

ARTICLE

Categorising emotion words: the influence of response options

Barbra Zupan^{1*}, Lynn Dempsey² and Katelyn Hartwell¹

¹College of Health Sciences, School of Medical and Applied Sciences, Central Queensland University, Bruce Highway, Rockhampton, Queensland, Australia; ²Department of Applied Linguistics, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, Ontario, Canada

*Corresponding author. Email: b.zupan@cqu.edu.au

(Received 24 April 2022; Revised 22 July 2022; Accepted 23 August 2022)

Abstract

Words used to describe emotion are influenced by experience, context and culture; nevertheless, research studies often constrain participant response options. We explored the influence of response options on how people conceptualise emotion words in two cross-sectional studies. In Study 1 participants rated the degree to which a large set of emotion words ($n = 497$) fit five basic emotion categories – Happy, Sad, Angry, Fearful, Neutral. Twenty-four words that fit well within these categories were included in Study 2. In Study 2 response options were expanded to include two additional basic emotions (Disgust, Joy), and six complex emotions (Amusement, Anxiety, Contentment, Irritated, Pride, Relief). Only half of the Study 1 words were categorised into the same emotion categories in Study 2. An increase in diversity of ratings for both positive and negative valenced words suggested overlaps in people’s conceptualisations of emotion words. Results suggest potential benefits of providing research participants complex emotion categories of varying intensity, which may better reflect people’s nuanced conceptualisations of emotion. Future research exploring varied response options may provide further insight into how people categorise and differentiate emotion words.

Keywords: basic emotions; complex emotions; response options; categorisation; conceptualisation

1. Introduction

Our ability to perceive emotion is vital to well-being (Milders et al., 2003), impacting our capacity to understand and respond to others’ feelings (Fugate et al., 2018). Emotion perception involves recognising, integrating, and interpreting different cues (Czimskey & Marquardt, 2019; Zupan & Eskritt, 2022), a process at least partially driven by one’s internal awareness and representation of emotion (Ben-David et al.,

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

2011; Czimskey & Marquardt, 2019; Herpertz et al., 2016). In other words, our internal representation of emotion, or conceptual knowledge, helps us both describe our own emotional states and identify emotions in others (Adolphs, 2002; Adolphs et al., 2000; Phillips et al., 2003).

The words used to label emotional states are a key indicator of emotion knowledge (Barrett et al., 2011). Since words are influenced by experience, context and culture (Niedenthal et al., 2004; Strauss & Allen, 2008), how people think about and describe emotion varies widely. Ultimately, the label attached to an emotion is dependent on how it is internally represented in our lexicon (Barsalou, 1999, 2003), including the complexity which characterises the relationships among lexical entries (e.g. Shaver et al., 1987; Storm & Storm, 1987).

Perspectives regarding how the emotion lexicon is structured differ. According to the prototype model (e.g. Shaver et al., 1987), the structure is based on idealised representations (i.e. prototypes) formed from repeated experiences, with some words more central than others (Fehr & Russell, 1984; Hupka et al., 1999). For example, in the fear prototype, *anxiety* might be more central than *upset*. This prototype model aligns with the basic and discrete theory of emotion in which all emotions are captured in a core set of classes, for which there exists a linguistic label in every language (Ekman, 1992; Russell et al., 2011). Alternatively, the emotion lexicon might be a taxonomy, organised hierarchically with broad features housed at the superordinate level and finer distinctions captured at each subordinate level (Storm & Storm, 1987). The superordinate level could indicate whether a word carries an emotion or not. Subordinate levels might differentiate valence, intensity, and so on. The characteristics or properties that lead to finer distinctions might differ according to context and people's previous emotion experiences (Barrett, 2004). This taxonomic perspective is similar to the dimensional theory of emotion in which emotion is characterised by the degree to which various dimensions (e.g. valence, intensity) are activated (Russell, 1980). In other words, the final percept arises from a pattern of cues (Posner et al., 2005). These perspectives have been used to develop standardised lists of both emotional (e.g. cruel, delight, pleasant) and non-emotional (e.g. ambulance, chair, garage) English words that are structured according to the degree to which words represent emotion dimensions (Bradley & Lang, 1999) or according to the discrete emotion categories the words fit within (Stevenson et al., 2007; Strauss & Allen, 2008).

Regardless of the perspective one subscribes to, the emotion word people assign a stimulus is constrained by available response options. Many studies in emotion perception use a forced-choice paradigm in which participants are instructed to select from a list the word that best describes the stimulus. This list is typically limited to words representing four to six basic (universal) emotion categories (i.e. happy sad, angry, fearful, disgust, surprise) (Airdrie et al., 2018; Grosbras et al., 2018; Hauschild et al., 2020; Melendez et al., 2020; Zupan & Babbage, 2017). This format permits ease of analysis, reproducibility, and comparison across studies but also constrains responses, potentially resulting in artificial and inflated agreement among participants (Barrett et al., 2011; Limbrecht-Ecklundt et al., 2013; Nelson & Russell, 2013) and masking nuances in how people conceptualise and interpret emotion (Turkstra et al., 2017).

The labels people assign when identifying emotion are important because they guide responses to others' feelings (Torre & Lieberman, 2018), influencing interpersonal relationships (Trampe et al., 2015). In conversation, people use a wider variety of words to describe emotion than the four to six basic ones proposed by Ekman (1992), of which only one is clearly positive in valence. Turkstra et al. (2017) highlight

the functional importance of complex (e.g. anxiety, amusement) and social emotions (e.g. pride, guilt) in everyday human interactions. Given the option, people may well use the labels for these emotions, in addition to those for basic emotions, to describe research stimuli. If they are to capture nuances in how people think about and perceive emotion, researchers must consider alternative paradigms and/or response options for examining emotion perception. This is of particular import in research addressing developmental trajectories, individual variability, and cultural differences. Better understanding of the breadth of labels people use to classify emotions can also inform assessment and treatment for clinical populations.

The aim of this project was to test the hypothesis that people will identify words as portraying complex or social emotions if given the opportunity. To do so, we conducted two studies. In the first, we tested our assumptions about the limitations of traditional forced-choice labelling tasks using basic emotion categories. Specifically, we sought to identify how well emotion words, previously generated via a free labelling task (Zupan & Babbage, 2008), fit within a set of basic emotion categories. We used these data to make evidence-based decisions about the stimulus words we asked people to identify in Study 2. In Study 2, we investigated the influence of providing a broader array of response options on emotion categorisation using a subset of words from Study 1. Overall, our goal was to contribute to a more nuanced understanding of how people conceptualise and categorise emotion words and phrases. This work extends the work of studies focused on building a standardised corpora of emotion words in English (Stevenson et al., 2007; Strauss & Allen, 2008) by expanding the possible response options beyond the basic emotion categories proposed by Ekman (1992). More importantly, this work has implications for developing response formats and scoring guidelines for studies that ask participants to rate and/or classify emotion stimuli.

2. Study 1

2.1. Participants

Participants were recruited from a Canadian university via institutionally approved campus advertisements and invitations posted to course websites. They received either monetary compensation (\$10) or course credit for their participation. Only participants who were at least 18 years of age, spoke English as a primary language, and had no uncorrected vision deficits were eligible. Eighty-nine participants (17 males; 72 females), ranging in age from 18 to 54 years ($M = 23.49$; $SD = 6.41$) participated. The ratio of males to females in this sample is not balanced, but approximates that typically observed in introductory psychology courses of the type we recruited from (Dickinson et al., 2012). Given the imbalance we did not conduct gender-based analyses. On average, participants had completed 16.07 ($SD = 1.76$) years of education. The majority of participants (79.8%) self-identified as Caucasian.

2.2. Stimuli

Emotion words and/or phrases (maximum six words) were generated via an unpublished study (Zupan & Babbage, 2008) in which 82 participants (34 males; 48 females) freely labelled facial (dynamic and static) and vocal emotion expressions. These emotion expressions had been elicited in 10 male and 10 female speakers using

stimuli (film clips; novel excerpts; Zupan & Babbage, 2017) equally representing happy, sad, angry, fearful and neutral emotions. During this free labelling tasks, participants provided a total of 494 unique words/phrases (hereafter referred to as emotion words). Three words were erroneously included twice, resulting in a total of 497 words. As in Strauss and Allen (2008), the words were randomly divided into three lists, each consisting of 166–167 words. Each participant was presented with one list of words. The three duplicate words each appeared in different lists so no participant received any word twice. Ratings for the repeated words were combined in analyses.

2.3. Procedure

The cross-sectional design required that participants individually attend a single 60-minute face-to-face session. Each participant first completed a brief demographic questionnaire. Participants were then sequentially assigned to one of the three-word lists in the order in which they entered the study, with a total of 30 participants assigned to lists one and two, and 29 to list three. Using Cedrus SuperLab software, words were presented in size 36 black serif font on a white background via a 21.5-inch computer screen. To control for potential order effects, word order was randomised.

Using 9-point Likert scales, participants were prompted to rate each word first on intensity (1 = extremely low; 9 = extremely high), then valence (1 = extremely negative; 9 = extremely positive), then the degree to which the word represented *each* of five emotion categories (i.e. happy, sad, angry, fearful, neutral); the order of these categories remained consistent for each stimulus. The instructions for these ratings is provided in the Appendix. These five categories are considered distinctive without requiring complex appraisals, thoughts or judgments (Ekman & Cordaro, 2011) and are commonly used as the minimal response options in emotion studies (Airdrie et al., 2018; Ben-David et al., 2011; Zupan & Babbage, 2017). Evidence suggests that some emotion words present with neutral semantics (Larsen et al., 2006), hence the use of a neutral category. These categories also align to those used in previous studies aimed to build a standardised corpora of emotion words in English (Stevenson et al., 2007; Strauss & Allen, 2008). Following Strauss and Allen (2008), we opted to have participants rate words on valence and intensity. We opted to use intensity rather than arousal, another important dimension of emotion perception, because the former is easier to present to participants (i.e. it does not require visual depictions). Though intensity and arousal are not the same, Laukka et al. (2005) report a strong correlation between the two.

Participants used a response keypad with only 10 keys; numbers one through nine appeared on two lines (equally spaced); a tenth key labelled 'next' appeared on a third line. Limiting and spacing the keys reduced the likelihood of participants mistyping. Participants were not provided feedback at any time.

