Journal of Child Language (2024), 1–16
doi:10.1017/S0305000923000727

CAMBRIDGE UNIVERSITY PRESS

BRIEF RESEARCH REPORT

Same name, different representational levels? Misalignment of indirect parent-reported and direct alternative forced choice measures of emotion word comprehension in preschool children

Ida Torp ROEPSTORFF[1], Julien MAYOR[2], Sophie S. HAVIGHURST[2,3] and Natalia KARTUSHINA[4]

[1]Department of Psychology, Aarhus University, Bartholins Allé 11, 8000 Aarhus C, Denmark
[2]Department of Psychology, University of Oslo, Forskningsveien 3A, 0373 Oslo, Norway.
[3]Mindful: Centre for Training and Research in Developmental Health, Department of Psychiatry, University of Melbourne, Building C, 50 Flemington Street, Travancore 3032, Melbourne, Australia
[4]Centre for Multilingualism in Society across the Lifespan, Faculty of Humanities, University of Oslo, Henrik Wergelands hus, Niels Henrik Abels vei 36, 0313 Oslo, Norway
Corresponding author: Ida Torp Roepstorff; Email: ir@psy.au.dk

(Received 28 March 2022; revised 01 December 2023; accepted 17 December 2023)

Abstract

This study assessed the relationship between preschoolers' directly and indirectly assessed emotion word comprehension. Forty-nine two-to-five-year-old Norwegian children were assessed in a tablet-based 4-alternative forced choice (AFC) task on their comprehension of six basic and six complex emotions using facial expression photographs. Parents reported emotion word comprehension and production of the same words. Parent-reported emotion word production interacted with age to predict preschoolers' performance, with a parent-child alignment only observed for older children. Parent-reported word comprehension did not significantly predict accuracy. The results suggest that, in preschoolers, direct and indirect assessments might address distinct representational levels of emotion word comprehension.

Keywords: emotion word comprehension; parental reports; preschool children

Introduction

The ability to distinguish and communicate emotions is important to develop and maintain social relations, as well as good mental health (Nook et al., 2020). Despite the importance of emotion vocabulary for later emotional development (Streubel et al., 2020), there is no consensus on measurement tools in preschoolers and whether the measures are in concord.

In past research, emotion word comprehension has been examined using both indirect and direct measures. Indirect measures rely on parents to report whether their child understands (comprehension, Baron-Cohen et al., 2010; Dale & Fenson, 1996) or uses (production, Ridgeway et al., 1985) an emotion word. Direct measures of children's emotion word comprehension include Alternative Forced Choice (AFC) tasks where children match an emotion word with pre-selected facial expressions (Declercq et al., 2019; Wu et al., 2022) and definition tasks, where children describe what a given emotion word means (Nook et al., 2020). Both AFCs and parental reports may be useful tools to measure emotion word comprehension, as they are easily administered, and allow the creation of developmental age-related norms. There are, however, no studies measuring the agreement between parental reports and AFC measures of emotion word comprehension in preschoolers.

Research on general word comprehension, which is often conducted using AFC paradigms (e.g., Peabody Picture Vocabulary Test, PPVT, Dunn, 2019) and parental reports (e.g., Communicative Development Inventories, CDI, Dale & Fenson, 1996), presents mixed findings regarding the alignment between direct and indirect assessments. Libertus et al. (2015) found that parent-reported word comprehension explained the variance in two-to-six-year-old children's performance on the PPVT. However, while Sattler et al. (1985) also found a correlation between parent-reported and actual performance on the PPVT, parents overestimated their children's item-to-item performance, suggesting that parental reports align with children's relative vocabulary size, but not necessarily with the specific words that children understand. Furthermore, Lo et al. (2021) found that parental reports of infants' word comprehension were more reliable predictors of children's accuracy in a 2AFC when words were NOT presented together with semantically-related words (e.g., dog-airplane instead of dog-cat). Similarly, Arias-Trejo and Plunkett (2010) found that conceptually and perceptually similar words (e.g., dog and cat) were difficult for children to recognize, even when the parents expected the words to be understood. These findings suggest that toddlers' early word representations are not as differentiated as adult representations are: word recognition can be destabilized when the target word is encountered with a semantically similar word (Lo et al., 2021), and parents have little insight into these fine details, likely due to the lack or insufficiency of such ambiguous contexts to be encountered in real life.

Accuracy on the AFC task requires children to both understand an emotion word and map this word to a (often stereotypical, and validated by adults, see Ruba & Pollak, 2020) facial expression. Such a fine-grained differentiation of facial emotion expressions (Wu et al., 2022) is challenging for small children, as there is growing evidence that preschool children's emotion word comprehension and production is coarser and more valence-based than adults' (Widen, 2013). For instance, a majority of three-to-seven-year-old children label so-called disgusted faces "angry" (Widen & Russell, 2010b) and two-year-olds often confuse angry, sad, and scared faces (see Widen & Russell, 2008). However, Wu et al. (2022) found that two-year-olds were able to distinguish angry, sad, and scared faces (that are often confused) above chance level (see Widen & Russell, 2008), when employing a 2AFC design, making the task simpler. Thus, variations in direct AFC measurements in children may present diverging results, suggesting that emotion word comprehension in children can be measured on different representational levels.

Taken together, AFC assessment of emotion word comprehension, when presenting still emotions, may measure a fine-grained level of emotion word representation (i.e., adult-like, hence more difficult), extending the findings from Arias-Trejo and Plunkett (2010) that words that are conceptually (i.e., emotions, valence) and perceptually

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

(i.e., facial expressions) similar are the most difficult to separate for preschool children (Widen, 2013; Wu et al., 2022). Parental assessments, on the other hand, may report on coarser representational levels of emotion word comprehension and production (i.e., the first sign of emotion word comprehension and production).

