The psychometric assessment of alexithymia: Development and validation of the Perth Alexithymia Questionnaire

David Preece*, Rodrigo Becerra, Ken Robinson, Justine Dandy, Alfred Allan

Edith Cowan University, Psychology Department, School of Arts and Humanities, 270 Joondalup Drive, Joondalup, 6027 Perth, Western Australia, Australia

ABSTRACT

Alexithymia is a trait comprising people's ability to focus *attention* on and accurately *appraise* their own emotions. Its assessment is of clinical interest because people who have difficulty processing their negative and positive emotions are more vulnerable to developing psychopathology symptoms, however, existing alexithymia measures cannot comprehensively assess the construct across both negative and positive emotions. In this paper, we attempt to remedy these measurement limitations by developing and validating a new 24-item self-report measure, the Perth Alexithymia Questionnaire (PAQ), which is based on the *attention-appraisal model of alexithymia*. In Study 1, our confirmatory factor analyses in a sample of 231 adults suggested that the PAQ had a factor structure consistent with its theoretical basis; it could separately measure all components of the construct and do so across negative and positive emotions. All subscale and composite scores had high internal consistency reliability. Study 2 (*N* = 748) replicated these findings with respect to the PAQ's factor structure and internal consistency reliability, and statistical comparisons with measures of psychopathology and emotion regulation supported the PAQ's concurrent and discriminant validity. Our data therefore suggest that the PAQ has strong psychometric properties as a measure of alexithymia. Clinical and research applications are discussed.

1. Introduction

Emotions manifest as responses across the subjective-experiential (e.g., feeling of fear), physiological (e.g., increased breathing-rate), and behavioural channels (e.g., urge to escape) of the emotion system (Gross, 2014). Emotions may be negatively valenced, like sadness and anger, or positively valenced, like happiness (Bradley & Lang, 2007). People differ in their capacity to process their emotional responses, and these variations reflect individual differences in the trait *alexithymia* (Gross, 2014; Lane, Weihs, Herring, Hishaw, & Smith, 2015; Nemiah & Sifneos, 1970; Preece, Becerra, Allan, Robinson, & Dandy, 2017; Taylor, Bagby, & Parker, 1999; Vorst & Bermond, 2001).

The alexithymia construct is of substantial clinical interest, because high levels of alexithymia appear to be an important transdiagnostic risk factor for a range of psychopathologies, including depressive (Honkalampi, Hintikka, Laukkanen, & Viinamäki, 2001), anxiety (Zeitlin & McNally, 1993), psychosomatic (Duddu, Isaac, & Chaturvedi, 2003), substance use (Thorberg, Young, Sullivan, & Lyvers, 2009), eating (Taylor, Parker, Bagby, & Bourke, 1996) and personality (Berenbaum, 1996) disorders. High levels of alexithymia can also reduce the efficacy of some psychotherapy approaches (e.g., psychoanalysis; Leweke, Bausch, Leichsenring, Walter, & Stingl, 2009).

Researchers have consequently developed several measures of alexithymia, but as we will demonstrate below, these measures have some notable limitations that reduce their clinical and research utility. We, in this paper, attempt to remedy this by reporting on our development and validation of a new self-report measure of alexithymia called the Perth Alexithymia Questionnaire (PAQ).

1.1. Theoretical background

Alexithymia (meaning "no words for emotions" in Greek) was first coined by psychoanalytic practitioners (Nemiah & Sifneos, 1970; Sifneos, 1973) to describe the presentation of psychosomatic patients who were commonly unable to "describe their feelings or to differentiate among them" and displayed "an absence of the capacity to produce fantasies with the result that [their] thought content [was] restricted to a preoccupation with external objects, people, and environmental events" (Nemiah, 1984, p. 127). Early theoretical models of alexithymia were therefore mostly underpinned by psychoanalytic ideas, and conceptualised alexithymia as a multidimensional construct comprised of at least four components: difficulty identifying one's own feelings (DIF); difficulty describing feelings (DDF); an externally orientated thinking style (EOT) marked by an excessive focus on external

