A set of tasks to test interpretation and use of emoji by children

Jonna de Kruijff
Graduation Thesis
Media Technology MSc program, Leiden University
January 2019
Thesis advisors: Dr. Max van Duijn, Prof.dr. Carolien Rieffe

Abstract

In light of the difficulties children with an autism spectrum disorder (ASD) face when reading facial expressions, as well as their tendency to prefer online interaction to face-to-face communication, a study was planned to investigate whether children with ASD interpret and use emoji differently from typically developing children. Three tasks were developed to test how children group, interpret, and apply emoji. Due to setbacks in recruiting a test group in the limited time frame, an alternative criterion for division into two test groups was selected post hoc. Participants were divided according to their self-reported frequency of emoji use. Various differences were found and will be discussed, along with recommendations for future continuation of this research with participants with ASD.

Introduction

Autism Spectrum Disorders (ASD) form a group of developmental conditions associated with trouble with social skills. People with ASD generally find expressing their own emotions difficult, as well as interpreting the emotions displayed by others (Baron-Cohen, 1993). Subtleties like tone of voice and sarcasm may be lost on people with ASD. They also have been shown to have trouble interpreting facial expressions from illustrations (Celani, 1999).

As the label suggest, people with ASD fall on a spectrum. When someone with ASD has an IQ of 70 or greater, they are said to have high-functioning (HF) autism (Carpenter, 2009). Recent studies have shown that children with HF ASD are equally good at recognising the basic emotions of happiness, sadness, anger, surprise, fear, and disgust in photographs as typically-developing children (Castelli, 2005). These studies were conducted with photographs designed for testing recognition of basic emotions. There are currently no studies investigating if children with HF ASD are able to recognise more complex emotions from photographs, and if they can recognise emotions from different pictures, like icons.

People with ASD tend to be more comfortable in online communications (Benford, 2009). Digital devices such as tablets have also been shown to help people with ASD in social interactions (Hourcade, 2013). Software has been developed with the goal of training individuals with ASD on how to behave in social situations. These programs showed good results, with the best results being attained by games that involved a roleplay component that mirrored natural social interaction (Mitchell, 2007). With recent advances in technology, research has also been done to show how new technologies can be used in interventions for people with ASD. An example of this was the use of virtual reality environments designed to train children with ASD in interactions (Ke, 2013).

Today, digital communication is ubiquitous. Whenever we want to send a quick message to

anyone we do so through text messaging. However, text messaging lacks the interactive depth that face-to-face communication offers with added dimensions like body language. People without social developmental disorders are able to infer whether a text message has a positive or negative emotional valence. Communication partners used strategies like verbosity and punctuation to express negative affect, which was understood by both partners (Hancock, 2007). Aside from these strategies, emoji can also be used to help the reader understand the emotion behind the message. Emoji are a collection of icons representing a simplified human face with exaggerated emotional display. These pictures are often used in online communications to help convey emotion with text, something that is harder to accomplish without images (Garrison, 2011).

The difficulty with interpretation of emotions from facial expressions found in people with ASD may have some implications for online communications. This paper will investigate whether children with HF ASD have trouble interpreting emoji, in the same way they have trouble interpreting emotions from regular facial expressions.

Research questions

Originally this paper aimed at answering 2 research questions: in comparison with typically developed peers, to what degree are children with HF ASD

(1) able to interpret the use of emoji?; and

(2) able to pick a fitting emoji when given a sample conversation?

The idea was to recruit an equal number of participants at primary schools with and without ASD. However, only 5 participants with an ASD diagnosis were found within the time frame of this study, versus 44 participants without ASD. This ratio made the between-group comparison necessary for answering research questions (1) and (2) virtually impossible.

With this in mind, it was decided to add a third research question exhibiting a better fit with the available data:

> (3) Does frequency of emoji-use have an effect on emoji recognition and interpretation?

A difference in emoji recognition and interpretation between frequent and infrequent emoji users could show whether emoji use relies rather on learned versus innate mechanisms. If infrequent users interpret emoji very differently from frequent users, this would indicate a tendency towards learning; after all, it shows that infrequent users haven't used emoji enough to have acquired a consensus on their meaning. Conversely, if no difference is found it would point towards emoji interpretation relying more on innate mechanisms.