The current study

The current study aimed to assess the relationship between Norwegian two-to-five-year-old children's accuracy on a tablet-based 4AFC emotion word comprehension task and their parents' reports on both their emotion word comprehension and production, since we lack knowledge on the relationship between indirect and direct measures in emotion word comprehension and production. Children were included from the age of two to examine the developmental trajectory of emotion word knowledge from the first use of emotion words (see Ridgeway et al., 1985). We used an extended list of emotions expressed by child actors (previously validated for research, cf. Baron-Cohen, 2002) to extend previous research that primarily focused on variations of the basic emotions (Declercq et al., 2019; Wu et al., 2022).

In line with general word vocabulary findings (Libertus et al., 2015; Sattler et al., 1985), we expected that parent-reported emotion word comprehension would predict children's overall AFC accuracy. However, given findings that parents report on their toddlers' word knowledge on a coarse representational level (Lo et al., 2021) that may be difficult to capture in a more fine-grained AFC (see Widen & Russell, 2008; Woodard et al., 2022), it would also be plausible to find an inconsistent alignment, as the AFC and parental checklists may tackle different representational levels of emotion word comprehension. Furthermore, we expected that children's reported emotion word comprehension, production and performance in 4AFC task would improve with age, and that girls would outperform boys in both measures, paralleling earlier research on gender differences in preschoolers' language (Simonsen et al., 2014; Stangeland et al., 2018) and emotion recognition (Hall, 1979; McClure, 2000) skills.

Method

Participants

Parents (N = 350) of two-to-five-year-old children (age range 2;0-5;8 years, M = 4.18, SD = 1.06) from eight kindergartens in Asker and Oslo (Norway) participating in a larger intervention study (Havighurst et al. 2022) were invited to let their children participate in the current study. The study was approved by the Norwegian Center for Research Data. One-hundred parents filled out the consent form. Children that were in kindergarten on the testing day (N = 81) were assessed. Children whose mother tongue was not Norwegian were reported to have auditory or visual problems, or were born preterm were excluded from the analyses (N = 23).[1] Of the remaining 58 children, nine were excluded: four children did not complete the test, and five failed the familiarization task (see below). In total, 49 children (22 girls) between 24 and 68 months of age participated in the study (M = 50.2, SD = 12.8).

[1]These exclusion criteria were grouped together in the parental questionnaire and thus the breakdown per exclusion criterion is not available.

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

Table 1. Distribution of Children Across Age and Gender

	Age 2		Age 3		Age 4		Age 5		Full sample	
	n	%	n	%	n	%	n	%	n	%
Children	11		12		12		14		49	
Gender										
Girls	5	45.45	5	41.67	2	16.67	10	71.43	22	44.90
Boys	6	54.55	7	58.33	10	83.33	4	28.57	27	55.10

The gender distribution was skewed across the age groups ($\chi^2(3) = 7.90$, $p = .050$): there were only two four-year-old girls and four five-year-old boys in the sample (see Table 1).

Stimuli

We complemented the six so-called basic emotions (English in single quotation marks): *glad* 'happy', *sint* 'angry', *lei seg* 'sad', *overrasket* 'surprised', *redd* 'scared' and *ekkelt* 'yucky' (for 'yucky' as a proxy for 'disgust', see Camras & Allison, 1985; Widen & Russell, 2003) with six additional emotion words *fornøyd* 'content', *frustrert* 'frustrated', *bekymret* 'worried', *sjenert* 'shy' or 'intimidated', *irritert* 'irritated', *spent* 'excited'. The additional emotion words were selected based on words used in the intervention study (Havighurst et al., 2022) that followed the current assessment. For each emotion word, a child-portrayed facial expression photograph (half girls) was selected from Baron-Cohen (2002). Pictures ($n = 12$) were equalized for brightness, size and resolution and grouped into three 4-picture blocks: Block 1 contained 'happy', 'sad', 'scared', and 'angry', Block 2 contained 'disgusted', 'content', 'surprised', and 'frustrated', and Block 3 contained 'irritated', 'worried', 'intimidated', and 'excited'. The latter block was only presented to the four-to-five-year-old children. The combinations were selected partly based on earlier research on basic emotions (e.g., Widen, 2013), partly to maximize recognition by presenting frequently confused emotions separately (anger and disgust, fear and surprise, and the positive emotions of excitement, happiness and contentment, see Declercq et al., 2019; Widen & Russell, 2008). Further, we wished to make the task feasible and not too difficult to avoid discouraging the youngest children (see Wu et al., 2022). Four familiarization trials displayed a dog, an apple, a car, and a ball.

Procedure

Consent and demographic information

Parents filled in the consent form and a questionnaire including parents' educational status and language background. All mothers except for four held either a masters' or a doctoral degree. A Fisher's exact test revealed that there was no significant difference in maternal education level across ages, $p = .081$. Maternal educational level was therefore not assessed further.

Additionally, parents were, via questionnaire, instructed to check from 12 emotion words those that they believed their child understood ("understands") and understood AND produced ("understands and says") in line with the CDI formulation (Norwegian version: Kristoffersen & Simonsen, 2012). Instructions specified that children might not always say the words they understood.

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

Children's assessment

After receiving parents' consent forms, children were tested in their kindergartens by an experimenter. The assessment was performed using a Samsung tablet with a touch-screen function on an experimental e-Babylab platform developed by Lo et al. (2023). Children were first given four practice trials to familiarize themselves with the procedure. On each trial, the child saw four pictures and heard an instruction: *Kan du ta på eplet/bilen/hunden/ballen?* 'can you touch the apple/car/dog/ball?'. Upon a touch, the next trial was displayed on the screen. Children who failed three or more of the four familiarization trials (N = 5) or did not complete the entire task (N = 4) were excluded from the analysis.