2.4. Data analysis

Descriptive statistics were calculated for demographics, valence, intensity, and degree-of-fit ratings, but only degree-of-fit ratings are reported here. Mean degree-of-fit ratings ranged from 1 to 9 in each of five emotion categories; higher scores indicated a better fit for the stimulus word. Since nine was the highest possible score,

we considered ratings of 8.00 or greater as indicating the emotion category provided an excellent fit for the stimulus word.

Next, we calculated a Simpson Diversity Index Score (SDIS) using degree-of-fit ratings, an approach that has been applied in other studies with a psychosocial focus (Koffer et al., 2016; Quoidbach et al., 2014; Ram et al., 2012; Zupan & Eskritt, 2022). The SDIS takes into account the emotion category that had the highest degree-of-fit rating and a frequency count for the number of times each emotion category was rated highest for each word. The resulting index is a nonparametric statistic representing the dispersion of scores across the emotion categories for each word. Dispersion values fall between 0 and 1, with values closer to 0 representing a more definite categorisation of the emotion. As the value of the SDIS increases, so does diversity across participant responses (Gregorius & Gillet, 2008; Zupan & Eskritt, 2022). Using existing guidelines, words with values between 0.1 and 0.4 were considered to have a low degree of diversity, words with values between 0.41 and 0.6 were classified as having moderate diversity and those with values of 0.61 or higher were considered to have widely dispersed ratings (Guajardo, 2015). Thus, a value of 0.61 or higher indicates that when asked to rate the degree to which a word fit each emotion category, participants rated numerous emotion categories similarly.

2.5. Results

Appendix A¹ in the Supplementary Material provides the mean degree-of-fit ratings for all 497 words for each emotion category as well as the SDIS; valence and intensity ratings are also provided in the Appendix. A results overview is provided in Table 1. Since participants rated each emotion category for each word, we also identified which emotion categories were rated as having the second highest degree of fit (see Fig. 1). A summary for each emotion category follows.

2.5.1. Happy

A total of 128 words (25.75%), the most for any category, were identified as best fitting in the Happy category. Overall, mean degree-of-fit ratings for words categorised as Happy ranged from 4.40 (*drunk*; *expectant*) to 8.70 (*happy*). Thirteen words (10.16%) received a mean rating of at least 8.00. All 13 of these words were low in diversity as indicated by SDISs ranging from 0.06 (*merry*) to 0.36 (*excited*). Moderate diversity was evident in the words with degree-of-fit ratings less than 8.00. The category most commonly rated as the second best degree of fit was Neutral ($n = 118$; 92.19%), with mean ratings ranging from 1.87 (*energetic*) to 5.57 (*a tease*). One word, *earnest*, received the same mean rating (5.10) for both Happy and Neutral, but had a lower standard deviation for Happy ($SD = 2.28$) so was listed in that category.

2.5.2. Sad

Of the 88 words identified as Sad, seven (7.95%) received degree-of-fit ratings of 8.00 or higher, with ratings ranging from 8.03 (*devastated*) to 8.87 (*sad*). Only *sad* was classified as low diversity, with an SDIS of 0.36. Of the remaining six words with

¹Supplementary Materials have been stored with Open Science Framework [Zupan, B. (2022, April 4)]. Emotion Words. Retrieved from <https://osf.io/uw95c/>

Table 1. Ratings summary for categories identified as highest degree of fit in study 1

| Category (<i>n</i> ;% of total words) | Ratings | 8.00–8.90 | 7.00–7.99 | 6.00–6.99 | 5.00–5.99 | 4.00–4.99 | 3.00–3.99 |
|--|--------------|-------------|-------------|-------------|-------------|-------------|-----------|
| Happy (<i>n</i> = 128; 25.75%) | <i>n</i> (%) | 13 (10.16%) | 51(39.84%) | 30 (23.44%) | 28 (21.87%) | 6 (4.69%) | |
| | Range | 8.07–8.70 | 7.00–7.97 | 6.00–6.90 | 5.07–5.93 | 4.40–4.90 | |
| | SDI range | 0.06–0.36 | 0.12–0.56 | 0.46–0.70 | 0.58–0.75 | 0.61–0.79 | |
| Sad (<i>n</i> = 88; 17.71%) | <i>n</i> (%) | 7 (7.95%) | 13 (14.77%) | 27 (30.68%) | 27 (30.68%) | 14 (15.90%) | |
| | Range | 8.03–8.87 | 7.00–7.90 | 6.00–6.97 | 5.03–5.97 | 4.03–4.97 | |
| | SDI range | 0.36–0.66 | 0.49–0.67 | 0.58–0.79 | 0.68–0.79 | 0.75–0.79 | |
| Angry (<i>n</i> = 107; 21.53%) | <i>n</i> (%) | 10 (9.34%) | 17 (15.89%) | 26 (24.30%) | 32 (29.90%) | 21 (19.63%) | 1 (0.93%) |
| | Range | 8.03–8.76 | 7.03–7.87 | 6.00–6.87 | 5.00–5.97 | 4.21–4.93 | 3.97 |
| | SDI range | 0.12–0.59 | 0.29–0.67 | 0.53–0.75 | 0.66–0.79 | 0.74–0.79 | 0.79 |
| Fearful (<i>n</i> = 67; 45.58%) | <i>n</i> (%) | 6 (8.95%) | 12 (17.91%) | 18 (26.86%) | 17 (25.37%) | 11 (16.42%) | 3 (4.78%) |
| | Range | 8.00–8.40 | 7.00–7.97 | 6.00–6.80 | 5.10–5.90 | 4.30–4.97 | 3.43–3.97 |
| | SDI range | 0.30–0.58 | 0.47–0.71 | 0.48–0.75 | 0.66–0.78 | 0.70–0.79 | 0.76–0.79 |
| Neutral (<i>n</i> = 107; 21.53%) | <i>n</i> (%) | 1 (0.93%) | 3 (2.80%) | 24 (22.43%) | 41 (38.32%) | 36 (33.64%) | 2 (1.87%) |
| | Range | 8.47 | 7.00–7.23 | 6.00–6.97 | 5.00–5.97 | 4.00–4.97 | 3.90–3.90 |
| | SDI range | 0.40 | 0.54–0.66 | 0.53–0.75 | 0.60–0.78 | 0.73–0.79 | 0.75–0.79 |

Abbreviation: SDI, Simpson diversity index.

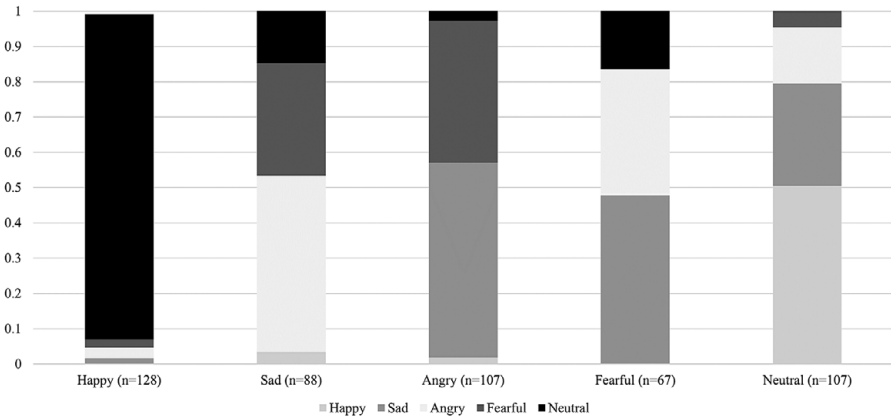


Fig. 1. Proportion of words in each category identified as second highest degree of fit.

degree-of-fit ratings of 8.00 or higher, five were moderate in diversity with scores ranging from 0.48 (*grief*) to 0.59 (*depressed*); one (*suicidal*) had an SDIS (0.66) indicating high diversity. For forty-five (51.14%) of the 88 words, Angry was the category with the second highest degree of fit. Degree-of-fit ratings for Angry ranged from 3.37 (*crest fallen*) to 6.45 (*disappointed*).

2.5.3. Angry

Ten (9.34%) of the 107 words best categorised as Angry received degree-of-fit ratings higher than 8.00, ranging from 8.03 (*ticked*) to 8.76 (*rage*). *Snickering* had the lowest degree-of-fit score (3.97) in the category. Interestingly, *angry* received only the third highest degree-of-fit rating (8.60), though it had the lowest SDIS (0.12). Overall, the SDIS for the 10 words ranged from 0.12 to 0.59 (*murderous*), with five of the 10 words receiving scores indicative of moderate diversity. Sad was the category most identified as having the second best degree of fit, with ratings ranging from 2.83 (*snarky*) to 7.20 (*betrayed*) across a total of 59 (55.14%) words.

2.5.4. Fearful

Only 67 (45.58%) words were identified as Fearful, the lowest proportion of words across the five categories. One of these words (*sly*) had the lowest degree-of-fit rating in any category (3.43). Six words (8.95%) had degree-of-fit ratings of 8.00 (*horrified*) or higher, with *frightened* rated highest (8.40), and *fearful* rated as third highest (8.28). Only *frightened* had a low diversity of ratings across emotion categories as indicated by its 0.30 SDIS. The remaining five words rated as 8.00 or higher were moderate in diversity, with scores ranging from 0.45 (*scared*) to 0.57 (*petrified*). For most of these words, Sad was the next best fit ($n = 32$; 47.76%), with ratings ranging from 3.27 (*hesitant*) to 6.72 (*traumatised*).

2.5.5. Neutral

One hundred and seven words (21.53%) were identified as Neutral. Overall, degree-of-fit ratings for Neutral were lower than for other categories, with the majority of

words ($n = 79$; 73.83%) rated as 5.97 (*unconcerned*) or below. Only *neutral* received a degree-of-fit rating over 8.00 ($M = 8.47$). However, the SDIS for *neutral* (0.40) just met the threshold for low diversity. This was the lowest SDIS in this category. Three additional words were classified as moderate diversity with scores of 0.53 (*bored*), 0.54 (*monotone*), and 0.60 (*cool*). The remaining 103 words (96.26%) had high SDISs that ranged from 0.61 (*blank*) to 0.79 (e.g. *nonplussed*; *pensive*).

