social-emotional knowledge to guide their visual processing. Moreover, such a mechanism may underlie not only the interpretation of emotions, but other areas that require social-emotional knowledge. For example, DHH children were reported to show fewer prosocial behaviors (e.g., helping and sharing) than their TH peers (Netten et al. 2015; Eichengreen et al. 2019). This may suggest that DHH children face difficulties approaching other people in social situations or that they do so in a different manner, which could further affect their peer relationships (Rieffe et al. 2018). More studies are needed to understand the relation between visual processing and how DHH children actually approach social situations.

CONCLUSIONS

This study showed that DHH children diverted their attention away from ambiguous information to more explicit, visually observable cues when processing emotions in social situations. This visual cue-based tendency is likely a strategy to minimize misunderstanding in their daily communication. However, DHH children might not have adequate knowledge about causes of emotions and social rules to support such a strategy. This, in turn, may lead to misinterpretation of emotions in social situations. Our findings underscore the need to provide extra support to DHH children. Such support could include more explicit discussion about and instruction on mental states, and providing an environment where these children can more easily follow what is happening around them, such as one that features the use of an FM system and acoustic paneling in classrooms. Moreover, professionals and parents may need to consider the possibility that children with hearing loss may use unique compensatory visual encoding patterns for understanding social situations. By taking the information processing patterns of these children into account, more appropriate support may be provided to these children.

ACKNOWLEDGMENTS

The authors would like to thank all the children and their parents who participated in this study. We would also like to express our genuine appreciation to Dr. C.M. Wu, Ms. C.J. Lin, and Ms. Y.Y. Lin at Chang-Gung Memorial Hospital for their tremendous help in recruiting participants, and to Ms. E. Chen from Tobii Pro and SOLO at Leiden University for their technical support. Last but not least, we would like to thank Ms. J. Schoerke for her suggestions and corrections on the language in this manuscript.

Y.T.T. applied for funding, designed and performed experiments, analyzed data, and wrote the article with input from all authors; B.L. designed and planned experiments, applied for funding, and provided critical revision;

M.E.K. designed experiments, supervised statistical analysis, and provided critical revision; J.H.M.F. conceived the study, contributed to the interpretation of the results, and provided critical revision; C.R. conceived the study, applied for funding, provided critical revision, and was in charge of overall direction and planning.

This work was supported by The Royal Netherlands Academy of Arts and Sciences under Grant 530-5CDP17; the Taiwan Ministry of Education under Grant GSSA1071007013. No potential conflict of interest was reported by the authors.

Following the policy of the Unit of Developmental and Educational Psychology, Leiden University, the dataset and associated information used in the present study will be shared publicly on the archiving platform DataverseNL (<https://dataverse.nl/>) once the manuscript is accepted.

Address for correspondence: Yung-Ting Tsou, Unit of Developmental and Educational Psychology, Institute of Psychology, Leiden University, Pieter de la Court Building, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands. E-mail: ytsou@fsw.leidenuniv.nl.

Received May 11, 2020; accepted October 28, 2020; published online ahead of print December 22, 2020.