The emotion word blocks were presented in a fixed order. Within every block, the order of emotion words was randomized across participants. To ensure consistency, all instructions were recorded by a native Norwegian female speaker through the tablet's internal speakers. For 'disgust', the instruction was *kan du ta på ansiktet som synes noe er ekkelt?* 'can you touch the face that finds something yucky?' For all other emotions, the instruction was *kan du ta på ansiktet som er [emotion word]?* 'can you touch the face that is [emotion word]?'. Between trials, a smiling emoticon presented small encouragement prompts, e.g., *Nå er du nesten ferdig, bra*! 'now you are almost done, well done!'. In the absence of an answer within 20 seconds, the study moved to the next task. This cut-off was based on a pilot and selected because longer waiting times could be perceived as pressuring the child. Missing answers were excluded from the analysis as they could be attributed to a range of reasons (e.g., child intimidation, distraction, or no answer). Average response time for correct answers varied from 8.09 seconds (95% CI = 6.99, 9.20) for the two-year-olds to 6.28 seconds (95% CI = 5.89, 6.68) for the five-year-olds.

Data processing

All descriptive statistics were performed using IBM SPSS statistics version 26. T-tests and the generalized mixed effects models were performed in R (R Development Core Team, 2012). Parent-reported emotion word comprehension and production, as well as child AFC accuracy were coded as binary scores. *Ekkelt* 'disgusted' was excluded from the analysis because it was miswritten as *forferdet* 'terrified' in the parental checklist. Post-hoc, we created a confusion matrix of the children's AFC responses.

Results

Table 2 depicts the descriptive data of the children's mean AFC accuracy, parent-reported emotion word comprehension and production, and agreement scores for each emotion word and age category. On average, children recognized 55% of emotion words and were reported by their parents to understand 92% and produce 69% of the emotion words. Agreement between parent-reported comprehension and child accuracy concerned 58% of the trials, underestimation concerned 2% and overestimation concerned 40% of the trials. Agreement between parent-reported production and child accuracy concerned 59% of the trials, underestimation concerned 13%, and overestimation concerned 28% of the trials. Binomial tests grouping the three foils against the target response indicated that for all emotion words except for 'frustrated', 'excited', and 'intimidated, AFC accuracy was above chance level of .25, $p < .05$ (two-tailed) for at least one age group.

There was a ceiling effect for parent-reported emotion word comprehension for some words: more than 90% of children were reported to understand the emotion words in

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

Table 2. Mean Emotion Word Recognition Accuracy, Parent-Reported Emotion Word Comprehension and Emotion Word Production and Agreement Scores by Age and Emotion Word

			Child accuracy			Comprehension						Production					
						Parent report			Agreement			Parent report			Agreement		
			M	95% CI		M	95 % CI		M	95 % CI		M	95 % CI		M	95 % CI	
Emotion	Age	N		LL	UL		LL	UL		LL	UL		LL	UL		LL	UL
Angry	2	10	0.50	0.19	0.81	1.00			0.50	0.20	0.80	0.80	0.50	1.00	0.50	0.20	0.80
	3	12	0.75***	0.43	0.95	1.00			0.75	0.50	1.00	0.92	0.75	1.00	0.67	0.42	0.92
	4	12	0.92***	0.62	1.00	1.00			0.92	0.75	1.00	1.00			0.92	0.75	1.00
	5	14	0.93***	0.66	1.00	1.00			0.93	0.79	1.00	1.00			0.93	0.79	1.00
Happy	2	11	0.64**	0.31	0.89	1.00			0.64	0.36	0.91	0.64	0.36	0.91	0.45	0.18	0.73
	3	12	0.92***	0.625	1.00	1.00			0.92	0.75	1.00	1.00			0.92	0.75	1.00
	4	12	0.92***	0.62	1.00	1.00			0.92	0.75	1.00	1.00			0.92	0.75	1.00
	5	14	1.00***	0.77	1.00	1.00			1.00			1.00			1.00		
Sad	2	11	0.27	0.06	0.61	1.00			0.27	0.00	0.55	0.82	0.55	1.00	0.27	0.00	0.55
	3	12	0.42	0.15	0.72	1.00			0.42	0.17	0.75	1.00			0.42	0.17	0.67
	4	12	0.67**	0.35	0.90	0.92	0.75	1.00	0.58	0.25	0.83	0.92	0.75	1.00	0.58	0.33	0.83
	5	13	0.85***	0.44	0.98	1.00			0.85	0.62	1.00	1.00			0.85	0.62	1.00
Scared	2	11	0.36	0.11	0.69	0.91	0.73	1.00	0.45	0.18	0.73	0.64	0.36	0.91	0.55	0.27	0.82
	3	12	0.58*	0.28	0.85	1.00			0.58	0.25	0.83	1.00			0.58	0.33	0.83
	4	12	0.50	0.21	0.79	1.00			0.50	0.25	0.75	0.92	0.75	1.00	0.58	0.33	0.83
	5	13	0.62**	0.32	0.86	1.00			0.62	0.38	0.85	1.00			0.62	0.31	0.85