2.6. Discussion

The aim of Study 1 was to explore the extent to which a diverse array of emotion words fit within traditional basic emotion categories. Participants characterised a set of emotion words in a categorisation task that required they rate the degree to which each word fit five basic emotion categories. According to SDISs, many words received degree-of-fit ratings that were dispersed across multiple categories. Not surprisingly, given that participants had several negatively valenced labels to consider, this dispersion was particularly evident for words categorised into one of the three negatively-valenced categories.

The Happy category included the largest number of words ($n = 128$), most likely because it was the only positive response option. Because of the limited options, any positively-valenced word was likely to be categorised as Happy, Neutral being the only other feasible alternative. In fact, Neutral was consistently identified as having the second best degree of fit for Happy words. If one or more additional positively-valenced emotion categories had been provided as response options, the proportion of words discretely categorised as Happy may have been reduced and participants may not have rated Neutral as highly. For example, *earnest* had the same degree-of-fit score for both Happy and Neutral (5.10) suggesting that participants identified this word as one with positive valence, but that Happy was not necessarily a good fit. Words such as *carefree*, *content*, and *satisfied* also had similar Happy and Neutral ratings.

Negatively-valenced emotions are reported to frequently co-occur with one another (Samson et al., 2016) and this appeared to be the case in Study 1, particularly evident for words identified as representing Sad or Angry. For example, *disappointed* had degree-of-fit ratings of 7.52 for Sad and 6.45 for Angry. Neutral was rarely identified as the second best degree of fit for words that were categorised into one of the three negatively-valenced categories. In fact, Neutral was only identified as second-best degree of fit for 18 of the 262 words (7%) categorised as Sad, Angry, or Fearful. Overall, results suggest that valence may be an important factor in structuring the emotional lexicon (Kuperman et al., 2014). This is consistent with research in lexical decision-making that showed valence has a strong effect on word processing. Results also suggest that when provided with several feasible response options (e.g. Angry, Fearful, Sad) there is diversity in how people categorise emotion words, at least in the case of those that are negatively-valenced.

A large number of words were categorised as Neutral in Study 1, but overall degree-of-fit ratings were lower than for other categories. In fact, only one word (*neutral*) had a degree-of-fit rating of 8.0 or higher. It is possible that participants do not routinely consider Neutral as representing an affective state and therefore tended to rate this category lower overall. For example, the degree-of-fit rating for *emotionless* was relatively low (5.45), despite the fact the word inherently suggests neutrality.

The overall low ratings for words categorised as Neutral seemed to be because Neutral was perceived as a 'catch all' category. Participants appeared to use this category for words that did not clearly signal an emotional state (e.g. *hungry; quiet; reading*), that tended to be better described as actions or behaviours (e.g. *hurried, preoccupied*) (Goddard, 2014), and those that described personality traits (e.g. *reserved; tolerant*) (Nettle & Penke, 2010).

In sum, results of Study 1 support our hypothesis that people's conceptualisations of emotion words may be quite nuanced, as indicated by the number of words (particularly negatively-valenced ones) with high SDISs. We conclude that a small-set of basic emotion categories is insufficient in capturing the degree of distinction with which people conceptualise emotion words. The use of Neutral as a 'catch all' category, particularly for Happy, is consistent with this conclusion. The lack of positively-valenced options in Study 1 means we do not know the extent to which the pattern observed in the negatively-valenced words is present across the full spectrum of emotion words.

3. Study 2

Study 2 was conducted to explore the influence of a broader set of response options on categorisation of a subset of emotion words from Study 1. We expanded the array of positive response options in order to explore whether the overlapping rating pattern observed for negatively valenced emotions in Study 1 might also occur for positively-valenced emotions. We also included complex emotions to explore whether a broader set of response options would provide a more nuanced picture of how people characterise emotion words. Due to public health restrictions in place at the time, Study 2, unlike Study 1, was conducted online.

3.1. Participants

Participants were recruited online using snowball sampling on the authors' social media platforms. We required that participants be at least 18 years of age and speak English as a primary language. All participants provided informed consent. A total of 113 participants began the study; however, six participants exited the survey before completing the demographic questions. Of the remaining 107 participants, only 56 (7 males; 49 females) proceeded beyond the demographic questions to the ratings. These 56 participants ranged in age from 19 to 72 years ($M = 47.5$; $SD = 12.63$) and had completed 16.8 ($SD = 3.12$) years of education on average. Thirty-four participants were from Canada, 12 from Australia, six from the United States, and four from the United Kingdom.

3.2. Stimuli

In selecting stimuli to be used in Study 2, we did not rely solely on degree of fit or SDISs from Study 1. We made this decision because it appeared that limited response options may have influenced Study 1 ratings. This was reflected in the data in two ways. First, review of individual participant data showed that for some words (e.g. *astonished, apprehensive, eager, miserable, sympathetic*) no categories were identified as a good fit. In these cases, participants either selected '1' for all five

categories or selected '1' for all but one category and gave the fifth category a low rating. Second, the SDISs for words in the three negatively-valenced categories (Sad, Angry, Fearful) were moderate to high, with very few words receiving low scores. It is unclear whether the higher overall diversity was because words that fell into these categories could be interpreted as any one of these three negatively-valenced emotions, or whether participants identified the word as negative, but did not feel that any of these three basic categories accurately fit the word. For example, *guilty* was identified as Sad with a degree-of-fit rating of 6.45. However, Angry and Fearful had similarly high ratings of 5.34 and 6.03, respectively, resulting in an SDIS of 0.68.

To reduce the length of task and minimise the likelihood of cognitive fatigue, a particular risk in online studies (Saleh & Bista, 2017), we opted to decrease the size of stimulus set in Study 2. We carefully reviewed the Study 1 words and excluded the following types of words: those that were better characterised as relating to appearance or character than emotion (e.g. *ugly, goofy, silly, dumb*), those considered slang (e.g. *snarky, miffed*), phrases (e.g. *creeped out, put upon*), and words that matched the category names of the response options (e.g. *happy, angry*). Next, all three authors independently reviewed the word list for each emotion category, and identified words that they would use in the sentence 'I feel [WORD]'. Only words identified by all three authors as fitting within that sentence frame were considered in the next step of the selection process. As was the case in Study 1, the total number of words remaining in each category varied widely, but Happy still included the highest total number ($n = 23$). Table 2 lists the words identified from Study 1 for possible use in Study 2, including mean degree-of-fit ratings for the five emotion categories and SDISs.

Next, the word list for each emotion category was analysed for word frequency to ensure, inasmuch as possible, that participants would be familiar with the words. Since Study 2 would be conducted online and include participants with English as a first language across different countries, we calculated word frequency using the average rating between the Corpus of Contemporary American English (Davies, 2008) and the British National Corpus (Davies, 2004) (see Table 2). We then selected 30% of the most frequently used words in each emotion category. This yielded a stimulus set of 24 words (identified in Table 2 using an *) distributed as follows: Happy (7); Sad (7); Angry (3); Fearful (6); Neutral (1).

3.3. Procedure

The study was conducted online using the Survey Monkey platform (<http://www.surveymonkey.com>). After viewing study details and providing informed consent, participants confirmed they met inclusion criteria and then answered demographic questions regarding gender, age, country of residence, and years of education. Next, they were presented with a word and asked to rate how well the word fit a number of different categories (see the Appendix), using a scale of 1 to 9 (1 = not at all; 9 = extremely well). Response options were ordered alphabetically and remained consistent for each stimulus.

Thirteen categories were provided as response options. Included were an equal number of basic and complex emotions, and an equal number of positive and negative valenced emotions. Neutral was also retained as a response option. To extend the basic emotion categories to six, we opted to include Disgust and Joy as response options. Disgust is well known as a basic emotion category (Ekman, 1992).

Table 2. Mean rating and Simpson diversity index for words from study 1 considered for inclusion in study 2