---

stimuli rather than internal experiences; and difficulty fantasising (DFAN) marked by the absence or scarcity of daydreams and fantasies (e.g., Nemiah, 1977; Taylor, Ryan, & Bagby, 1985). Proponents of these psychoanalytic models hypothesised that people with high levels of alexithymia were more vulnerable to somatic or psychiatric symptoms because they were unable to use mental elaboration or fantasy to regulate the energy of their instinctual drives (e.g., McDougall, 1974; Nemiah, 1977).

Several psychometric tools were subsequently designed to assess these proposed components of alexithymia (e.g., Bagby, Taylor, Parker, & Dickens, 2006; Bermond, Vorst, Vingerhoets, & Gerritsen, 1999; Taylor et al., 1985), and statistical analyses of these measures supported most specifications of the early theoretical models. DIF, DDF and EOT subscales, for example, were commonly found to correlate positively (e.g., Bagby, Parker, & Taylor, 1994; Vorst & Bermond, 2001) and load together on the same higher-order "alexithymia" factor in factor analyses (e.g., Gignac, Palmer, & Stough, 2007; Preece, Becerra, Robinson, & Dandy, 2017). Little statistical support emerged, however, for the inclusion of DFAN, suggesting that its inclusion in early models was likely a misspecification (for a review, see Preece, Becerra, Allan, et al., 2017). DFAN subscales were found, in most empirical studies, to be uncorrelated or negatively correlated with DIF, DDF and EOT subscales (e.g., Preece, Becerra, Allan, et al., 2017; Taylor et al., 1985; Vorst & Bermond, 2001; Watters, Taylor, & Bagby, 2016).

Some test developers consequently removed all DFAN items from their alexithymia measures (e.g., Bagby et al., 1994; Sekely, Taylor, & Bagby, 2018) or changed their scoring procedures so that DFAN items were not included when calculating an overall alexithymia score (e.g., Vorst & Bermond, 2001). Until recently, though, this body of empirical work had not resulted in any substantial modifications to alexithymia models and they all still included DFAN within their definition of alexithymia (e.g., Bermond et al., 1999; Sifneos, 1996; Taylor et al., 1999). This discrepancy between alexithymia models and the alexithymia measurement (Bagby, Taylor, Quilty, & Parker, 2007) was, however, recently addressed by Preece, Becerra, Allan, et al. (2017) via their introduction of the *attention-appraisal model of alexithymia*. The attention-appraisal model was an evolution of earlier alexithymia models (Sifneos, 1996; Taylor et al., 1999), but included modifications to be consistent with the abovementioned body of empirical findings (i.e., removing DFAN), and modifications to directly align it with established cognitive models of emotion regulation (Gross's (2015) *extended process model of emotion regulation*) and emotion processing (Lane and Schwartz's (1987) *cognitive-developmental theory of levels of emotional awareness*).

We think the conceptual clarity afforded by this new alexithymia model, alongside the accumulated body of data on existing alexithymia measures, consequently provides an excellent opportunity to now build on this work and develop the PAQ as a new alexithymia measure with better psychometric properties than existing measures. Prior to describing the structure of our proposed PAQ, we firstly provide a more detailed description of the theoretical model upon which it is based, outline the psychometric criteria against which we think the utility of any alexithymia measure must be evaluated, and briefly review how well existing measures meet these criteria.