REFERENCES

- Ahtola, E., Stjerna, S., Yrttiäho, S., Nelson, C. A., Leppänen, J. M., Vanhatalo, S. (2014). Dynamic eye tracking based metrics for infant gaze patterns in the face-distractor competition paradigm. *PLoS One*, 9, e97299.
- Birmingham, E., Bischof, W. F., Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Res*, 49, 2992–3000.
- Calderon, R., & Greenberg, M. T. (2011). Social and Emotional Development of Deaf Children: Family, School, and Program Effects. In M. Marschark & P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (2nd ed., Vol. 1, pp. 188–199). New York, NY: Oxford University Press.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychol Bull*, 115, 74–101.
- Dael, N., Mortillaro, M., Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, 12, 1085–1101.
- de Castro, B. O., Merk, W., Koops, W., Veerman, J. W., Bosch, J. D. (2005). Emotions in social information processing and their relations with reactive and proactive aggression in referred aggressive boys. *J Clin Child Adolesc Psychol*, 34, 105–116.
- Deneault, J., & Ricard, M. (2013). Are emotion and mind understanding differently linked to young children's social adjustment? Relationships between behavioral consequences of emotions, false belief, and SCBE. *J Genet Psychol*, 174, 88–116.
- Dirks, E., Stevens, A., Kok, S., Frijns, J., Rieffe, C. (2020). Talk with me! Parental linguistic input to toddlers with moderate hearing loss. *J Child Lang*, 47, 186–204.
- Durand, K., Galloway, M., Seigneuric, A., Robichon, F., Baudouin, J. Y. (2007). The development of facial emotion recognition: The role of configural information. *J Exp Child Psychol*, 97, 14–27.
- Eichengreen, A., Broekhof, E., Güroğlu, B., Rieffe, C. (2019). Fairness decisions in children and early adolescents with and without hearing loss. *Soc Dev*, 0, 1–15.
- Embregts, P., & van Nieuwenhuijzen, M. (2009). Social information processing in boys with autistic spectrum disorder and mild to borderline intellectual disabilities. *J Intellect Disabil Res*, 53, 922–931.
- End, A., & Gamer, M. (2017). Preferential processing of social features and their interplay with physical saliency in complex naturalistic scenes. *Front Psychol*, 8, 418.
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *PNAS*, 102, 17245–17250.
- Frank, M. C., Amso, D., Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *J Exp Child Psychol*, 118, 13–26.
- Frijda, N. H. (1986). *The emotions*. Paris, France: Editions de la Maison des Sciences de l'Homme.
- Gray, C., Hosie, J., Russell, P., Scott, C., Hunter, N. (2007). Attribution of emotions to story characters by severely and profoundly deaf children. *J Dev Phys Disabil*, 19, 145–159.
- Hess, U., Blaison, C., Kafetsios, K. (2015). Judging facial emotion expressions in context: The influence of culture and self-construal orientation. *J Nonverbal Behav*, 40, 55–64.
- Horsley, T. A., de Castro, B. O., Van der Schoot, M. (2010). In the eye of the beholder: eye-tracking assessment of social information processing in aggressive behavior. *J Abnorm Child Psychol*, 38, 587–599.
- Kolb, B., Wilson, B., Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain Cogn*, 20, 74–84.
- Klein, M. R., Moran, L., Cortes, R., Zalewski, M., Ruberry, E. J., Lengua, L. J. (2018). Temperament, mothers' reactions to children's emotional experiences, and emotion understanding predicting adjustment in pre-school children. *Soc Dev*, 27, 351–365.
- Kret, M. E., & de Gelder, B. (2013). When a smile becomes a fist: the perception of facial and bodily expressions of emotion in violent offenders. *Exp Brain Res*, 228, 399–410.
- Kret, M. E., Stekelenburg, J. J., de Gelder, B., Roelofs, K. (2017). From face to hand: Attentional bias towards expressive hands in social anxiety. *Biol Psychol*, 122, 42–50.
- Leibold, L. J., Hillock-Dunn, A., Duncan, N., Roush, P. A., Buss, E. (2013). Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear Hear*, 34, 575–584.
- Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Dev*, 71, 107–118.
- Letourneau, S. M., & Mitchell, T. V. (2011). Gaze patterns during identity and emotion judgments in hearing adults and deaf users of American Sign Language. *Perception*, 40, 563–575.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc*, 83, 1198–1202.
- Loots, G., Devisé, I., Jacquet, W. (2005). The impact of visual communication on the intersubjective development of early parent-child interaction with 18- to 24-month-old deaf toddlers. *J Deaf Stud Deaf Educ*, 10, 357–375.
- Main, J. C., DeBruine, L. M., Little, A. C., Jones, B. C. (2010). Interactions among the effects of head orientation, emotional expression, and physical attractiveness on face preferences. *Perception*, 39, 62–71.
- Morgan, G., Meristo, M., Mann, W., Hjelmquist, E., Surian, L., Siegal, M. (2014). Mental state language and quality of conversational experience in deaf and hearing children. *Cogn Dev*, 29, 41–49.
- Nelson, N. L., & Mondloch, C. J. (2018). Children's visual attention to emotional expressions varies with stimulus movement. *J Exp Child Psychol*, 172, 13–24.
- Netten, A. P., Rieffe, C., Theunissen, S. C., Soede, W., Dirks, E., Briaire, J. J., Frijns, J. H. (2015). Low empathy in deaf and hard of hearing (pre)adolescents compared to normal hearing controls. *PLoS One*, 10, e0124102.
- Proksch, J., & Bavelier, D. (2002). Changes in the spatial distribution of visual attention after early deafness. *J Cogn Neurosci*, 14, 687–701.
- Pyykkö, J., Ashorn, P., Ashorn, U., Niehaus, D. J. H., Leppänen, J. M. (2019). Cross-cultural analysis of attention disengagement times supports the dissociation of faces and patterns in the infant brain. *Sci Rep*, 9, 14414.
- Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychol Methods*, 5, 199–213.
- Reynolds, G. D., & Roth, K. C. (2018). The development of attentional biases for faces in infancy: a developmental systems perspective. *Front Psychol*, 9, 222.
- Rieffe, C., Netten, A. P., Broekhof, E., Veiga, E. (2015). The role of the environment in children's emotion socialization. In H. Knoors & M. Marschark (Eds.), *Educating deaf learners: Creating a global evidence base* (pp. 369–388). New York, NY: Oxford University Press.
- Rieffe, C., Broekhof, E., Eichengreen, A., Kouwenberg, M., Veiga, G., da Silva, B. M. S., van der Laan, A., Frijns, J. H. M. (2018). Friendship and emotion control in pre-adolescents with or without hearing loss. *J Deaf Stud Deaf Educ*, 23, 209–218.
- Rieffe, C., & Terwogt, M. M. (2000). Deaf children's understanding of emotions: desires take precedence. *J Child Psychol Psychiatry*, 41, 601–608.
- Rieffe, C., Terwogt, M. M., Cowan, R. (2005). Children's understanding of mental states as causes of emotions. *Infant Child Dev*, 14, 259–272.