Table 2. *(Continued)*

			Child accuracy			Comprehension						Production					
						Parent report			Agreement			Parent report			Agreement		
			M	95% CI		M	95 % CI		M	95 % CI		M	95 % CI		M	95 % CI	
Emotion	Age	N		LL	UL		LL	UL		LL	UL		LL	UL		LL	UL
Content	2	11	0.36	0.11	0.70	0.64	0.36	0.91	0.55	0.27	0.82	0.09	0.00	0.27	0.73	0.45	1.00
	3	12	0.17	0.02	0.48	1.00			0.17	0.00	0.42	0.42	0.17	0.67	0.42	0.17	0.75
	4	11	0.45	0.17	0.77	1.00			0.45	0.18	0.73	0.82	0.55	1.00	0.27	0.00	0.55
	5	14	0.71***	0.42	0.92	0.93	0.79	1.00	0.64	0.36	0.86	0.64	0.36	0.86	0.64	0.36	0.86
Disgusted	2	11	0.73**	0.39	0.94												
	3	12	0.67**	0.35	0.90												
	4	12	0.75***	0.43	0.96										.		
	5	14	0.86***	0.57	0.98										.		
Frustrated	2	11	0.18	0.02	0.52	0.45	0.18	0.73	0.73	0.45	1.00	0.18	0.00	0.45	0.64	0.36	0.91
	3	11	0.27	0.06	0.61	0.82	0.55	1.00	0.27	0.00	0.55	0.27	0.00	0.55	0.45	0.18	0.73
	4	12	0.25	0.05	0.57	0.67	0.42	0.92	0.25	0.00	0.50	0.17	0.00	0.42	0.75	0.50	1.00
	5	14	0.29	0.08	0.58	0.71	0.50	0.93	0.57	0.29	0.86	0.29	0.07	0.57	0.57	0.29	0.79
Surprised	2	11	0.36	0.11	0.69	0.45	0.18	0.73	0.55	0.27	0.82	0.27	0.00	0.55	0.36	0.09	0.64
	3	12	0.08	0.00	0.38	1.00			0.08	0.00	0.25	0.83	0.58	1.00	0.25	0.00	0.50
	4	11	0.27	0.06	0.61	1.00			0.27	0.00	0.55	0.73	0.45	1.00	0.36	0.09	0.64
	5	13	0.69**	0.39	0.91	1.00			0.69	0.46	0.92	0.77	0.54	1.00	0.62	0.31	0.85

Table 2. *(Continued)*

			Child accuracy			Comprehension						Production					
						Parent report			Agreement			Parent report			Agreement		
			M	95% CI		M	95 % CI		M	95 % CI		M	95 % CI		M	95 % CI	
Emotion	Age	N		LL	UL		LL	UL		LL	UL		LL	UL		LL	UL
Excited	4	10	0.30	0.07	0.65	1.00			0.30	0.00	0.60	0.70	0.40	0.90	0.60	0.30	0.90
	5	14	0.50	0.23	0.77	0.86	0.64	1.00	0.64	0.36	0.86	0.79	0.57	1.00	0.57	0.29	0.79
Irritated	4	11	0.64**	0.31	0.89	0.91	0.73	1.00	0.73	0.45	1.00	0.45	0.18	0.73	0.45	0.18	0.73
	5	14	0.93***	0.66	1.00	0.86	0.64	1.00	0.93	0.79	1.00	0.57	0.29	0.79	0.64	0.36	0.86
Intimidated	4	11	0.27	0.06	0.61	0.82	0.55	1.00	0.27	0.00	0.55	0.45	0.18	0.73	0.45	0.18	0.73
	5	13	0.38	0.14	0.68	1.00			0.38	0.15	0.69	0.85	0.62	1.00	0.38	0.15	0.62
Worried	4	11	0.18	0.02	0.52	0.91	0.73	1.00	0.27	0.09	0.55	0.27	0.00	0.55	0.55	0.27	0.82
	5	14	0.64**	0.35	0.87	0.93	0.79	1.00	0.71	0.50	0.93	0.29	0.07	0.50	0.50	0.21	0.79

Note. Binomial t-test against $p = .25$. * $p < .05$. ** $p < .01$. *** $p < .001$. CI = confidence interval; LL = lower limit; UL = upper limit

Table 3. Output From Generalized Linear Mixed Effects Model

Variable	χ2	Df	*p*
Age	22.0194	1	< .001***
Gender	0.9381	1	0.333
Comprehension	1.8248	1	0.177
Production	1.7741	1	0.183
Stimulus	59.6377	10	< .001***
Age*Gender	0.8687	1	0.351
Age*Production	4.6005	1	0.032*

* $p < .05$. ** $p < .01$. *** $p < .001$.

Block 1 across all ages. Only 'frustrated', 'irritated', and 'intimidated' were not reported by parents to be understood by at least one age group. Parent-reported emotion word production had a wider range of scores. Only the emotion words used in Block 1 reached a ceiling effect and only in the older age groups.

A binomial generalized linear mixed effects model was run to predict 4AFC accuracy (accurate, wrong) on each trial. Age (in years), gender of the child, parent-reported comprehension, parent-reported production, emotion word, and the interactions between age and gender and between age and production were added as fixed effects. Child was added as a random effect[2]. Age 2 was the reference level. To examine the significance of the main effects, the *anova* function (lmerTest package by Kuznetsova et al., 2017) was run on the model. Theoretical conditional R^2 (R_c^2) was calculated using the *MuMin* package (Barton, 2013) for significant factors. Table 3 presents the model output.

The results revealed that age significantly predicted accuracy. As expected, AFC accuracy increased significantly with age, $\chi^2(1) = 21.60$, $p < .001$, $R_c^2 = .06$. The factor emotion word was also significant, $\chi^2(10) = 59.23$, $p < .001$, $R_c^2 = .36$, indicating that some emotions, e.g., happiness, were easier recognized than others, e.g., frustration (see Table 2). The interaction between age and production was also significant: parent-reported emotion word production was a better predictor of the older children's AFC accuracy compared to the younger children, $\chi^2(1) = 4.63$, $p = .031$, $R_c^2 = .10$ (see Figure 1). Neither gender, gender-age-interaction, comprehension nor production were significant predictors of AFC accuracy (all *ps* > .1).