1.2. Attention-appraisal model of alexithymia

The attention-appraisal model (Preece, Becerra, Allan, et al., 2017), which underpins the proposed PAQ, defines alexithymia as a continuous and multidimensional construct comprised of three interrelated components: DIF, DDF, and EOT. These components are conceptualised within a valuation systems framework; valuation systems being systems comprised of a four-stage *situation-attention-appraisal-response* sequence, through which a person valuates (evaluates) the meaning of a stimulus (see Gross, 2015). Normally when an emotional response becomes the stimulus (situation stage) that is target of valuation, the person focuses

his or her attention on the emotional response (attention stage), he or she then appraises the emotional response in terms of what it is and what it means (appraisal stage) and, based on this appraisal, he or she might activate a goal to try to modify the emotion (response stage; i.e., emotion regulation; Gross, 2015). EOT is conceptualised as difficulty at the *attention* stage of this valuation system, and DIF and DDF are difficulties at the *appraisal* stage. In other words, when an emotional response occurs, people with high levels of alexithymia have trouble focusing their attention on it and trouble accurately appraising what it is. There is a subtle shift in emphasis here when describing EOT relative to early psychoanalytic models (e.g., Nemiah, 1984); the pertinent point is not that alexithymic people focus excessively on external stimuli, but rather, from the reverse perspective, that they do not properly focus their attention on their emotions. Thus, DIF, DDF and EOT are considered components of a common latent construct because they are deficits specific to the emotion valuation process (Preece, Becerra, Allan, et al., 2017).

The severity of these attention and appraisal difficulties is further understood and categorised in this model according to the five Piagetian cognitive-developmental levels of emotional awareness first described by Lane and Schwartz (1987). People operating at a low developmental level (i.e., high alexithymia) experience emotions only as undifferentiated pleasant or unpleasant states (e.g., "I am feeling *bad*" or "I am feeling *good*"), whereas people operating at a higher developmental level (i.e., low alexithymia) experience emotions in a more nuanced and differentiated manner (e.g., "I am feeling *angry*, not *sad*", or "I am feeling *excited*, not *amused*"; Lane & Schwartz, 1987). Preece, Becerra, Allan, et al. (2017) posit, consistent with the current body of empirical findings, that people's level of alexithymia depends on the developmental level of their emotion schemas (i.e., those cognitive structures used to process emotions; Lane et al., 1996; Luminet, Vermeulen, Demaret, Taylor, & Bagby, 2006; Lundh, Johnsson, Sundqvist, & Olsson, 2002; Suslow & Junghanns, 2002; Vermeulen, Luminet, & Corneille, 2006) and the extent to which they use experiential avoidance of emotions as an emotion regulation strategy (Bilotta, Giacomantonio, Leone, Mancini, & Coriale, 2015; Coriale et al., 2012; Panayiotou et al., 2015).

1.3. Criteria for judging measures of alexithymia

A measure of alexithymia must have good levels of validity and reliability to have research and clinical utility (Groth-Marnat, 2009). A fundamental starting point for such validity is that an alexithymia measure's content should capture all facets of the construct. Because alexithymia is a multidimensional construct, this assumes that there is some statistical or theoretical value in being able to assess each component of the construct separately, as well as some value in being able to combine all components together into an overall composite score (Reise, Moore, & Haviland, 2010). An alexithymia measure should therefore include DIF, DDF and EOT items, and should allow for separate subscales to be derived for each of these components.

A measure of alexithymia, which is an affective phenomenon, should also be able to assess it across both *negatively* and *positively* valenced emotions (John & Eng, 2014). Valence-specific measurement is, indeed, now common in newer measures of other affective phenomena like emotion regulation and emotional reactivity (e.g., Becerra, Preece, Campitelli, & Scott-Pillow, 2017; Ripper, Boyes, Clarke, & Hasking, 2018; Weiss, Gratz, & Lavender, 2015; Zou, Plaks, & Peterson, 2017), and recent empirical work has highlighted that valence is an important consideration in alexithymia assessments. Feldman Barrett, Gross, Christensen, and Benvenuto (2001) have, for example, demonstrated that people's ability to differentiate between their various negative emotions is not equivalent to their ability to differentiate between their various positive emotions, and van der Velde et al.'s (2013) recent meta-analysis demonstrated that alexithymia has different neural correlates depending on whether the emotions being processed are