| Word (n) | Happy M(SD) | Sad M(SD) | Angry M(SD) | Fear M(SD) | Neutral M(SD) | Simpson diversity | Word frequency score |
|----------------------------|-------------------|-------------------|-------------|------------|---------------|-------------------|----------------------|
| Ecstatic (29) | 8.31 (1.5) | 1.24 (0.6) | 1.59 (1.6) | 1.59 (1.5) | 1.90 (1.9) | 0.32 | 1,458 |
| Cheerful (30) | 8.27 (1.0) | 1.43 (1.1) | 1.37 (1.3) | 1.20 (0.5) | 2.50 (1.9) | 0.12 | 2,918.5 |
| Delighted (29) | 8.24 (1.0) | 1.55 (1.4) | 1.55 (1.4) | 1.55 (1.6) | 2.76 (2.2) | 0.18 | 5,495.5 |
| * <i>Excited</i> (29) | 8.07 (2.0) | 1.41 (1.1) | 1.38 (1.0) | 1.59 (1.6) | 2.28 (1.8) | 0.36 | 20,870 |
| Enthusiastic (29) | 7.97 (1.3) | 1.52 (1.2) | 1.52 (1.3) | 1.48 (1.4) | 2.79 (2.4) | 0.37 | 4,742.5 |
| Thrilled (30) | 7.90 (1.7) | 1.80 (1.7) | 1.87 (1.8) | 2.10 (2.1) | 2.40 (1.8) | 0.42 | 5,295 |
| * <i>Confident</i> (29) | 7.86 (1.7) | 1.28 (0.5) | 1.24 (0.4) | 1.41 (1.2) | 2.86 (2.3) | 0.19 | 14,628 |
| Upbeat (29) | 7.86 (1.5) | 1.45 (1.2) | 1.38 (1.3) | 1.45 (1.4) | 2.66 (2.1) | 0.35 | 1,706.5 |
| * <i>Peaceful</i> (30) | 7.70 (1.4) | 2.07 (1.5) | 1.40 (1.0) | 1.60 (1.2) | 4.70 (2.5) | 0.36 | 9,267.5 |
| Exhilarated (29) | 7.69 (1.7) | 1.59 (1.0) | 1.72 (1.3) | 1.69 (1.5) | 2.21 (2.0) | 0.31 | 316 |
| Hopeful (29) | 7.52 (1.2) | 2.38 (2.4) | 1.45 (0.9) | 3.17 (2.6) | 2.66 (2.1) | 0.42 | 5,495 |
| * <i>Glad</i> (30) | 7.50 (1.5) | 1.80 (1.6) | 1.50 (0.9) | 1.53 (1.0) | 2.77 (1.9) | 0.38 | 31,692 |
| Thankful (30) | 7.50 (1.6) | 2.10 (2.1) | 1.90 (2.1) | 1.90 (2.2) | 3.90 (2.6) | 0.36 | 4,417.5 |
| Optimistic (29) | 7.45 (2.2) | 1.62 (1.1) | 1.48 (1.1) | 1.83 (1.7) | 2.83 (2.4) | 0.18 | 7,120 |
| Relaxed (30) | 7.23 (1.9) | 1.67 (1.0) | 1.43 (1.0) | 1.23 (0.6) | 5.53 (2.9) | 0.52 | 7,084.5 |
| * <i>Pleased</i> (30) | 7.20 (2.0) | 2.00 (1.7) | 1.63 (1.3) | 1.57 (1.1) | 3.53 (2.2) | 0.50 | 13,926.5 |
| * <i>Satisfied</i> (30) | 6.87 (1.4) | 2.33 (1.6) | 1.87 (1.6) | 1.80 (1.3) | 4.63 (2.6) | 0.59 | 10,953 |
| Elated (30) | 6.80 (2.5) | 2.47 (2.1) | 2.10 (1.5) | 2.10 (1.5) | 3.63 (2.7) | 0.70 | 731 |
| Carefree (30) | 6.40 (2.1) | 2.23 (1.9) | 1.80 (1.5) | 2.13 (1.9) | 5.30 (2.8) | 0.62 | 901 |
| * <i>Eager</i> (30) | 6.40 (1.9) | 2.17 (1.7) | 1.90 (1.3) | 2.20 (1.8) | 3.33 (2.1) | 0.56 | 9,975 |
| Awestruck (30) | 6.30 (1.8) | 2.67 (1.7) | 2.60 (2.1) | 2.73 (2.3) | 3.97 (2.7) | 0.69 | 290 |
| Bemused (30) | 5.07 (1.6) | 3.07 (1.9) | 3.07 (2.0) | 2.67 (1.9) | 4.40 (2.0) | 0.78 | 687.5 |
| Astonished (30) | 4.67 (1.9) | 3.80 (1.9) | 3.63 (2.3) | 2.90 (1.7) | 3.67 (2.4) | 0.78 | 2,093 |
| Heartbroken (29) | 1.14 (0.4) | 8.66 (0.9) | 6.00 (2.7) | 4.59 (3.1) | 1.83 (1.8) | 0.54 | 836 |
| Depressed (29) | 1.31 (0.8) | 8.10 (2.1) | 5.76 (2.5) | 5.93 (2.6) | 1.93 (2.0) | 0.59 | 7,511.5 |
| Devastated (30) | 1.40 (1.2) | 8.03 (1.7) | 5.77 (2.8) | 5.43 (2.9) | 1.50 (1.1) | 0.54 | 3,730 |
| * <i>Upset</i> (30) | 1.40 (0.9) | 7.67 (1.8) | 5.50 (2.8) | 4.40 (2.6) | 2.37 (2.1) | 0.65 | 20,427.5 |
| * <i>Disappointed</i> (59) | 1.40 (0.7) | 7.44 (1.8) | 6.14 (2.3) | 3.87 (5.4) | 2.23 (1.7) | 0.56 | 11,628 |
| * <i>Hurt</i> (30) | 1.63 (1.1) | 7.37 (1.5) | 6.17 (2.3) | 4.70 (2.7) | 2.37 (1.7) | 0.65 | 48,392 |
| Lonely (30) | 1.73 (1.5) | 7.37 (2.2) | 3.87 (2.7) | 4.77 (2.8) | 2.93 (2.5) | 0.63 | 9,357 |
| Miserable (30) | 1.40 (1.2) | 7.37 (2.2) | 5.73 (2.7) | 4.63 (2.9) | 1.67 (1.5) | 0.67 | 6,125 |
| Disheartened (29) | 1.52 (0.9) | 7.34 (1.5) | 4.48 (2.7) | 3.83 (2.4) | 2.59 (2.0) | 0.51 | 315 |
| Hopeless (29) | 1.59 (1.1) | 7.10 (2.3) | 4.69 (2.6) | 6.59 (2.4) | 2.17 (1.6) | 0.66 | 3,051 |

(Continued)

Table 2. (Continued)

| Word (n) | Happy M(SD) | Sad M(SD) | Angry M(SD) | Fear M(SD) | Neutral M(SD) | Simpson diversity | Word frequency score |
|--------------------|-------------|-------------------|-------------------|-------------------|---------------|-------------------|----------------------|
| Remorseful (30) | 2.30 (2.1) | 6.93 (2.2) | 4.23 (2.8) | 3.97 (2.8) | 2.67 (2.2) | 0.63 | 271.5 |
| *Down (30) | 1.70 (1.2) | 6.77 (2.9) | 3.57 (2.5) | 3.93 (2.7) | 2.97 (2.2) | 0.61 | 507,205.5 |
| Helpless (29) | 1.48 (1.0) | 6.72 (2.3) | 4.45 (2.7) | 6.28 (2.5) | 1.97 (1.5) | 0.71 | 3,831 |
| Defeated (30) | 1.80 (1.4) | 6.57 (2.5) | 5.20 (2.6) | 4.90 (2.7) | 2.37 (1.7) | 0.75 | 7,950 |
| *Guilty (29) | 1.38 0.(6) | 6.45 (2.3) | 5.34 (2.4) | 6.03 (2.8) | 2.52 (2.3) | 0.68 | 24,726.5 |
| *Desperate (29) | 1.66 (1.1) | 6.21 (1.6) | 4.55 (2.3) | 5.83 (2.5) | 2.72 (1.7) | 0.70 | 12,820.5 |
| Anguished (30) | 1.73 (1.1) | 5.70 (2.2) | 5.50 (2.3) | 4.87 (2.7) | 2.40 (1.5) | 0.73 | 622.5 |
| *Overwhelmed (29) | 2.10 (1.6) | 5.62 (2.6) | 5.00 (2.2) | 5.21 (2.7) | 2.83 (2.5) | 0.75 | 5,252.5 |
| Melancholic (29) | 3.52 (2.1) | 5.52 (2.4) | 3.62 (1.9) | 3.41 (1.9) | 4.00 (2.3) | 0.79 | 231.5 |
| Dismayed (30) | 1.80 (1.1) | 5.20 (2.3) | 4.97 (2.0) | 4.30 (2.3) | 3.27 (2.0) | 0.74 | 1,069 |
| Sympathetic (30) | 4.70 (2.6) | 4.73 (2.7) | 2.33 (1.9) | 2.43 (2.0) | 3.33 (2.3) | 0.76 | 5,233 |
| Disillusioned (30) | 1.93 (1.3) | 4.63 (2.4) | 4.43 (2.5) | 4.33 (2.4) | 3.57 (2.3) | 0.77 | 1,037 |
| Outraged (29) | 1.55 (1.6) | 4.31 (3.0) | 8.55 (1.0) | 3.79 (2.9) | 1.76 (1.4) | 0.46 | 2,972.5 |
| Enraged (30) | 1.43 (1.1) | 3.60 (2.6) | 8.37 (1.2) | 3.63 (2.9) | 1.73 (1.4) | 0.41 | 1,381 |
| Furious (30) | 1.50 (1.3) | 3.73 (2.6) | 8.17 (2.0) | 3.87 (2.6) | 1.73 (1.3) | 0.53 | 5,567.5 |
| *Mad (29) | 1.38 (0.9) | 3.90 (2.4) | 8.10 (1.7) | 4.03 (2.6) | 1.86 (1.4) | 0.44 | 21,392.5 |
| *Frustrated (30) | 1.60 (1.3) | 4.80 (2.5) | 7.77 (1.7) | 4.00 (2.5) | 2.27 (1.8) | 0.29 | 8,424.5 |
| Agitated (30) | 1.50 (1.5) | 4.20 (2.8) | 7.50 (1.6) | 4.27 (2.6) | 2.20 (1.5) | 0.46 | 1,715.5 |
| Annoyed (29) | 1.83 (1.5) | 4.00 (2.5) | 7.21 (1.4) | 3.28 (2.5) | 3.28 (2.4) | 0.52 | 3,864 |
| Appalled (30) | 1.93 (1.5) | 4.13 (2.3) | 6.57 (2.1) | 3.90 (2.6) | 2.50 (1.6) | 0.65 | 293.5 |
| *Tense (29) | 1.93 (1.5) | 3.79 (2.4) | 5.90 (2.4) | 5.52 (2.9) | 3.03 (2.2) | 0.71 | 5,508 |
| Exasperated (30) | 3.10 (2.4) | 4.23 (2.4) | 4.90 (2.6) | 3.23 (2.3) | 3.47 (2.1) | 0.78 | 1,043 |
| Petrified (30) | 1.50 (1.1) | 4.00 (2.9) | 4.23 (3.0) | 8.37 (1.5) | 2.37 (2.2) | 0.57 | 871.5 |
| *Scared (30) | 1.53 (1.3) | 4.17 (2.8) | 3.13 (2.5) | 8.27 (1.2) | 2.50 (1.9) | 0.45 | 22,724.5 |
| Horrified (30) | 1.40 (1.2) | 4.53 (2.8) | 5.20 (2.8) | 8.00 (1.7) | 1.77 (1.7) | 0.50 | 2,600.5 |
| Threatened (30) | 1.50 (1.3) | 4.47 (2.7) | 4.97 (2.8) | 8.03 (1.3) | 2.23 (1.6) | 0.58 | 15,732 |
| *Afraid (30) | 1.33 (0.8) | 4.23 (2.7) | 3.33 (2.2) | 7.73 (2.2) | 2.33 (1.7) | 0.47 | 42,234 |
| Uneasy (30) | 1.57 (0.9) | 4.63 (2.4) | 4.20 (2.4) | 7.70 (1.4) | 2.80 (1.9) | 0.48 | 3,042.5 |
| Terrified (29) | 1.83 (1.9) | 5.69 (2.5) | 4.72 (2.8) | 7.69 (2.2) | 1.90 (1.5) | 0.65 | 5,761.5 |
| Distressed (30) | 1.87 (1.9) | 6.03 (2.6) | 5.43 (2.6) | 7.17 (2.1) | 1.97 (1.4) | 0.65 | 2,010 |
| *Worried (30) | 1.80 (1.6) | 4.87 (2.6) | 3.60 (2.3) | 7.10 (1.9) | 2.80 (1.2) | 0.65 | 30,027 |
| *Nervous (30) | 2.17 (1.5) | 4.40 (2.4) | 3.43 (2.3) | 6.63 (2.4) | 2.87 (2.3) | 0.70 | 18,318.5 |
| Uncomfortable (30) | 1.83 (1.1) | 4.87 (2.5) | 5.13 (2.6) | 6.60 (2.4) | 2.67 (1.9) | 0.63 | 10,778.5 |