A confusion matrix (inspired by Widen & Russell, 2008) was added post-hoc to assess whether children's recognition errors were systematic, see Figure 2. The confusion matrix revealed that children most often selected the target emotion, except for frustrated and intimidated faces. Post hoc, two-tailed binomial t-tests compared the selection frequency with a chance level of $p = .25$ to assess whether any non-target emotion expressions were selected above chance level. Children selected the scared face above chance level (39.6%) when the target emotion was sadness, $N = 19$, $p = .029$ and the surprised face (39.6%)

[2]The first binomial model containing an interaction between comprehension and accuracy failed to converge. The *anova* that was produced reported that comprehension*age contributed the least to the model fit and was not significant, $\chi^2(1) = 1.74$, $p = .19$. Therefore, a new model without this interaction was run to reach convergence.

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

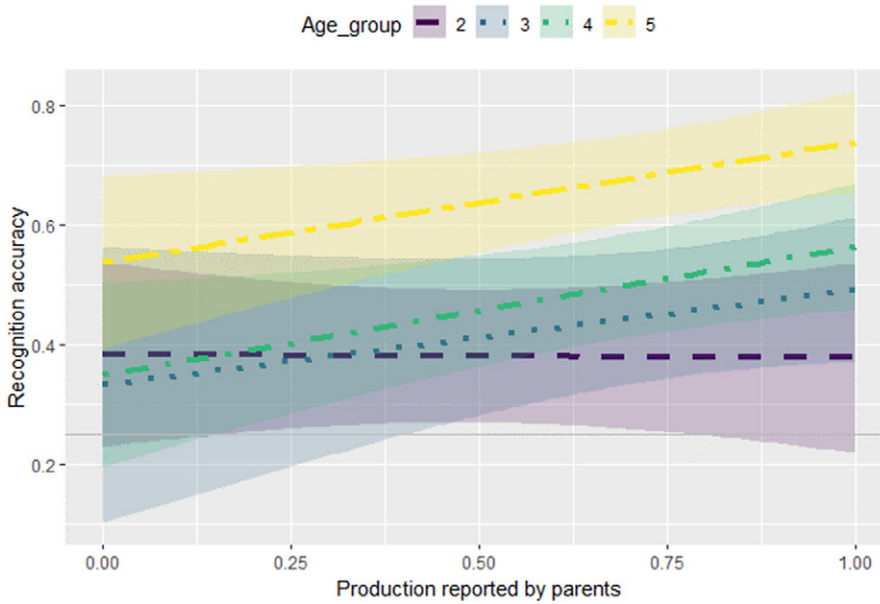

Figure 1. Average Production and Accuracy Across Emotion Words for Every Child Sorted by Age Group. Lightly Colored Areas Indicate 95% Confidence Intervals. The Dashed Line Indicates Recognition Accuracy Chance Level of 25%.

when the target emotion was frustration, $N = 19$, $p = .029$). No other (non-target) emotion faces were selected significantly above chance level. Note, however, that these results should be treated with caution, as emotion expressions that were expected to be confused (e.g., anger and disgust) were presented in separate blocks. The post hoc nature of the findings requires further replication of these results.

Discussion

The current study aimed to assess the relationship between parent-reported emotion word comprehension and production and children's accuracy on a 4AFC tablet-based emotion word comprehension task in Norwegian two-to-five-year-old children. The results revealed a general effect of age, with higher AFC accuracy in older children, in line with both emotion (Baron-Cohen et al., 2010; Ridgeway et al., 1985; Widen, 2013) and general vocabulary research (Dale & Fenson, 1996).

There was a significant effect of emotion words, indicating different developmental trajectories across emotions. Of the basic emotions, happy and disgusted faces were selected above chance level by all age groups, followed by angry, scared and, finally, surprised faces. The complex emotions of fear, contentment and worry were all recognized above chance level by the five-year-old children, and the irritated face was recognized earlier than the basic emotion of surprise. Except for the face expressing disgust, which is often confused with anger (Widen & Russell, 2010b), this trajectory parallels findings from previous emotion recognition research (Widen, 2013; Widen & Russell, 2008) and emotion word comprehension and production (Baron-Cohen et al., 2010; Ridgeway et al., 1985). No gender differences were found in the current study,

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

Block 1

Emotion Word	N	Response (Selected Face)			
		Angry	Happy	Sad	Scared
Angry	48	79.2	6.3	10.4	4.2
Happy	49	4.1	87.8	6.1	2.0
Sad	48	0.0	4.2	56.3	39.6
Scared	48	6.3	8.3	33.3	52.1

Block 2

		Content	Disgusted	Frustrated	Surprised
Content	48	43.8	12.5	25.0	18.8
Disgusted	49	8.2	75.5	10.2	6.1
Frustrated	48	14.6	20.8	25.0	39.6
Surprised	47	31.9	19.1	12.8	36.2

Block 3

		Excited	Irritated	Intimidated	Worried
Excited	24	41.7	8.3	33.3	16.7
Irritated	25	4.0	80.0	8.0	8.0
Intimidated	24	8.3	25.0	33.3	33.3
Worried	25	16.0	20.0	20.0	44.0

Figure 2. Confusion Matrix Representing the Percentage of Each Emotion Word That was Categorized by the Children as Each Emotion Face.

contrary to our expectations (Hall, 1979; McClure, 2000; Simonsen et al., 2014; Stangeland et al., 2018). Potential gender effects may, however, be obscured by the imbalanced gender distribution in the sample. Hence, the gender effect must be confirmed in future better gender balanced studies.