Methods and test materials

This study was approved by the ethics board of the Faculty of Social Sciences at Leiden University in the Netherlands on the 2nd of June 2017. Participants were recruited from elementary schools in the western part of the Netherlands. Recruitment took place from September through November 2018. In total, 49 participants agreed to participate, 5 of which had an ASD diagnosis. The participants' ages ranged from 10 to 15, with most participants being 12 years old. 45% of the participants were boys, 55% were girls.

The tasks that are described below were developed by the researcher in Qualtrics. This

allowed for a test that could be completed on various devices, as well as an easy way to manage responses. Responses were coded to separate the groups based on ASD diagnosis. Data and consent forms will be saved at the faculty of social sciences at Leiden University for 5 years after the study.

The tasks consisted of two main parts. The first part asked participants for their age, gender, and their experience with emoji. The second part contained the main tasks. The main tasks were a grouping task, an interpretation task, and an application task. Each part will be described in detail below.

In the first part, participants were asked some questions about their own use of emoji. We asked participants how often they used emoji, how clear they thought others' emoji use is, and whether they thought others understood their use of emoji. Participants rated their answers on a 5-point scale (see figure 1).

Figure 1. Example question. Participants were asked how often they use emoji and had to rate their answer on a 5-point scale, with the low end being labelled "never" and the top end "very often."

Next, participants did three different tasks: a grouping task, an interpretation task, and an application task. Originally, the first two tasks were designed to test research question (1), the third task to test research question (2). With each question, the differences between groups were analysed.

For the grouping task, participants were shown a word for one of four basic emotions, "happy," "angry," "sad," or "surprised." They then had to choose as many emoji as they wanted from a group of 10 that they thought fit the word. See figure 2 for an sample of these questions.

We will focus on differences between groups, rather than trying to decide on the 'correct' use of emoji a priori and using this as a benchmark for measuring participants' 'performance'. In other words, our view is that there is never a 'right' answer to the task, but it should be used to measure differences in average answers between groups.

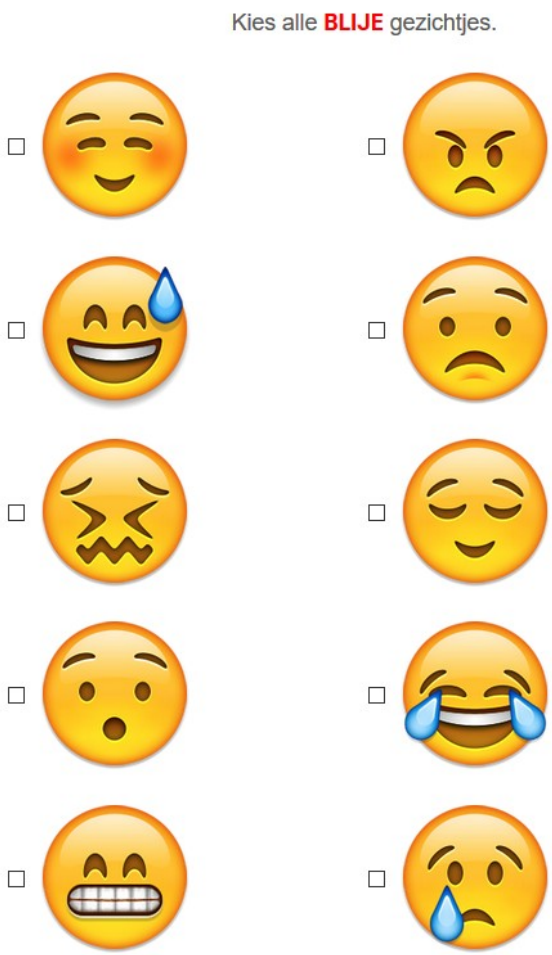

Figure 2. Sample of the grouping task. Participants are asked to pick all happy faces from a set of 10. Some emoji clearly depict happy or sad faces, while others are more ambiguous.

In the interpretation task participants were shown an emoji (target emoji). They were asked to match 1 of 4 emoji representing the basic emotions of happiness, sadness, anger, and surprise to the target emoji. Then they were also asked which word they thought best described the test emoji. This is to control for the participant selecting a word not describing the chosen emoji, thus possibly misunderstanding the chosen emoji in the first place. See figure 3 for a sample of this task.