(Continued)

Table 2. (Continued)

| Word (<i>n</i>) | Happy M(SD) | Sad M(SD) | Angry M(SD) | Fear M(SD) | Neutral M(SD) | Simpson diversity | Word frequency score |
|-------------------------|-------------|------------|-------------|-------------------|-------------------|-------------------|----------------------|
| Distraught (29) | 1.66 (1.1) | 6.21 (2.3) | 5.41 (2.4) | 6.28 (2.3) | 2.17 (1.8) | 0.75 | 1,265 |
| * <i>Concerned</i> (30) | 2.23 (1.5) | 4.17 (2.6) | 3.50 (2.3) | 6.00 (2.4) | 3.40 (2.4) | 0.64 | 45,094.5 |
| Uncertain (30) | 1.93 (1.1) | 4.10 (2.4) | 3.57 (2.4) | 5.60 (2.7) | 3.43 (2.3) | 0.70 | 6,908 |
| Embarrassed (30) | 1.97 (1.6) | 5.13 (2.7) | 4.40 (2.6) | 5.50 (2.5) | 2.63 (2.0) | 0.72 | 8,058 |
| * <i>Shocked</i> (30) | 3.67 (2.3) | 3.73 (2.2) | 4.80 (2.4) | 5.37 (2.4) | 2.80 (2.2) | 0.78 | 10,978 |
| Doubtful (30) | 1.87 (1.3) | 5.10 (2.0) | 3.77 (2.6) | 5.27 (3.2) | 3.37 (2.3) | 0.66 | 2,569 |
| Apprehensive (30) | 2.97 (2.2) | 2.90 (1.8) | 2.47 (1.8) | 5.13 (2.7) | 3.17 (2.0) | 0.75 | 842.5 |
| Confused (30) | 1.97 (1.2) | 4.37 (2.5) | 4.33 (2.8) | 4.67 (2.6) | 3.67 (2.8) | 0.77 | 14,366.5 |
| Indifferent (30) | 2.90 (2.1) | 3.37 (2.3) | 2.97 (2.2) | 2.47 (1.8) | 6.97 (2.6) | 0.63 | 2,340.5 |
| * <i>Calm</i> (30) | 5.90 (2.3) | 2.70 (1.9) | 1.60 (1.1) | 1.93 (1.5) | 6.77 (2.4) | 0.62 | 20,311.5 |
| Bored (30) | 1.67 (1.1) | 3.40 (2.4) | 2.37 (1.9) | 1.77 (1.1) | 6.33 (2.8) | 0.53 | 7,034.5 |

Joy has also been identified as basic and although this category is understood to contain Happy, people appear to differentiate these words by intensity (Zupan & Eskritt, 2020). Surprise was not included because it cannot be clearly identified as positive or negative in valence (Schlegel et al., 2012; Zupan & Eskritt, 2020). The final six emotion categories represented complex emotions – four that were positively-valenced (Amusement, Contentment, Pride, Relief) and two that were negatively-valenced (Anxiety, Irritated).

To maximise participant engagement, Revilla and Ochoa (2017) report that the ideal length of an online study should be less than 20 minutes. In an effort to meet this target, participants were informed that they could rate as many of the categories as they wished for each word, but if they felt a word did not fit into a category, they could leave the rating blank and a rating of '1' would be assumed. However, a minimum of three ratings was required for each word to increase the likelihood that participants were fully reflecting on the degree of fit across the range of response options. The study was expected to take approximately 15 minutes to complete.

3.4. Data analysis

The same analyses used in Study 1 were applied to Study 2.

3.5. Results

Table 3 lists the mean degree-of-fit rating for each emotion category for each of the Study 2 words, alongside the word's corresponding SDIS. Of the 24 words, only half ($n = 12$; 50%) were identified as best fitting one of the same five categories used in Study 1, with four words identified as Happy, four as Sad, one as Angry, and three as Fearful. No words were categorised as Neutral.

Words were categorised into nine of the 13 possible categories. The diversity of ratings was higher overall in Study 2; only three words (*nervous*, *afraid*, *scared*), were found to have a moderate SDIS. The remaining 21 words had a high diversity as indicated by SDISs ranging from 0.62 (*mad*) to 0.91 (*shocked*). This included the positively-valenced words, which had been found to have lower diversity overall in Study 1. A summary of results follows, organised by the emotion category in which the word was initially placed in Study 1.

3.5.1. Words identified as happy in study 1

Four of the seven words (57%) identified as Happy in Study 1, were also identified as Happy in Study 2. The SDIS was higher for all four words, ranging from 0.70 to 0.87 in Study 2 compared to 0.36 to 0.56 in Study 1. Degree-of-fit ratings for Neutral for all four of these words were low, ranging from 1.06 to 1.15. Instead, Joy was identified as the second best degree of fit for three words (*glad*, *excited*, *eager*) and Contentment for the fourth (*pleased*). The remaining three words identified as Happy in Study 1 were categorised as complex, positively-valenced emotion in Study 2; two words (*satisfied*, *peaceful*) were identified as Contentment and the third word (*confident*) was identified as Pride. Happy was identified as the second-best degree of fit for all three of these words.

Table 3. Mean degree-of-fit ratings and Simpson diversity index for each word in study 2