Contrary to our predictions, neither parent-reported emotion word comprehension nor production predicted children's AFC accuracy alone. Parents seemingly overestimated their children's emotion word comprehension. However, the interaction between parent-reported emotion word production and age was significant and predicted children's accuracy on the AFC, suggesting that parental reports of emotion word production predicted accuracy on the 4AFC better in older than in younger children. This finding differs from research that reports a relationship between indirect and direct measures of word comprehension in general vocabulary development (Libertus et al., 2015) but is in line with the finding from Sattler et al. (1985) that parents overestimated their children's AFC performance.

The lack of alignment between parental reports and children's emotion word comprehension could be due to the nature of the AFC task. It is possible that parents reported a coarse emotion word comprehension level, while the AFC task tackled a more fine-grained comprehension level. This could be due to the conceptual and perceptual similarities of the emotions. Such an interpretation would parallel findings from Lo et al. (2021) and Arias-Trejo and Plunkett (2010) that children's early word recognition is less stable in AFCs using semantically similar words, suggesting an underspecified representational level. This interpretation is further supported by the finding from Wu

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

Table 4. Percent of Children That Understood (Baron-Cohen et al., 2010), Understood and Produced (Ridgeway et al., 1985) and Understood, Produced, and Recognized (Current Study) Each Emotion Word

	Current Study (Norwegian)			Baron-Cohen et al. (2010) (British)	Ridgeway et al. (1985) (American)	
Emotion	Accuracy	Comprehension	Production	Comprehension	Comprehension	Production
Happy	96	100	100	97	98	96
Angry	93	100	100	88	98	85
Disgusted	81	.	.	49	50	29
Irritated	80	88	52	23	72[a]	33[a]
Sad	76	96	96	100	97	92
Content	60	96	72	13	9	1
Scared	56	100	96	93	98	92
Surprised	50	100	75	80	90	79
Worried	44	92	23	77	70	45
Excited	42	92	75	75	88	71
Intimidated	33	92	67	80	78	63
Frustrated	27	69	23	6	29	6

Note: Comprehension from Baron-Cohen et al. (2010) is the percentage of 4-to-6-year-olds that "clearly understood" each emotion word. For Ridgeway et al. (1985), percentages are calculated for children between four and five years of age that comprehended emotion words ("understand (..) when someone else used them to describe a feeling") or produced emotion words ("ever used to refer to his or her or to other people's mood or feelings") (Ridgeway et al., 1985, p. 902). Results from the current study present percent of 4- and 5-year-olds that understood, produced, and recognised every emotion word. Results are sorted by emotion word recognition accuracy in the current task.
[a]Irritated was not assessed in Ridgeway et al (1985). Results are therefore from *annoyed.*

et al. (2022), where children were able to distinguish emotion words in a 2AFC design when they were tested at home with their parents present. These findings suggest that characteristics of the AFC task can influence the word comprehension accuracy with it being more fine-grained (and thereby more difficult for toddlers) when emotion words are conceptually and perceptually similar (see Wu et al., 2022).

To evaluate the role of distractor stimuli on target identification (drawing parallels to the studies by Lo et al., 2021; Widen & Russell, 2008), we established a confusion matrix of responses (see Figure 2). The results showed that the scared face was selected significantly above chance level when sadness was the target emotion, replicating findings from Widen and Russell (2008). Furthermore, the face expressing disgust was recognized significantly above chance level, whereas earlier findings reported that disgust and anger were confused due to perceptual similarities (Widen, 2013; Widen & Russell, 2008, 2010b). Since children in the current study were presented to anger and disgust in separate blocks, this finding may suggest a burgeoning concept of disgust that was not captured in other studies that combined the angry and disgusted expressions in the same set of pictures.

An alternative (but not mutually exclusive) interpretation for the misalignment between direct and indirect measures of emotion comprehension could be that parents overestimated their children's emotion word comprehension and production, as suggested by the near-ceiling effect for many emotion words. Similar overestimation of children's performance was also found in general vocabulary tests (Sattler et al., 1985) and

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

emotion recognition (Kårstad et al., 2014). In general, the Norwegian parents in the current study reported their children to understand and, partly, produce more emotion words than their English-speaking counterparts, as reported in Baron-Cohen et al. (2010) and Ridgeway et al. (1985, see Table 4). For example, 92% of Norwegian children were reported to understand 'worried', compared to 77% of British and 70% of American children.

Reporting word comprehension is a difficult task for parents, as this must be inferred from their children's behavior, meaning that it is sensitive to the exact instructions provided to parents (Tomasello & Mervis, 1994). In the current study, parents were asked whether the child "understood" an emotion word, whereas Baron-Cohen et al. (2010) asked whether the child "clearly understood" an emotion word, and Ridgeway et al. (1985) asked whether the child "understood the word when it described a feeling". The specification of what it means to understand an emotion word in the two other studies may have led those parents to be more conservative in their emotion word comprehension reports than the parents in the current study. This interpretation suggests that the parental reports may not be reliable on an item-to-item basis, in line with general vocabulary research (Sattler et al., 1985).

However, the levels of parent-reported emotion word comprehension and production may also reflect the high educational level of the parents in the current sample. Children with highly educated parents are more exposed to language compared to families from lower educational backgrounds (Dailey & Bergelson, 2022), and highly-educated parents are less likely to overestimate vocabulary on the CDI compared to other groups (Feldman et al., 2000). These findings contradict the interpretation of parental overestimation and instead suggest that parents may have accurately reported high levels of emotion word vocabulary.