Figure 3. Sample of the interpretation task. The emoji matching and word matching questions were presented on separate pages. They are shown in one image for clarity. The word matching task asks participants which word they think matches the target emoji, with options for sad, happy, surprised, and angry.

The application task involved text conversations written by the researcher. In each text conversation a space was marked. Participants were asked to fill in this space with one of six provided emoji. Each set of 6 emoji contained 2 with a positive emotional valence, 2 with a negative emotional valence, and 2 ambiguous emoji. The text conversations were about subjects a participant could encounter in daily life, like losing a pencil case. The conversations were written in a way that multiple emoji answers could be appropriate. See figure 4 for a sample of this task.

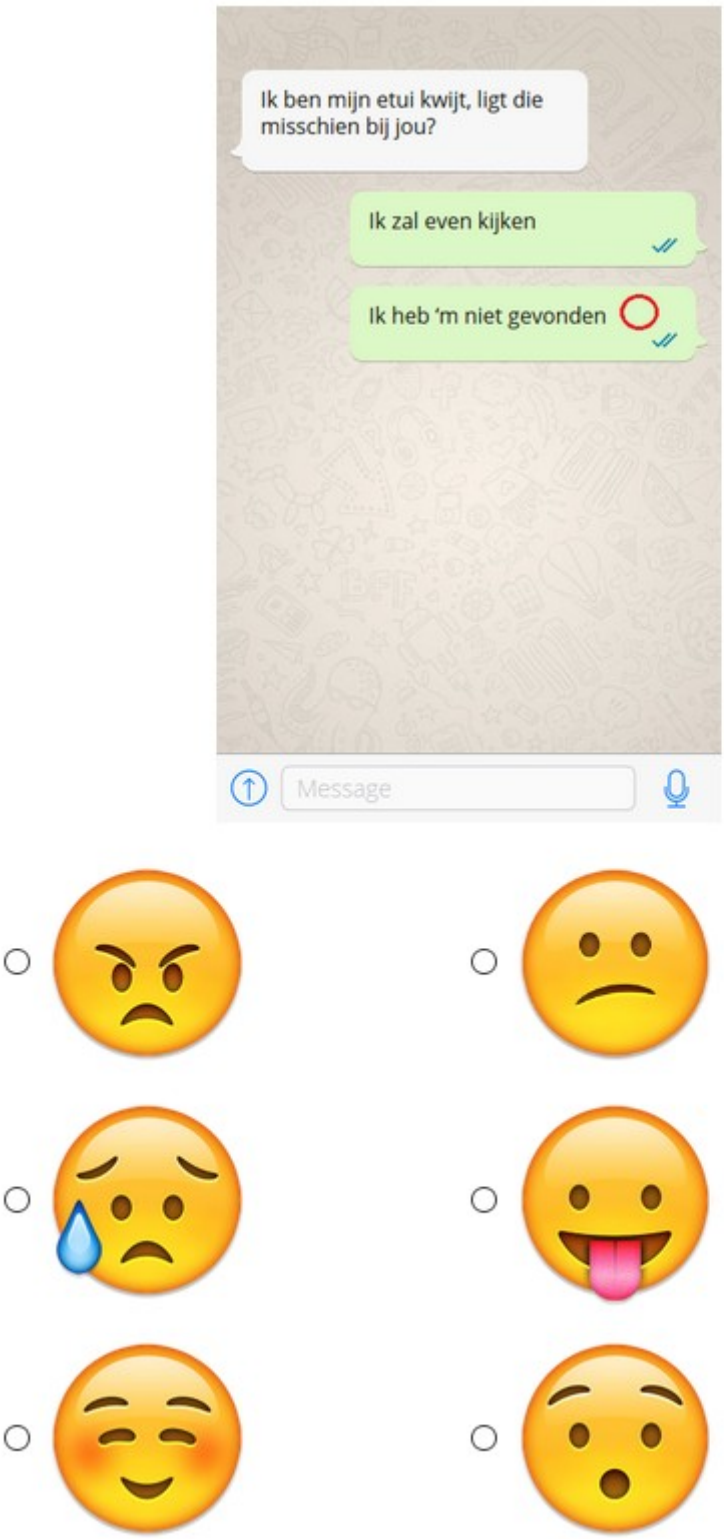

Figure 4. Sample of the application task. The text conversation reads: "I've lost my pencil case, is it at your place?" "I'll have a look." "I haven't found it." The participant was asked to choose which emoji they would put in the red circle.