negative or positive. Localised brain injury may therefore cause valence-specific emotion processing deficits (Becerra, Amos, & Jongenelis, 2002). Emotional valence is of most relevance when attempting to assess functioning at the appraisal stage of emotional valuation (i.e., DIF and DDF) because, theoretically speaking, it is not until the appraisal stage of emotion valuation that a valence judgement is made (Gross, 2015; Ochsner & Gross, 2014; Preece, Becerra, Allan, et al., 2017). It is, hence, less appropriate to include valence when attempting to isolate the earlier attention stage (i.e., EOT).[1] An alexithymia measure should therefore allow for separate DIF and DDF subscales to be derived for negative and positive emotions.

Outside of these content considerations, an alexithymia measure should also meet accepted statistical standards when its validity and reliability are formally tested (Kline, 2013). Many psychometricians agree, for example, that when subjected to factor analysis the pattern of factor loadings and factor intercorrelations in a measure should be consistent with its theoretical basis, and all items should load well (factor loadings ≥ 0.40) on their intended latent factor (i.e., factorial validity; Groth-Marnat, 2009; Kline, 2013). Scores should, similarly, correlate in expected ways with established measures of other constructs (i.e., concurrent validity), and reliability coefficients should be at least 0.70 for a score to be used in research and ideally 0.90 or above for a score to be used in clinical decision making (Groth-Marnat, 2009; Nunnally & Bernstein, 1994).

For sake of clarity, in the remainder of this paper, we categorise and label these abovementioned considerations into *three* broad measurement criteria, against which we will evaluate all available alexithymia measures:

(1) Allows separate DIF, DDF, and EOT subscales to be derived;
(2) Accounts for emotional valence when assessing functioning at the appraisal stage of emotion valuation (i.e., DIF and DDF);
(3) Subscale and composite scores have adequate validity and reliability when tested statistically.

1.4. Existing measures of alexithymia

We identified 14 psychometric tools that are either specifically designed to measure alexithymia or are designed to measure a broader construct (e.g., emotional intelligence) but have some alexithymia subscales (see Table 1). In our view, none of these measures meet all three of the abovementioned measurement criteria. With respect to our first criterion, ten measures cannot produce separate DIF, DDF and EOT subscales (ASC, CAM, CAQ-AP, DERS, LEAS, M-BIQ, OAS, PTI-AS, RAS, TMMS). With respect to our second criterion, no measures can produce valence-specific scores. With respect to our third criterion, two measures have at least one subscale that has low validity or reliability (TAS-20, BVAQ).

Because the TAS-20 (Bagby et al., 1994) and BVAQ (Vorst & Bermond, 2001) are presently the most widely used measures of alexithymia, we briefly examine these two measures against our criteria in more detail below.

1.4.1. 20-item Toronto Alexithymia Scale

The TAS-20 (Bagby et al., 1994) is a 20-item self-report measure of alexithymia. Items correspond to three subscales, designed to measure DIF (7 items), DDF (5 items), or EOT (8 items). All items are also summed into a total scale score as an overall marker of alexithymia (Bagby et al., 1994). The TAS-20 therefore meets the first of our three measurement criteria.

Table 1
A list of existing psychometric tools designed to measure alexithymia or designed to measure a broader construct (i.e., emotional intelligence) but which have some alexithymia subscales.