- Rieffe, C., Terwogt, M. M., Smit, C. (2003). Deaf children on the causes of emotions. *Educ Psychol*, 23, 159–168.
- Rollman, S. A., & Harrison, R. D. (1996). A comparison of deaf and hearing subjects in visual nonverbal sensitivity and information processing. *Am Ann Deaf*, 141, 37–41.
- Saarni, C. (1999). *The development of emotional competence*. New York, NY: Guilford Press.
- Schreitmüller, S., Frenken, M., Bentz, L., Ortmann, M., Walger, M., Meister, H. (2018). Validating a method to assess lipreading, audiovisual gain, and integration during Speech Reception With Cochlear-Implanted and Normal-Hearing Subjects Using a Talking Head. *Ear Hear*, 39, 503–516.
- Sladen, D. P., Tharpe, A. M., Ashmead, D. H., Grantham, D. W., Chun, M. M. (2005). Visual attention in deaf and normal hearing adults. *J Speech Lang Hear Res*, 48, 1529–1537.
- Slater, A., & Quinn, P. C. (2001). Face recognition in the newborn infant. *Infant Child Dev*, 10, 21–24.
- Torres, J., Saldaña, D., Rodríguez-Ortiz, I. R. (2016). Social information processing in deaf adolescents. *J Deaf Stud Deaf Educ*, 21, 326–338.
- Troop-Gordon, W., Gordon, R. D., Vogel-Ciernia, L., Ewing Lee, E., Visconti, K. J. (2018). Visual attention to dynamic scenes of ambiguous provocation and children's aggressive behavior. *J Clin Child Adolesc Psychol*, 47, 925–940.
- Tye-Murray, N., Hale, S., Spehar, B., Myerson, J., Sommers, M. S. (2014). Lipreading in school-age children: the roles of age, hearing status, and cognitive ability. *J Speech Lang Hear Res*, 57, 556–565.
- Wang, Y., Zhou, W., Cheng, Y., Bian, X. (2017). Gaze Patterns in Auditory-Visual Perception of Emotion by Children with Hearing Aids and Hearing Children. *Front Psychol*, 8, 2281.
- Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wiefferink, C. H., Rieffe, C., Ketelaar, L., De Raeve, L., Frijns, J. H. (2013). Emotion understanding in deaf children with a cochlear implant. *J Deaf Stud Deaf Educ*, 18, 175–186.
- Worster, E., Pimperton, H., Ralph-Lewis, A., Monroy, L., Hulme, C., MacSweeney, M. (2018). Eye movements during visual speech perception in deaf and hearing children. *Lang Learn*, 68(Suppl Suppl 1), 159–179.