To determine which emoji would be investigated the tasks were first given to a group of 8

adults. This group was recruited among friends and family of the researcher. Based on their answers emoji would be labelled ‘ambiguous’ or ‘clear.’ For the grouping task emoji were labelled ‘ambiguous’ if an emoji was not chosen for an emotion unanimously. For the application task, emoji were labelled ‘ambiguous’ when answers were not chosen unanimously. The emoji that were labelled ambiguous can be found in appendix I.

Due to the low amount of children with an ASD diagnosis we were able to recruit, research questions (1) and (2) will remain unanswered for now. The tasks described above were made with these research questions in mind and can be used for further research.

Research question (3) concerned differences in emoji use and interpretation between frequent users and infrequent users. Frequent users were defined as answering with a 3 or 4 on the question “How often do you use emoji?” None of the participants rated their emoji use with a 5. This yields a division into two groups: 17 infrequent users and 32 frequent users. This group was formed using data from the 3 questions about participants’ experience with emoji that were asked before the tasks, which can be found in table 1. Data for all questions is shown, the test groups were formed based on data in table 1A.

a

		Frequency	Percent
Valid	.00	1	2.0
	1.00	6	12.2
	2.00	10	20.4
	3.00	22	44.9
	4.00	10	20.4
	Total	49	100.0

b

		Frequency	Percent
Valid	1.00	5	10.2
	2.00	13	26.5
	3.00	22	44.9
	4.00	9	18.4
	Total	49	100.0

c

		Frequency	Percent
Valid	1.00	1	2.0
	2.00	12	24.5
	3.00	23	46.9
	4.00	13	26.5
	Total	49	100.0

Table 1. Frequency tables for the questions about participants’ emoji use. A: “How often do you use emoji?” B: “Do you understand the emoji other people send to you?” C: “Do other people understand the emoji you send?” 0 indicates “Never,” while 4 indicates “Very often.”

For analysis we decided to look at how the groups interpreted the more ambiguous emoji and text conversations differently. For the selected ambiguous emoji, Pearson’s chi-squared tests were performed to see if the frequent users interpreted and used these emoji differently from infrequent users ($P<0.05$). Due to the nature of the grouping task, participants could select multiple emoji for each emotion, analysis was done for separate ambiguous emoji in each emotion-group.

Results

Results will be elaborated upon when the emoji the experimental group chooses are significantly different ($p <= 0.05$) from the one chosen by the control group. For the grouping task, the groups showed no difference in assigning any of the ambiguous emoji as ‘happy’ or ‘surprised’. When asked to label emoji as ‘angry’, the infrequent users grouped the emoji labeled ‘brr’ (this emoji can be seen in figure 5A) under ‘angry’ slightly more often than the frequent users, but this difference was not significant. Assigning emoji as ‘sad’ showed the largest differences between the groups. The emoji labeled ‘brr’ ($P=0.022$) and ‘weary’ ($P=0.027$)(the leftmost and middle emoji in figure 5B) was grouped under ‘sad’ more often by the frequent users than the infrequent users. The emoji labeled ‘meh’ ($P=0.069$) (the rightmost emoji in figure 5B) was grouped under ‘sad’ somewhat more often by the frequent users, but this difference was not significant. An overview of the analysed emoji can be

found in table 2.

GROUPING TASK		
category	label	p
blij	tanden	0.959
boos	*brr*	*0.077*
	tanden	0.128
verdrietig	tranen	0.293
	brr	**0.022**
	weary	**0.027**
	disappoint	0.166
	meh	*0.069*
verrast	disappoint	0.185
	smile	0.461

Table 2. P-values for the analysed emoji in the grouping task. Significant differences are bolded, differences that are approaching significant (P<0.10) are italicised.