| Word | Amuse-ment | Angry | Anxiety | Content-ment | Disgust | Fearful | Happy | Irritated | Joy | Neutral | Pride | Relief | Sad | Simpson diversity |
|--------------|------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------------|------------|-------------------|------------|-------------------|-------------------|
| Mad | 1.09 (0.4) | 8.69 (0.7) | 2.89 (2.5) | 1.09 (0.6) | 5.28 (2.9) | 2.63 (2.1) | 1.11 (0.5) | 7.28 (2.1) | 1.15 (1.0) | 1.28 (0.8) | 1.46 (1.6) | 1.13 (0.6) | 3.07 (2.5) | 0.62 |
| Worried | 1.11 (0.8) | 2.22 (2.0) | 8.41 (1.2) | 1.17 (0.8) | 1.91 (2.0) | 7.41 (2.1) | 1.15 (1.0) | 3.31 (2.5) | 1.15 (0.9) | 1.50 (1.2) | 1.41 (1.5) | 1.74 (2.1) | 3.52 (2.7) | 0.63 |
| Nervous | 1.69 (1.9) | 1.96 (1.9) | 8.26 (1.7) | 1.20 (1.1) | 1.54 (1.4) | 6.93 (2.3) | 1.44 (1.4) | 2.70 (2.3) | 1.30 (1.3) | 1.37 (1.3) | 1.30 (1.2) | 1.50 (1.5) | 2.52 (2.3) | 0.60 |
| Overwhelmed | 1.52 (1.7) | 3.65 (3.1) | 7.94 (1.8) | 1.28 (1.2) | 2.20 (2.3) | 5.09 (3.1) | 1.76 (2.2) | 4.02 (3.0) | 1.81 (2.3) | 1.33 (1.1) | 1.67 (2.0) | 1.91 (2.2) | 4.43 (3.1) | 0.78 |
| Tense | 1.31 (1.1) | 5.04 (3.3) | 7.50 (2.1) | 1.15 (0.9) | 2.61 (2.4) | 6.11 (2.8) | 1.04 (0.2) | 6.19 (2.4) | 1.09 (0.4) | 1.15 (0.6) | 1.26 (1.2) | 1.63 (1.8) | 2.56 (2.3) | 0.74 |
| Desperate | 1.19 (0.9) | 2.81 (2.3) | 6.76 (2.5) | 1.17 (0.8) | 2.00 (1.9) | 5.57 (2.8) | 1.07 (0.4) | 3.46 (2.7) | 1.02 (0.1) | 1.22 (0.9) | 1.46 (1.6) | 1.85 (2.0) | 3.83 (3.1) | 0.81 |
| Guilty | 1.11 (0.5) | 3.00 (2.5) | 5.54 (3.0) | 1.41 (1.5) | 3.33 (2.9) | 4.98 (2.9) | 1.24 (1.1) | 2.56 (2.1) | 1.13 (0.6) | 1.28 (0.8) | 1.48 (1.4) | 1.63 (1.6) | 3.63 (3.0) | 0.86 |
| Satisfied | 2.91 (2.7) | 1.15 (0.7) | 1.31 (1.2) | 8.70 (0.7) | 1.19 (0.9) | 1.11 (0.5) | 7.24 (2.2) | 1.19 (0.9) | 5.54 (2.9) | 3.57 (3.0) | 4.02 (3.0) | 4.59 (2.9) | 1.07 (0.3) | 0.73 |
| Peaceful | 1.94 (1.9) | 1.02 (0.1) | 1.22 (0.9) | 8.30 (1.3) | 1.09 (0.4) | 1.11 (0.6) | 6.52 (2.4) | 1.06 (0.3) | 5.80 (2.8) | 4.39 (3.0) | 2.43 (2.2) | 5.26 (2.8) | 1.15 (0.8) | 0.70 |
| Calm | 1.89 (1.6) | 1.24 (1.1) | 1.50 (1.8) | 7.31 (2.6) | 1.04 (0.2) | 1.22 (1.1) | 5.06 (2.9) | 1.06 (0.2) | 4.09 (2.8) | 5.93 (3.1) | 2.07 (2.1) | 4.56 (3.1) | 1.46 (1.4) | 0.75 |
| Shocked | 2.46 (2.3) | 3.30 (2.6) | 3.15 (2.6) | 1.28 (0.9) | 3.81 (3.1) | 3.48 (2.5) | 2.30 (2.1) | 2.70 (2.4) | 1.96 (1.9) | 1.57 (1.7) | 1.33 (1.3) | 1.98 (2.0) | 2.67 (2.6) | 0.91 |
| Afraid | 1.41 (1.3) | 2.61 (2.4) | 7.67 (1.7) | 1.19 (1.1) | 2.09 (2.0) | 8.57 (1.6) | 1.06 (0.4) | 2.76 (2.1) | 1.07 (0.5) | 1.26 (1.0) | 1.22 (1.1) | 1.61 (1.8) | 2.76 (2.4) | 0.54 |
| Scared | 1.69 (1.8) | 2.91 (2.4) | 7.76 (1.6) | 1.31 (1.3) | 2.02 (2.0) | 8.50 (1.7) | 1.13 (0.6) | 2.67 (2.0) | 1.09 (0.4) | 1.28 (0.9) | 1.41 (1.5) | 1.72 (1.9) | 3.04 (2.5) | 0.57 |
| Concerned | 1.09 (0.7) | 2.28 (1.9) | 6.11 (2.5) | 1.28 (1.1) | 1.98 (2.1) | 6.11 (2.3) | 1.15 (1.1) | 2.93 (2.3) | 1.13 (1.0) | 1.43 (1.3) | 1.24 (1.3) | 1.76 (2.0) | 3.37 (2.5) | 0.70 |
| Glad | 3.57 (2.6) | 1.00 (0.0) | 1.13 (0.7) | 5.28 (3.0) | 1.06 (0.2) | 1.17 (0.9) | 8.07 (2.0) | 1.07 (0.3) | 7.39 (2.1) | 2.09 (2.0) | 3.70 (2.9) | 3.98 (3.2) | 1.06 (0.3) | 0.70 |
| Excited | 6.06 (2.8) | 2.09 (2.3) | 2.20 (2.1) | 3.02 (2.6) | 1.17 (.5) | 1.56 (1.5) | 7.48 (1.5) | 1.70 (1.7) | 6.94 (2.3) | 1.52 (1.4) | 3.31 (2.6) | 2.24 (2.4) | 1.11 (.5) | 0.76 |
| Pleased | 4.59 (2.9) | 1.13 (0.6) | 1.22 (1.0) | 7.07 (2.2) | 1.19 (0.9) | 1.30 (1.4) | 7.24 (2.7) | 1.20 (0.9) | 7.02 (2.4) | 2.39 (2.2) | 5.28 (3.3) | 4.15 (2.9) | 1.11 (0.6) | 0.80 |
| Eager | 2.81 (2.4) | 1.24 (0.9) | 2.63 (2.3) | 2.41 (2.3) | 1.19 (0.8) | 1.44 (1.3) | 4.72 (2.8) | 1.30 (1.3) | 4.67 (3.0) | 2.07 (2.0) | 2.81 (2.6) | 2.00 (2.2) | 1.15 (0.6) | 0.87 |
| Frustrated | 1.11 (0.6) | 6.61 (2.2) | 4.72 (2.8) | 1.02 (0.1) | 3.33 (2.8) | 2.76 (2.5) | 1.02 (0.1) | 7.89 (1.8) | 1.04 (0.2) | 1.26 (0.9) | 1.31 (1.2) | 1.31 (1.3) | 3.11 (2.5) | 0.65 |
| Confident | 1.72 (1.6) | 1.26 (0.9) | 1.39 (1.4) | 5.74 (2.7) | 1.31 (1.0) | 1.39 (1.1) | 5.85 (2.5) | 1.28 (0.9) | 3.93 (2.7) | 2.46 (2.2) | 6.76 (2.7) | 2.33 (2.3) | 1.13 (0.6) | 0.73 |
| Down | 1.22 (1.0) | 2.59 (2.2) | 5.11 (2.9) | 1.50 (1.6) | 2.07 (1.9) | 3.72 (2.9) | 1.30 (1.4) | 2.41 (2.0) | 1.20 (1.1) | 1.83 (1.7) | 1.33 (1.4) | 1.39 (1.6) | 8.35 (1.8) | 0.64 |
| Hurt | 1.15 (1.0) | 5.65 (2.6) | 4.15 (3.0) | 1.09 (0.5) | 3.17 (2.4) | 3.41 (2.7) | 1.11 (0.6) | 3.67 (2.5) | 1.13 (0.6) | 1.43 (1.1) | 2.02 (2.3) | 1.43 (1.5) | 7.6 (2.1) | 0.65 |
| Disappointed | 1.02 (0.1) | 3.81 (2.2) | 3.11 (2.5) | 1.11 (0.7) | 3.39 (2.6) | 1.78 (1.6) | 1.06 (0.3) | 4.67 (2.7) | 1.04 (0.2) | 1.48 (1.1) | 1.57 (1.7) | 1.57 (1.6) | 6.81 (2.4) | 0.67 |
| Upset | 1.17 (0.9) | 6.44 (2.4) | 4.80 (3.0) | 1.17 (1.0) | 4.02 (3.0) | 3.89 (2.8) | 1.07 (0.4) | 6.07 (2.6) | 1.07 (0.4) | 1.33 (1.0) | 1.35 (1.3) | 1.69 (1.8) | 6.61 (3.1) | 0.81 |

Note: The bold values highlight the highest mean category response for each word.

3.5.2. Words identified as sad in study 1

Four words (57%) were identified as Sad in both studies – *down*, *hurt*, *disappointed*, and *upset*. However, SDISs were high for all four words (Range = 0.64–0.81), including the word *disappointed*, which had only moderate diversity in Study 1. Angry was identified as the second best degree of fit for two of these four words (*hurt*, *upset*). The remaining words (*overwhelmed*, *desperate*, *guilty*) were all identified as Anxiety in Study 2; Fearful was the second best degree of fit.

3.5.3. Words identified as angry in study 1

Only the word *mad* was identified as Angry in both studies, with a degree-of-fit score of 8.10 in Study 1 and 8.69 in Study 2. The SDIS of this word increased from moderate diversity in Study 1 (0.44) to high diversity in Study 2 (0.62); Irritation was identified as the second best degree of fit. The remaining two words were identified as Irritated (*frustration*) and Anxiety (*tense*) in Study 2.

3.5.4. Words identified as fearful in study 1

Fearful was identified as the best degree of fit for three of the six words (50%) that had been categorised as Fearful in Study 1 – *scared*, *afraid*, and *concerned*. All three words had slightly higher degree-of-fit ratings for Fearful in Study 2 (Range = 6.11–8.57) compared to Study 1 (Range = 6.00–8.27). Diversity ratings were also similar across studies with moderate SDISs for *scared* (0.57) and *afraid* (0.54), and a high SDIS for *concerned* (0.70). Where ratings differed most was in the second best degree-of-fit category. In Study 1, participants chose Sad as the second-best category for all three words with ratings ranging from 4.17 to 4.47. However, in Study 2, Anxiety was identified as having the second best degree of fit with ratings similar to those for Fearful (Range = 6.11–7.76). In fact, *concerned* received the same mean degree-of-fit rating (6.11) for both Fearful and Anxiety (it was categorised as Fearful based on standard deviation). Two of the remaining three words identified as Fearful in Study 1 (*worried*, *nervous*) were categorised as Anxiety in Study 2; the third word, *shocked*, was categorised as Disgust.

3.5.5. Words identified as neutral in study 1

The word *calm*, which had been categorised as Neutral in Study 1, was instead identified as Contentment in Study 2 with a degree-of-fit rating of 7.32. Neutral was still identified as the second-best degree of fit (5.93). SDIS was high (0.75) due to moderate degree-of-fit ratings for Happy (5.06) and Relief (4.56), suggesting an overall positive valence for this word.

3.6. Discussion

Overall, the diversity of ratings was higher in Study 2 than in Study 1, ranging from 0.54 to 0.91 ($M = 0.70$) for negatively-valenced words, and 0.70 to 0.87 ($M = 0.75$) for positive ones. This suggests that when participants are given a broader set of options, categorisation becomes more dispersed regardless of whether the word is positive or negative in valence. While this dispersion in ratings may suggest that people categorise words differently to one another, it may also reflect that even though a broader set of options were provided, there was still no one ‘best’ fit for the word.

Future research might explore if further increasing the set of response options leads to greater or less dispersion in ratings to further explore this concept.

However, the final categorisation of words still remained limited to a small set of categories, particularly for positively-valenced emotions where words were only categorised into three of the possible eight categories – Contentment, Happy, Pride – but at least moderate ratings still occurred across multiple positive categories. For example, *pleased*, which was identified as Happy with a degree-of-fit rating of 7.24, attained ratings of four or higher (Range = 4.15–7.07) in five additional positively-valenced categories (Amusement, Contentment, Joy, Pride, Relief). This suggests that, as is the case with negative emotions, there may be overlap in how people perceive positive emotions. This was particularly notable across Happy, Contentment and Joy, which were identified as the best or second best degree of fit for seven of the eight words identified as positively-valenced in Study 2. Data suggest that these three emotion categories may form an emotion family, with all three words representing a similar, general feeling of positive well-being, differing only in the level of intensity portrayed (Crane & Gross, 2007; Diener et al., 1995; Ekman, 1992).

The availability of more complex, lower intensity emotions in Study 2 also appeared to influence participant responses for negatively-valenced words. For instance, *guilty*, *desperate*, and *overwhelmed* had all been categorised as Sad in Study 1 when participants were given only basic emotion categories to rate, with Fearful identified as the second best degree of fit. In Study 2, all three words were categorised as Anxiety, the lower intensity alternative within the Fearful emotion family; Fearful remained the second best degree of fit. This same pattern was observed for words categorised as Angry or Irritated. This pattern suggests that offering participants different intensity options can provide further insight into their conceptualisation of emotion.

Unlike in Study 1, participants only identified Neutral as the second best category for one of the positively-valenced emotion words – *calm* – a word that had been previously categorised as Neutral in Study 1. In Study 2, it was instead categorised as Contentment. Interestingly, the three words categorised as Contentment in Study 2 (*satisfied*, *peaceful*, *calm*) had the highest Neutral ratings (Range = 3.57–5.93) of all 24 words included in Study 2. Though not always labelled an emotion per se, Neutral is considered an affective state, either positive or negative, in which an individual feels nothing in particular (Gasper et al., 2019). In other words, Neutral does not necessarily mean the absence of emotion or feeling. However, it has also been suggested that simply recognising an affective state as Neutral can result in feelings of pleasure (Gasper et al., 2019). Providing participants with a positively valenced, low intensity option like Contentment may have enabled them to more accurately reflect their conceptualisation of words like *calm*. The purpose of the studies presented here was not to explore Neutral per se; however, in hindsight, we could have gained insight into this matter had we selected a subset of words identified as Neutral in Study 1 on the basis of valence ratings to explore whether participants would be more inclined to attach affect to these words if low-intensity descriptors were available.