Name and type of measure
Self-report measures
Levels of Emotional Awareness Scale (LEAS; Lane, Quinlan, Schwartz, Walker, & Zeitlin, 1990).
20-item Toronto Alexithymia Scale (TAS-20; Bagby, Taylor, & Parker, 1994).
Trait Meta-Mood Scale (TMMS; Salovey, Mayer, Goldman, Turvey, & Palfai, 1995).
Bermond-Vorst Alexithymia Questionnaire (BVAQ; Vorst & Bermond, 2001).
Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004).
Emotion Awareness Questionnaire (EAQ-30; Rieffe, Oosterveld, Miers, Terwogt, & Ly, 2008).
Psychological Treatment Inventory – Alexithymia Scale (PTI-AS; Gori, Giannini, Palmieri, Salvini, & Schuldberg, 2012).
Observer-rated measures
California Q-set Alexithymia Prototype (CAQ-AP; Haviland & Reise, 1996).
Alexithymia Scale for Children (ASC; Fukunishi, Yoshida, & Wogan, 1998).
Modified Beth Israel Hospital Psychosomatic Questionnaire (M-BIQ; Taylor, Bagby, & Parker, 1999).
Observer Alexithymia Scale (OAS; Haviland, Warren, & Riggs, 2000).
Toronto Structured Interview for Alexithymia (TSIA; Bagby, Taylor, Parker, & Dickens, 2006).
Children's Alexithymia Measure (CAM; Way et al., 2010).
Projective tests
Rorschach Alexithymia Scale (RAS; Porcelli & Mihura, 2010).

Note. Some of these measures are revisions of earlier measures. We list only the latest version of each measure here.

Two of the DIF items specify a negative valence (e.g., "I often don't know why I'm angry"), but all other DIF and DDF items specify no valence (e.g., "I am often confused about what emotion I am feeling"). Thus factor analytic studies of the TAS-20 suggest no valence-specific subscales can be derived (e.g., Kooiman, Spinhoven, & Trijsburg, 2002). The TAS-20 therefore does not meet our second measurement criterion.

The total scale score and DIF and DDF subscales commonly have adequate factorial validity and reliability coefficients over 0.70, but the EOT subscale does not (Preece, Becerra, Robinson, & Dandy, 2017). Cronbach's alpha reliability coefficients for the EOT subscale are often under 0.60, and many EOT items load poorly (factor loadings < 0.40) on their intended factor (e.g., Kooiman et al., 2002; Preece, Becerra, Robinson, & Dandy, 2017; Taylor, Bagby, & Parker, 2003). Available data therefore suggest that the TAS-20 is of limited utility for any clinical or research purposes that require the isolation of EOT (e.g., Leweke, Leichsenring, Kruse, & Hermes, 2012; Lyvers, McCann, Coundouris, Edwards, & Thorberg, 2018). Additionally, whilst the TAS-20 total scale score consistently meets minimum reliability standards for use in research (e.g., Taylor et al., 2003), we are not aware of any studies where it reaches the 0.90 value that is desired for clinical decision making. The TAS-20 therefore does not meet our third measurement criterion. Recent psychometric studies have suggested that these reliability problems are due to its factor structure being disrupted by the poor content validity of several EOT items (Preece, Becerra, Robinson, & Dandy, 2017) and the reverse-scored format[2] of five items (Gignac et al., 2007; Meganck, Vanheule, & Desmet, 2008; Preece, Becerra, Robinson, & Dandy, 2017).

---

[1] When EOT items in other self-report measures have been modified to include a negative valence (e.g., a modified form of the Difficulties in Emotion Regulation Scale where the phrase "When I'm upset, …" was added to the start of all DIF and EOT items), these EOT items become statistically indistinguishable from DIF items (see Bardeen, Fergus, Hannan, & Orcutt, 2016).

[2] Reverse-scored items are items that mean the opposite to the other items in the scale. In the case of the TAS-20, the five reverse-scored items describe a *low* level of alexithymia, rather than a *high* level of alexithymia. Participant responses on the 5-point Likert scale for these reverse-scored items must, therefore, be reversed by the examiner prior to calculating subscale and total scale scores. Recent research has found that including reverse-scored items within a self-report scale is problematic because it tends to increase cognitive burden on the examinee, it produces a method factor within the scale's factor structure, and it decreases internal consistency reliability (see van Sonderen et al., 2013).