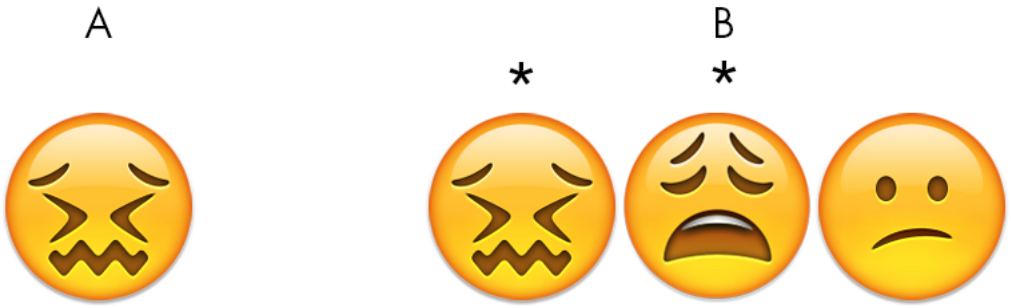

*Figure 5. A: The emoji that showed a difference that approached significance in how the groups grouped it as 'angry'. B: The emoji that showed a difference that approached significance and were significant (marked with a *) in how the groups grouped them as 'sad'.*

No significant differences between the two groups were found in the ambiguous emoji in the interpretation task. P-values of the analysed ambiguous emoji can be found in table 3.

INTERPRETING TASK	
label	p
tanden	0.268
tandenogen	0.721
druppelsad	0.997
disappoint	0.76
brr	0.486
oeps	0.798

Table 3. P-values for the analysed ambiguous emoji in the interpretation task.

Most of the conversations in the application task were not assigned different emoji by the two groups. There was one exception: the conversation about someone having a lift to school (figure 6) was given significantly different emoji by the two groups (P=0.009). Infrequent users' choices were more spread out over the available emoji, while 43.75% of

frequent users chose the same emoji. P-values of every conversation can be found in table 4.

Figure 6. The conversation where significantly different emoji were assigned by the two groups. The texts read: "Want to bike to school together?" "I'd like to, but my dad is already giving me a ride."

APPLICATION TASK	
label	p
beterschap	0.996
etui	0.23
tvkijken	0.167
zomervakantie	0.839
gebracht	**0.009**
allergisch	0.216
youtube	0.501
opruimen	0.922
liegen	0.361

Table 4. P-values for the various conversations participants were asked to apply emoji to. Significant differences are bolded.

Discussion

For this study the target emoji were pre-selected based on a pilot study among friends and family since investigating all emoji in existence is not feasible. Some emoji display their emotions very clearly, reflected in highly consistent answers in the pilot; therefore we decided to omit those emoji. Especially with the more ambiguous emoji, people can give their own meaning and use them in their own ways. This is where effects are to be expected, and thus the group of emoji we focussed on.

Some significant differences between groups were found in the grouping task and the application task, no differences were found in the interpretation task. One explanation would be that the groups are indeed different in some respects relevant to precisely these questions, i.e. that frequency of emoji use drive the found effects. Alternatively, given the relatively high amount of questions where the groups did not behave differently, it could be that the found effects have other explanations and can not be attributed to frequency of use. For instance, the lack of differences found in the interpretation task could be explained by the selection of tested emoji. In future testing, this selection could be evaluated to see whether this is the case.

One conversation in the application task was given a different emoji by the two groups, as seen in figure 5 and table 2. The characters in this conversation want to bike to school together, but one party already has a ride. The emoji applied to this conversation could depend on the participant's relationship with the character that gives the ride, in this case the character's father. A difference wouldn't necessarily be expected for this task between frequent and infrequent users, since the question can be interpreted in more than one way.

The biggest differences in the grouping task were seen when participants were asked to select 'sad' emoji. Two emoji showed a significant difference, and another showed a

difference approaching significant. These emoji all had features that can be associated with sadness, like a pouting mouth, and can be seen in figure 6B. One of the emoji that was chosen significantly differently for ‘sad’ also tended to be chosen differently for ‘angry.’ All the emoji that showed some difference were ambiguous and can be used in multiple ways. Infrequent users could have seen the ambiguous emoji used less, so they may not be as familiar with how they can be used as frequent users are. It would be interesting if regular users of emoji could be compared to never-users to see how much of emoji use relies on innate mechanisms, like facial recognition (Slater, 2001).