Taken together, the results of the current study suggest that parent-reported and AFC measures of emotion word comprehension do not measure the same representational levels in two-to-five-year-olds. This may reflect either that parents reported on a broad representational level of emotion word comprehension (Widen & Russell, 2008; Woodard et al., 2022) not captured by the fine-grained AFC task, or, alternatively, that parents overestimated their children's emotion word comprehension. There was an alignment between parental report of production and emotion word comprehension in older children, suggesting that parent-reported production and AFC emotion word comprehension tasks may measure similar representational levels as children grow. This could suggest that word production (hearing the child produce the word in the right context) might approach the representational level of knowledge tested in the AFC task. These results are tentative and require further replication due to the exploratory nature of the confusion matrix and post hoc tests. However, the findings suggest that researchers should carefully consider which representational level of emotion word comprehension they assess in preschool children.

Limitations and future directions

Several attempts were made to make the current study feasible for the two-year-olds in the current task; they were tested in known surroundings in their kindergarten, the tablet automatically continued if they did not answer, and the level of difficulty increased with each experimental block. However, the 4AFC task may be too demanding to capture coarse emotion word comprehension in the youngest children (Widen, 2013). To

https://doi.org/10.1017/S0305000923000727 Published online by Cambridge University Press

enhance feasibility in young children, measurements could be inspired by Wu et al. (2022) who ran the experiment at home, let the children sit with their parents, and only presented the children to one target emotion and one foil. However, note that such a task will likely reflect early and coarse emotion word comprehension.

Researchers interested in direct assessment of fine-grained emotion word comprehension, however, will notice that a problem with the 4AFC task is that children may have 'correct' responses in 25% of cases by chance alone. A different paradigm could help address this problem by incorporating more elements of the emotion concept such as related behavior and causes (Widen & Russell, 2010a) and ask children for emotions that are not in the display (as done by Nelson & Russell, 2016), or display emotions that are not mentioned to avoid systematic guesses. However, such a task may run the risk of intimidating the children. Both tasks could fruitfully be compared to indirect parental reports to further assess the potential alignment of the methods.

Conclusion

Parental reports of emotion word comprehension did not predict Norwegian two-to-five-year-old children's accuracy in a 4AFC emotion word comprehension task, suggesting that direct and indirect measures of emotion word comprehension do not measure the same representational level of emotion word comprehension in preschoolers. However, parental reports of emotion word production predicted emotion word comprehension on the AFC in the older children. These findings suggest that widely cited parental reports of emotion word comprehension (Baron-Cohen et al., 2010; Ridgeway et al., 1985) do not necessarily report the same representational level of emotion word comprehension as measured by AFCs in preschool children (Declercq et al., 2019). This points to a need for future development of reliable tools to assess emotion word comprehension for both research and for determining children's early emotion competence (e.g., Pons et al., 2004).

Acknowledgements. The research was conducted as a part of a larger project called Tuning in to Kids (Havighurst et al., 2022), where the current results were part of baseline results for an intervention. The project was approved by Norwegian Centre for Research Data.

Funding. This research was supported by FUS Kindergartens Norway. No grant number provided. FUS Kindergartens had no involvement in the research design or other parts of the research process. NK was supported by the Research Council of Norway through project number 301625, and its Centres of Excellence funding scheme [project number 223265].

Competing interest. The authors declare none.

Link to data repository. https://osf.io/nfzpw/?view_only=fc6fb34e93f74bdc8c77ab25951e8580

References

Arias-Trejo, N., & **Plunkett, K.** (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, **105**(1), 63–80. https://doi.org/10.1016/j.jecp.2009.10.002

Baron-Cohen, S. (2002). *Mind reading: The interactive guide to emotions (version 1.0/DVD ROM)*. Sheffield: Human emotions.

Baron-Cohen, S., **Golan, O.**, **Wheelwright, S.**, **Granader, Y.**, & **Hill, J.** (2010). Emotion Word Comprehension from 4 to 16 Years Old: A Developmental Survey. *Frontiers in evolutionary neuroscience*, **2**, 109. https://doi.org/10.3389/fnevo.2010.00109

Barton, K. (2013). *Package "MuMin". Version,* 1.*18.* In

Camras, L. A., & **Allison, K.** (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior*, **9**(2), 84–94. https://doi.org/10.1007/BF00987140

Dailey, S., & **Bergelson, E.** (2022). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. *Developmental Science*, **25**(3), e13192. https://doi.org/10.1111/desc.13192

Dale, P. S., & **Fenson, L.** (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, **28**(1), 125–127. https://doi.org/ckg9kq

Declercq, C., **Marlé, P.**, & **Pochon, R.** (2019). Emotion word comprehension in children aged 4–7 years. *The Educational and Developmental Psychologist*, **36**(2), 82–87. https://doi.org/10.1017/edp.2019.17

Dunn, D. M. (2019). *PPVT-5: Peabody picture vocabulary test* (5 ed.). Pearson Assessments.

Feldman, H. M., **Dollaghan, C. A.**, **Feldman, H. M.**, **Dollaghan, C. A.**, **Campbell, T. F.**, **Kurs-Lasky, M.**, **Janosky, J. E.**, & **Paradise, J. L.** (2000). Measurement properties of the MacArthur communicative development inventories at ages one and two years. *Child Development*, **71**(2), 310–322. https://doi.org/10.1111/1467-8624.00146

Hall, J. A. (1979). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, **85**(4), 845. https://doi.org/10.1037/0033-2909.85.4.845

Havighurst, S. S., **Edvoll, M.**, **Tidemann, I.**, **Bølstad, E.**, **Holme, H.**, **Bergum Hansen, M.**, **Eikseth, H.C.**, & **Nygaard, E.**, (2022). A Randomized Controlled Trial of an Emotion Socialization Intervention in Norwegian Kindergartens, Early Education and Development, https://doi.org/10.1080/10409289.2022.2160617

Kårstad, S. B., **Kvello, Ø.**, **Wichstrøm, L.**, & **Berg-Nielsen, T. S.** (2014). What do parents know about their children's comprehension of emotions? Accuracy of parental estimates in a community sample of preschoolers. *Child: Care, Health and Development*, **40**(3), 346–353. https://doi.org/10.1111/cch.12071

Kristoffersen, K. E., & **Simonsen, H. G.** (2012). *Tidlig språkutvikling hos norske barn: MacArthur-Bates foreldrerapport for kommunikativ utvikling*. Novus forlag.