1.4.2. Bermond-Vorst Alexithymia Questionnaire

The BVAQ (Vorst & Bermond, 2001) is a 40-item self-report measure of alexithymia. The BVAQ is based on Bermond et al.'s (1999) theoretical model of alexithymia, in which they hypothesise that DIF, DDF, EOT, DFAN, and reduced emotional reactivity (difficulty emotionalising; DEMO) are components of alexithymia. Items therefore correspond to five subscales; DIF (8 items), DDF (8 items), EOT (8 items), DFAN (8 items), and DEMO (8 items). In practice, however, the DFAN and DEMO subscales are not summed into the same total scale score as the DIF, DDF and EOT subscales. Standard scoring, based on factor analytic results (e.g., Bermond et al., 2007), involves the DIF, DDF and EOT subscales being summed into a composite score that Bermond et al. (2007) label *cognitive alexithymia*, and the DFAN and DEMO subscales being summed into a separate (uncorrelated) composite score labelled *affective alexithymia*. Proponents of the attention-appraisal model consider only the cognitive alexithymia score to be a marker of alexithymia (Preece, Becerra, Allan, et al., 2017). The BVAQ therefore meets the first of our three measurement criteria.

Five DIF items and two DDF items specify a negative valence (e.g., "When I am hard on myself, it remains unclear to me whether I am sad or afraid or unhappy"), and two DIF items specify a positive valence (e.g., "When I am in a sunny mood, I know whether I am enthusiastic or cheerful or elated [reverse-scored]"), but the remaining seven DIF or DDF items specify no valence (e.g., "I can express my feelings verbally [reverse-scored]"). Factor analyses of the items, hence, suggest that no valence-specific subscales can be derived (Vorst & Bermond, 2001). The BVAQ therefore does not meet our second measurement criterion.

The DIF and DDF subscales commonly reach minimum reliability standards for use in research (Cronbach's alpha > 0.70; Bermond et al., 2007; Vorst & Bermond, 2001), but the EOT subscale sometimes does not (e.g., Bermond et al., 2007; Müller, Bühner, & Ellgring, 2004). In all studies we have seen, reliability coefficients for the cognitive alexithymia composite score have also not reached the 0.90 value desired for clinical decision making (e.g., Preece, Becerra, Allan, et al., 2017; Vorst & Bermond, 2001). The BVAQ therefore does not meet our third measurement criterion. Reliability coefficients are likely lowered by the high proportion of reverse-scored items (20 items) in the measure (van Sonderen, Sanderman, & Coyne, 2013). Recent research has, furthermore, found that the DEMO subscale score is likely not a valid marker of the emotional reactivity construct (see Preece, Becerra, Allan, et al., 2017) because this subscale does not distinguish between negative reactivity and positive reactivity (that is, reactivity with respect to negative emotions and reactivity with respect to positive emotions), which researchers using dedicated emotional reactivity measures have found to be separable dimensions that are negatively correlated with each other (e.g., Becerra et al., 2017; Ripper et al., 2018). Indeed, in factor analysis, the BVAQ DEMO subscale does not load on the same factors as other emotional reactivity questionnaires (Preece, Becerra, Allan, et al., 2017). Thus, even for clinicians and researchers who want to measure alexithymia *and* emotional reactivity, available evidence suggests that the BVAQ is not an efficient measure.

1.5. Perth Alexithymia Questionnaire

To provide an alexithymia measure that meets all three of the abovementioned measurement criteria, we have developed the PAQ. The PAQ is intended for clinicians and researchers who want to work within the framework of the attention-appraisal model (Preece, Becerra, Allan, et al., 2017) and assess alexithymia in adults and adolescents. It is a 24-item self-report measure, with all items comprised of a statement designed to assess the DIF, DDF, or EOT components of alexithymia. Respondents answer each item on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree), with higher scores indicating higher levels of alexithymia. We used a 7-point Likert scale format because it is common in emotional assessment tools (e.g., Emotion Regulation Questionnaire [ERQ]; Gross & John, 2003) and

there is some evidence that 7-point Likert scales perform better than 5-point (or less) scales when measuring a continuous construct like alexithymia (e.g., Preston & Colman, 2000). In line with the recommendations of van Sonderen et al. (2013) and others (e.g., Rodebaugh, Woods, & Heimberg, 2007), no items are reverse-scored.