To answer research question (3), there seems to be no overall difference in emoji use and interpretation between frequent and infrequent users. The difference found for the application task can also be due to the way the question was worded. An explanation for the differences found for the ambiguous emoji in the grouping task could be that frequent users have seen these ambiguous emoji more than infrequent users have, and will have had more opportunity to learn how they are used. Other studies have shown emoji use seems to follow conventions within groups (Roele, submitted), so differences between participants can also be explained with them being in a different emoji-user group. This could indicate that emoji use is at least in part learned, but definitive answers require more research into the frequency of use of various emoji, as well as how frequent and infrequent users use these emoji.

A weakness of the questionnaire asking for frequency of use is that it is self-reported. ‘Frequent’ and ‘infrequent’ are relative measures, and we can’t be sure what kind of environment our participants compared themselves to. The same goes for the question that asked whether participants felt others understood their emoji use and whether they understood others’ emoji use.

Conclusion

The original plan for this study was to compare children with ASD with typically developing children. While we managed to recruit enough children for the control group, recruiting children with ASD proved more difficult. Since this research project had to be wrapped up for the time being, we decided to add another research question that implied dividing the groups by a different criterium than ASD versus typically developing. Of the other metrics that were asked about the participants, we decided frequency of use would be most interesting. Future research would be focused on recruiting participants with HF ASD so a comparison can be made with the already recruited control group. If direct recruiting of children with HF ASD proves to be very difficult, the questionnaire could be expanded to include questions to measure traits associated with ASD among participants, and divide the groups using those results. This study would further use the methods already outlined in this paper.

Aside from the continuation of this study to include a test group with ASD, more research can be done on digital communication, and especially on expressing emotions in a medium that lacks face-to-face interaction. Research could focus on including new ways to simulate face-to-face contact with technologies like virtual reality and how well this replaces real-life interaction. These new technologies could be especially interesting for people with ASD, for instance as a way to train for social interactions. If a future study finds that children with ASD use emoji very differently from typically developing children, more research could be done to see what improvements could be made to enhance clarity of emoji. This would benefit online

communication as a whole, as interpreting emoji, with all the nuances of irony and personal meaning, is a steep task as it is.

References

Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of autism and developmental disorders*, *29*(1), 57-66.

Baron-Cohen, S., & Bolton, P. (1993). *Autism: the facts*. Oxford University Press.

Laura Arnstein Carpenter PhD, B. C. B. A., & Latha Soorya PhD, B. C. B. A. (2009). Asperger's syndrome and high-functioning autism. *Pediatric Annals*, *38*(1), 30.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, *9*(4), 428-449.

Benford, P., & Standen, P. (2009). The internet: a comfortable communication medium for people with Asperger syndrome (AS) and high functioning autism (HFA)?. *Journal of Assistive Technologies*, *3*(2), 44-53.

Parsons, S., Mitchell, P., & Leonard, A. (2005). Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments?. *Autism*, *9*(1), 95-117.

Hourcade, J. P., Williams, S. R., Miller, E. A., Huebner, K. E., & Liang, L. J. (2013, April). Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3197-3206). ACM.

Ke, F., & Im, T. (2013). Virtual-reality-based social interaction training for children with high-functioning autism. *The Journal of Educational Research*, *106*(6), 441-461.

Mitchell, P., Parsons, S., & Leonard, A. (2007). Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *Journal of autism and developmental disorders*, *37*(3), 589-600.

Hancock, J. T., Landrigan, C., & Silver, C. (2007, April). Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 929-932). ACM.

Garrison, A., Remley, D., Thomas, P., & Wierszewski, E. (2011). Conventional faces: Emoticons in instant messaging discourse. *Computers and Composition*, *28*(2), 112-125.

Roele, M., van Duijn, M. J., Ward, J. Submitted. Tweet with a Smile and a Unicorn: the use of emoji on Twitter in the Netherlands and England.

Slater, A., & Quinn, P. C. (2001). Face recognition in the newborn infant. *Infant and Child Development: An International Journal of Research and Practice*, *10*(1-2), 21-24.

Appendix I

Grouping task

Happy					
Angry					
Sad					
Surprised					

Interpretation task