Kuznetsova, A., **Brockhoff, P. B.**, & **Christensen, R. H. B.** (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, **82**(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Libertus, M. E., **Odic, D.**, **Feigenson, L.**, & **Halberda, J.** (2015). A Developmental Vocabulary Assessment for Parents (DVAP): Validating Parental Report of Vocabulary Size in 2- to 7-Year-Old Children. *Journal of Cognition and Development*, **16**(3), 442–454. https://doi.org/10.1080/15248372.2013.835312

Lo, C. H., **Hermes, J.**, **Kartushina, N.**, **Mayor, J.**, & **Mani, N.** (2023). e-Babylab: An open-source browser-based tool for unmoderated online developmental studies. *Behavior Research Methods*. https://doi.org/10.3758/s13428-023-02200-7

Lo, C. H., **Rosslund, A.**, **Chai, J. H.**, **Mayor, J.**, & **Kartushina, N.** (2021). Tablet assessment of word comprehension reveals coarse word representations in 18–20-month-old toddlers. *Infancy*. https://doi.org/10.1111/infa.12401

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, **126**(3), 424. https://doi.org/10.1037/0033-2909.126.3.424

Nelson, N. L., & **Russell, J. A.** (2016). Building emotion categories: Children use a process of elimination when they encounter novel expressions. *Journal of Experimental Child Psychology*, **151**, 120–130. https://doi.org/10.1016/j.jecp.2016.02.012

Nook, E. C., **Stavish, C. M.**, **Sasse, S. F.**, **Lambert, H. K.**, **Mair, P.**, **McLaughlin, K. A.**, & **Somerville, L. H.** (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, **20**(5), 773–792. https://doi.org.ez.statsbiblioteket.dk:2048/10.1037/emo0000609

Pons, F., **Harris, P. L.**, & **Rosnay, M. d**. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, **1**(2), 127–152. https://doi.org/fh34p6

R Development Core Team. (2012). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing.

Ridgeway, D., **Waters, E.**, & **Kuczaj, S. A.** (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology*, **21**(5), 901–908. https://doi.org/bgmv79

Ruba, A. L., & **Pollak, S. D.** (2020). The Development of Emotion Reasoning in Infancy and Early Childhood. *Annual Review of Developmental Psychology*, **2**(1), 503–531. https://doi.org/10.1146/annurev-devpsych-060320-102556

Sattler, J. M., **Feldman, J.**, & **Bohanan, A. L.** (1985). Parental estimates of children's receptive vocabulary. *Psychology in the Schools*, **22**(3), 303–307. https://doi.org/fjr8dk

Simonsen, H. G., **Kristoffersen, K. E.**, **Bleses, D.**, **Wehberg, S.**, & **Jørgensen, R. N.** (2014). The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, **34**(1), 3–23. https://doi.org/10.1177/0142723713510997

Stangeland, E. B., **Lundetræ, K.**, & **Reikerås, E.** (2018). Gender differences in toddlers' language and participation in language activities in Norwegian ECEC institutions. *European Early Childhood Education Research Journal*, **26**(3), 375–392. https://doi.org/10.1080/1350293X.2018.1463905

Streubel, B., **Gunzenhauser, C.**, **Grosse, G.**, & **Saalbach, H.** (2020). Emotion-specific vocabulary and its contribution to emotion understanding in 4- to 9-year-old children. *Journal of Experimental Child Psychology*, **193**, 104790. https://doi.org/fm6k

Tomasello, M., & **Mervis, C. B.** (1994). The Instrument Is Great, but Measuring Comprehension Is Still a Problem. *Monographs of the Society for Research in Child Development*, **59**(5), 174–179. https://doi.org/10.1111/j.1540-5834.1994.tb00186.x

Widen, S. C. (2013). Children's Interpretation of Facial Expressions: The Long Path from Valence-Based to Specific Discrete Categories. *Emotion Review*, **5**(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & **Russell, J. A.** (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, **39**(1), 114–128. https://doi.org/cjfxbs

Widen, S. C., & **Russell, J. A.** (2008). Children acquire emotion categories gradually. *Cognitive Development*, **23**(2), 291–312. https://doi.org/dwtx96

Widen, S. C., & **Russell, J. A.** (2010a). Differentiation in preschooler's categories of emotion. *Emotion*, **10**(5), 651–661. https://doi.org/10.1037/a0019005

Widen, S. C., & **Russell, J. A.** (2010b). The "disgust face" conveys anger to children. *Emotion*, **10**(4), 455–466. https://doi.org/10.1037/a0019151

Woodard, K., **Zettersten, M.**, & **Pollak, S. D.** (2022). The representation of emotion knowledge across development. *Child Development*, **93**(3). https://doi.org/10.1111/cdev.13716

Wu, Y., **Matteson, H.**, **Baker, C. M.**, & **Frank, M. C.** (2022). Angry, sad, or scared? Within-valence mapping of emotion words to facial and body cues in 2- to 4-year old children. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **44**(44).

Cite this article: Roepstorff, I.T., Mayor, J., Havighurst, S.S., & **Kartushina, N.** (2024). Same name, different representational levels? Misalignment of indirect parent-reported and direct alternative forced choice measures of emotion word comprehension in preschool children. *Journal of Child Language* 1–16, https://doi.org/10.1017/S0305000923000727