4. Overall discussion

The overall aim of the two studies was to determine if emotion words categorised into a small set of basic emotion categories would retain that categorisation in the

presence of a broader set of response options. In Study 1, participants were asked to categorise emotions into one of four basic emotion categories (Happy, Sad, Angry, Fearful) or Neutral. In Study 2, these categories were extended to a total of 13 options. Of the 24 words included in Study 2, only half were categorised using the same basic emotion categories as Study 1. The remaining 12 words were instead categorised across five of the broader response options, four of which represented complex emotions.

Overall, results suggest that there is considerable diversity in how people conceptualise words and that basic emotion categories may not sufficiently represent the way people think about and define emotion. The provision of more options had an interesting impact on how people categorise emotion words. First, when participants were provided the opportunity to utilise complex emotion categories, they did so. In fact, nearly half of the words in Study 2 ($n = 11$) were categorised into one of the complex emotion categories. Second, when participants were provided a greater array of emotion categories from which to choose, they rated words differently, sometimes considerably so. In some cases, words in Study 2 were categorised in entirely different emotion families than Study 1. For example, *guilty*, which had been categorised as Sad in Study 1, was categorised as Anxiety in Study 2. This result was particularly surprising given that *guilt* has previously been identified as belonging to a family of related emotions that includes both anger and frustration (Roseman, 1994). Though further study is required, findings like these suggest that reliance on a narrow range of basic emotion categories may obfuscate our understanding of how people conceptualise emotion, such that we fail to appreciate the nuance of which they are capable.

We had expected that having categories that represented more complex and disparate (e.g. Amusement, Pride) and/or lower intensity emotions (e.g. Anxiety) would result in more distinct ratings in one category over others. This did not occur. Instead, it appeared that participants primarily differentiated emotion broadly on the basis of valence, and were less discerning within that valence. This was particularly evident in the case of positively-valenced emotions where the emotion categories were primarily complex in nature. It may be that participants have more difficulty differentiating complex emotion categories or that words that are best represented by complex emotions tend to have more overlap in meaning and appraisal. For instance, the word *pleased*, which was categorised as Happy by participants in both studies, had moderate to high ratings for six different emotion categories in Study 2 including Amusement, Pride, and Relief. Although these categories might generally be considered quite distinct from one another, words such as *pleased* may simply have multiple meanings depending on how they are appraised. For example, *pleased* could easily fit the category Pride if interpreted as pertaining to increased feelings of self-worth related to an accomplishment (Lewis, 2008). Similarly, *pleased* could fit the category Relief, describing a situation where an aversive situation has been resolved (Roseman, 1994).

4.1. Limitations and future directions

Findings must be considered in light of two types of limitations. The first involves choice of emotion words, both in regard to response options and the word lists. In regard to response options, there are a few notable absences. Though Surprise is one

of the six basic emotion categories, we opted not to include it, both because it is not clearly positive or negative and because we sought an equal number words of each valence. However, future studies might consider including Pleasant Surprise, which has been successfully used in studies of vocal emotion recognition (Dupuis & Pichora-Fuller, 2015), to make this distinction. We also did not include Compassion (considered a social emotion) or Contempt (a negative emotion, related to Anger and Disgust) (Ekman, 1992). Also, some of the words we had participants rate could be considered their own category, particularly if we look to more social emotions (e.g. *guilt*, *worry*). We acknowledge that inclusion of these words may have influenced the pattern of responding seen in Study 2. Future studies might vary the types of emotion words included as response options; specifically, we suggest studies include more social emotions (e.g. compassion) so that researchers can explore how such words are conceptualised. Results also suggest that people may conceptualise words as belonging to word families. Future studies might explore this (e.g. by asking participants to group words that belong together). Such studies have the potential to provide insight into hierarchies and the degree to which intensity influences how people conceptualise emotion words.

Participants were not provided with a 'none' or 'I do not know' option, something typical of forced-choice paradigms, which require participants to select only one option, regardless of how many they may attribute to the stimulus (Nelson & Russell, 2013). Instead, we allowed participants the flexibility to rate each stimulus according to its degree of fit with multiple response options. There may still have been times where participants felt uncertain of the meaning of the word; in such instances an escape option such as 'none', 'I do not know', or 'unfamiliar' could have reduced the likelihood of artificial agreements. However, other work (Frank & Stennett, 2001) suggests that the absence of an escape option is unlikely to have had much effect on the pattern of results. Finally, we acknowledge that order effects are possible since the order of response options remained consistent for each stimulus.

The word list, which was quite extensive in Study 1, was significantly reduced in Study 2. As a result conclusions are based on a relatively small number of words in each category. This makes further investigation of our findings important. It was necessary to reduce the list in Study 2 for pragmatic reasons. To do this, we had to make choices (detailed previously) about which emotion words we would include. Though we used word frequency as a guiding principal it is still possible that some words were unknown to participants. In future studies, researchers might consider including a response option that allows participants to indicate this. Also, for Study 2 we sought to keep word counts in each category in proportion to the category counts in Study 1. In retrospect, we should have selected an equal number of words from each category. The fact that we had relatively few Neutral words made it difficult to assess whether it was being used as a 'catch all' category or whether participants actually perceived some words as emotionally neutral. Finally, for Study 2 we chose words which people identified as fitting strongly into one of the five categories from Study 1. However, there were some emotion words where ratings in two categories were quite similar (e.g. *sympathetic* had similar Happy and Sad ratings, categories that were opposite in valence). Perhaps it is these not easily categorised words that require further investigation. Future studies involving such words might yield more valuable insights into how people conceptualise emotion words.

A second type of limitation concerns context. Participants were not provided any context to guide interpretation of word meaning which may have contributed to the

high diversity ratings. Fehr and Russell (1984) suggest that the degree to which people rate an emotion as belonging to an emotion category is dependent upon their appraisal of that emotion, which itself depends on their interpretation of the situation in which the emotion occurs (Siemer et al., 2007). It is possible that the absence of context in Studies 1 and 2 limited participants' ability to categorise words. With no prescribed context participants may have created their own, which would differ from person to person. For example, in Study 1 the primary rating for *console* was Sad, while the secondary rating was Happy. One would not expect the primary and secondary categories to be opposite in valence. However, perhaps respondents imagined a situation in which they were being consoled but differently appraised the valence of the word depending on whether they focused on the need to be consoled (negative), or its impact (positive). Individual differences in meaning appraisal could similarly have contributed to diversity scores in Study 2. To determine the extent to which the lack of context (or individual variation in inferred context) influences people's interpretation of a word, future studies may consider using a standard carrier phrase (e.g. If someone said 'I am feeling X', I would categorise that word as meaning: happy, sad, etc.). The use of such a carrier phrase does not provide context indicative of a specific word meaning but could support more consistency in approach.

5. Conclusion

Results of the two studies presented here showed that the response options we provide participants for judgements of emotion stimuli influences the way they conceptualise and categorise emotion words. Though desirable for data analysis and replication purposes, the use of forced-choice with only four to six basic emotion categories is likely too constraining, failing to adequately capture how emotion words are conceptualised. It may not be possible to identify one 'best' list of response options; nevertheless, results suggest that a broader array of positive emotion words is needed. Further, it appears that including more complex or social emotion words would enable participants to reflect on emotion intensity. Extending response options has the potential to provide important insight into how people (of various cultures) conceptualise and label emotion across the lifespan, informing our understanding of developmental trajectories. It may also highlight deficits in emotion differentiation and enable us to better characterise people's ability to interpret a broader scope of emotions associated with real-life contexts.

Supplementary Material. To view supplementary material for this article, please visit <http://doi.org/10.1017/langcog.2022.24>.

Competing interests. The authors declare none.

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, 12(2), 169–177.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(7), 2683–2690.

- Airdrie, J. N., Langley, K., Thapar, A., & van Goozen, S. H. M. (2018). Facial emotion recognition and eye gaze in attention-deficit/hyperactivity disorder with and without comorbid conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(8), 561–570. <https://doi.org/10.1016/j.jaac.2018.04.016>
- Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, 87(2), 266–281. <https://doi.org/10.1037/0022-3514.87.2.266>
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, 20(5), 286–290. <https://doi.org/10.1177/0963721411422522>
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 637–660. <https://doi.org/10.1017/S0140525X99532147>
- Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18(5–6), 513–562. <https://doi.org/10.1080/01690960344000026>
- Ben-David, B. M., van Lieshout, P. H. H. M., & Leszcz, T. (2011). A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury*, 25(2), 206–220. <https://doi.org/10.3109/026699052.2010.536197>
- Bradley, M. M. & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. [Technical Report C-1]. The Center for Research in Psychophysiology.
- Crane, E. & Gross, M. (2007). Motion capture and emotion: Affect detection in whole body movement. In A. Paia, R. Prada, & W. Picard (Eds.), *Affective computing and intelligent interaction. ACCII 2007. Lecture notes in computer science* (Vol. 4738, pp. 95–101). Springer. https://doi.org/10.1007/978-3-540-74889-2_9
- Czimskey, N. & Marquardt, T. P. (2019). Recall of emotional and neutral words and paragraphs in traumatic brain injury. *Aphasiology*, 33(8), 1019–1034. <https://doi.org/10.1080/02687038.2018.1512079>
- Davies, M. (2004). *British National Corpus*. Oxford University Press. <https://www.english-corpora.org/bnc/>
- Davies, M. (2008). *The Corpus of Contemporary American English (COCA)*. <https://www.english-corpora.org/coca/>
- Dickinson, E. R., Adelson, J. L., & Owen, J. (2012). Gender balance, representativeness, and statistical power in sexuality research using undergraduate student samples. *Archives of Sexual Behavior*, 41, 325–327. <https://doi.org/10.1007/s10508-011-9887-1>
- Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology*, 69(1), 130–141. <https://doi.org/10.1037/0022-3514.69.1.130>
- Dupuis, K. & Pichora-Fuller, M. K. (2015). Aging affects identification of vocal emotions in smenatically neutral sentences. *Journal of Speech, Language, and Hearing Research*, 58(3), 1061–1076. https://doi.org/10.1044/2015_JSLHR-H-14-0256
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169–200.
- Ekman, P. & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364–370. <https://doi.org/10.1177/1754073911410740>
- Fehr, B. & Russell, J. A. (1984). Concept of emotion viewed from prototype perspective. *Journal of Experimental Psychology*, 11(3), 464–486.
- Frank, M. G. & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology*, 80, 75–85. <https://doi.org/10.1037/0022-3514.80.1.75>
- Fugate, J. M. B., Hare, A. J. O., & Emmanuel, W. S. (2018). Emotion words: Facing change. *Journal of Experimental Social Psychology*, 79, 264–274. <https://doi.org/10.1016/j.jesp.2018.07.015>
- Gasper, K., Spencer, L. A., & Hu, D. (2019). Does neutral affect exist? How challenging three beliefs about neutral affect can advance affective research. *Frontiers in Psychology*, 10, 2476. <https://doi.org/10.3389/fpsyg.2010.02476>
- Goddard, C. (2014). Interjections and emotion (with special reference to ‘surprise’ and ‘disgust’). *Emotion Review*, 61(1), 53–63. <https://doi.org/10.1177/1754073913491843>
- Gregorius, H. & Gillet, E. M. (2008). Generalized Simpson-diversity. *Ecological Modelling*, 211, 90–96. <https://doi.org/10.1016/j.ecolmodel.2007.08.026>
- Grosbras, M.-H., Ross, P. D., & Belin, P. (2018). Categorical emotion recognition from voice improves during childhood and adolescence. *Scientific Reports*, 8, 14791. <https://doi.org/10.1038/s41598-018-32868-3>

- Guajardo, S. A. (2015). Measuring diversity in police agencies. *Journal of Ethnicity in Criminal Justice*, 13(1), 1–15. <https://doi.org/10.1080/15377938.2014.893220>
- Hauschild, K. M., Felsman, P., Keifer, C. M., & Lerner, M. (2020). Evidence of own-age bias in facial emotion recognition for adolescents with and without autism spectrum disorder. *Frontiers in Psychiatry*, 11, 428. <https://doi.org/10.3389/fpsy.2020.00428>
- Herpertz, S., Schutz, A., & Nezelek, J. (2016). Enhancing emotion perception, a fundamental component of emotional intelligence: Using multiple-group SEM to evaluate a training program. *Personality and Individual Differences*, 95, 11–19. <https://doi.org/10.1016/j.paid.2016.02.015>
- Hupka, R. B., Lenton, A. P., & Hutchison, K. A. (1999). Universal development of emotion categories in natural language. *Journal of Personality and Social Psychology*, 77(2), 247–278. <https://doi.org/10.1037/0022-3514.77.2.247>
- Koffer, R. E., Ram, N., Conroy, D. E., Pincus, A. L., & Almedia, D. M. (2016). Stressor diversity: Introduction and empirical intergration into the daily stress model. *Psychology and Aging*, 31(4), 301–320. <https://doi.org/10.1037/pag0000095>
- Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology, General*, 143(3), 1065–1081. <https://doi.org/10.1037/a0035669>
- Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion*, 6(1), 62–72. <https://doi.org/10.1037/1528-3542.6.1.62>
- Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, 19(5), 633–653. <https://doi.org/10.1080/02699930441000445>
- Lewis, M. (2008). Self conscious emotions: Embarrassment, pride, shame and guilt. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 742–756). The Guilford Press.
- Limbrecht-Ecklundt, K., Scheck, A., Jerg-Bretzke, L., Walter, S., Hoffmann, H., & Traue, H. C. (2013). The effect of forced choice on facial emotion recognition: A comparison to open verbal classification of emotion labels. *Psychosocial Medicine*, 10, Doc04. <https://doi.org/10.3025/psm000094>
- Melendez, J. C., Satorres, E., Reyes-Olmedo, M., Delhom, I., Real, E., & Lora, Y. (2020). Emotion recognition changes in a confinement situation due to COVID-19. *Journal of Environmental Psychology*, 72, 101518. <https://doi.org/10.1016/j.jenvp.2020.101518>
- Milders, M., Fuchs, S., & Crawford, J. R. (2003). Neuropsychological impairments and changes in emotional and social behaviour following severe traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 25(2), 157–172. <https://doi.org/10.1076/j.jcen.25.2.157.13642>
- Nelson, N. L. & Russell, J. A. (2013). Universality revisited. *Emotion Review*, 5(1), 8–15. <https://doi.org/10.1177/1754073912457227>
- Nettle, D. & Penke, L. (2010). Personality: Bridging the literatures from human psychology and behavioural ecology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1560), 4043–4050. <https://doi.org/10.1098/rstb.2010.0061>
- Niedenthal, P. M., Auxiette, C., Nugier, A., Dalle, N., Bonin, P., & Fayol, M. (2004). A prototype analysis of the French category ‘emotion’. *Cognition and Emotion*, 18(3), 289–312.
- Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry*, 54(5), 504–514. [https://doi.org/10.1016/S0006-3223\(03\)00168-9](https://doi.org/10.1016/S0006-3223(03)00168-9)
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Developmental Psychopathology*, 17(3), 715–724. <https://doi.org/10.1017/S0954579405050340>
- Quoidbach, J., Gruber, J., Mikolajczak, M., Kogan, A., Kotsou, I., & Norton, M. I. (2014). Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology*, 143(6), 2057–2066. <https://doi.org/10.1037/a0038025>
- Ram, N., Conroy, D., Pincus, A., Hyde, A., & Molloy, L. (2012). Tethering theory to method: Using measures of intraindividual variability to operationalize individuals’ dynamic characteristics. In J. R. Harring & G. R. Hancock (Eds.), *Advances in longitudinal methods in the social and behavioral sciences* (pp. 81–110). Information Age Publishing.
- Revilla, M. & Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, 59(5), 557–565. <https://doi.org/10.2501/IJMR-2017-039>

- Roseman, I. J. (1994). Emotions and emotion families. In N. H. Frijda (Ed.), *Proceedings of the 8th international conference of the international society for research on emotions* (pp. 171–175). Rutgers University. <https://doi.org/10.7282/T31N836X>
- Russell, J. A. (1980). Circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Russell, J. A., Rosenberg, E. L., & Lewis, M. D. (2011). Introduction to a special section on basic emotion theory. *Emotion Review*, 3(4), 363. <https://doi.org/10.1177/1754073911411606>
- Saleh, A. & Bista, K. (2017). Examining factors impacting online survey response rates in education research: Perceptions of graduate students. *Journal of Multidisciplinary Evaluation*, 13(29), 63–74. <http://jmde.com>
- Samson, A. C., Kreibig, S. D., Soderstrom, B., Wade, A. A., & Gross, J. J. (2016). Eliciting positive, negative and mixed emotional states: A film library for affective scientists. *Cognition and Emotion*, 30(5), 827–856. <https://doi.org/10.1080/02699931.2015.1031089>
- Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or set of modality- and emotion-specific skills? *Personality and Individual Differences*, 53, 16–21.
- Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52(6), 1061–1086. <https://doi.org/10.1037/0022-3514.52.6.1061>
- Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation—Different emotions: How appraisals shape our emotions. *Emotion*, 7(3), 592–600. <https://doi.org/10.1037/1528-3542.7.3.592>
- Stevenson, R. A., Mikels, J. A., & James, T. W. (2007). Characterization of the affective norms for English words by discrete emotional categories. *Behavior Research Methods*, 39(4), 1020–1024.
- Storm, C. & Storm, T. (1987). A taxonomic study of the vocabulary of emotions. *Journal of Personality and Social Psychology*, 53(4), 805–816. <https://doi.org/10.1037//0022-3514.53.4.805>
- Strauss, G. & Allen, D. (2008). Emotional intensity and categorisation ratings for emotional and non-emotional words. *Cognition & Emotion*, 22(1), 114–133. <https://doi.org/10.1080/02699930701319154>
- Torre, J. B. & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review*, 10(2), 116–124. <https://doi.org/10.1177/1754073917742706>
- Trampe, D., Quoidbach, J., & Taquet, M. (2015). Emotions in everyday life. *PLoS ONE*, 10(12), e0145450. <https://doi.org/10.1371/journal.pone.0145450>
- Turkstra, L. S., Kranning, S. G., Riedeman, S. K., Mutlu, B., Duff, M., & VanDenHeuvel, S. (2017). Labelling facial affect in context in adults with and without TBI. *Brain Impairment*, 18(1), 49–61. <https://doi.org/10.1017/BrImp.2016.29>
- Zupan, B. & Babbage, D. R. (2008). *Investigating the effectiveness of film versus text in eliciting emotion in a laboratory setting*. Brock University.
- Zupan, B. & Babbage, D. R. (2017). Film clips and narrative text as subjective emotion elicitation techniques. *Journal of Social Psychology*, 157(2), 194–210. <https://doi.org/10.1080/00224545.2016.1208138>
- Zupan, B. & Eskritt, M. (2020). Eliciting emotion ratings for a set of contemporary film clips. *Journal of Social Psychology*, 160(6), 768–789. <https://doi.org/10.1080/00224545.2020.1758016>
- Zupan, B. & Eskritt, M. (2022). Validation of affective sentences: Extending beyond basic emotion categories. *Journal of Psycholinguistic Research*. <https://doi.org/10.1007/s10936-022-09906-3>

Appendix. Instructions given to participants for Study 1 and Study 2

A.1. Study 1

The following directions were provided following each stimulus:

- Rate the intensity level of the word from 1 (extremely low) to 9 (extremely high).
- Rate how positive or negative you think this word is from 1 (extremely negative) to 9 (extremely positive).
- Rate the degree to which the word fits the emotion category [Happy] from 1 (not at all) to 9 (extremely well).

A.2. Study 2

A.2.1. Initial instructions

For the following questions, you will be presented with a WORD and asked to rate how well that word fits a number of different emotion categories on a scale of 1 to 9, with 9 indicating the highest degree of fit and 1 indicated that the word does not fit into that category at all.

You can select a rating for as many categories as you wish for each word. If you do not feel the word fits into a category, you can leave that category blank and we will assume a rating of 1. However, you will be required to rate at least 3 emotion categories for each word, with a minimum rating of 1.

A.2.2. Instructions for each stimulus

- How well does the word [X] fit into each of the following emotion categories (1 = not at all; 9 = extremely well)?