All items corresponding to the appraisal stage of emotion valuation (i.e., the DIF and DDF components) account for valence, and are designed to assess people's ability to appraise either *negative* or *positive* emotions; thus all DIF and DDF items begin with some variant of the phrase "When I'm feeling *bad*, …" or "When I'm feeling *good*, …". The intention of this phrasing style is to describe an undifferentiated unpleasant or pleasant state that reflects how people with a low developmental level of emotional awareness experience emotions (Lane & Schwartz, 1987). The remainder of the item then describes one's ability to move beyond this low developmental level (e.g., "When I'm feeling *bad*, I can't tell whether I'm sad, scared, or angry" or "When I'm feeling *good*, I can't find the right words to describe those feelings"). Respondents who agree with these DIF and DDF items are therefore indicating that, during the appraisal stage of emotion valuation, they operate at a low developmental level. For all items corresponding to the attention stage of emotion valuation (i.e., the EOT component) we do not specify a valence. This is because, as aforementioned, it is not until the appraisal stage of emotion valuation that a valence judgement is made (Gross, 2015; Ochsner & Gross, 2014; Preece, Becerra, Allan, et al., 2017). To maximise the measurement distinction between the attention and appraisal stages, our EOT items therefore include no valence appraisal and measure people's tendency to not focus attention on their emotions in the first place (e.g., "I prefer to just let my emotions happen in the background, rather than focus on them" or "I don't pay attention to my emotions").

We structured the 24-item PAQ so that there is an equal number of items (8 items) corresponding to the DIF, DDF, and EOT components of the construct. This number of items was selected with a view to achieving high levels of reliability, whilst keeping the measure reasonably brief. Because the DIF and DDF components feature negatively and positively valanced items, half their items correspond to negative feelings and half correspond to positive feelings. Five subscale scores can therefore be derived (see Table 2): *Negative-Difficulty identifying feelings* (N-DIF; 4 items, e.g., "When I'm feeling bad, I get confused about what emotion it is"), *Positive-Difficulty identifying feelings* (P-DIF; 4 items, e.g., "When I'm feeling good, I can't tell whether I'm happy, excited, or amused"), *Negative-Difficulty describing feelings* (N-DDF; 4 items, e.g., "When something bad happens, it's hard for me to put into words how I'm feeling"), *Positive-Difficulty describing feelings* (P-DDF; 4 items, e.g., "When I'm feeling good, I can't talk about those feelings in much depth or detail"), and *General-Externally orientated thinking* (G-EOT, 8 items, e.g., "I prefer to focus on things I can actually see or touch, rather than my emotions").

The five subscales of the PAQ are, furthermore, designed to be combined into a number of theoretically meaningful composite scores (see Table 2). To generate overall markers of DIF or DDF, generalised across both valence types, the N-DIF and P-DIF subscales can be combined into a *General-Difficulty identifying feelings* composite (G-DIF, 8 items), and the N-DDF and P-DDF subscales can be combined into a *General-Difficulty describing feelings* composite (G-DDF, 8 items). Moreover, because DIF and DDF are hypothesised to be particularly closely linked (i.e., both correspond to the appraisal stage of emotion valuation), broader scores reflecting the appraisal stage can be derived. The N-DIF and N-DDF subscales can be combined into a *Negative-Difficulty appraising feelings* composite (N-DAF, 8 items), the P-DIF and P-DDF subscales can be combined into a *Positive-Difficulty appraising feelings* composite (P-DAF, 8 items), and the N-DIF, N-DDF, P-DIF, and P-DDF subscales can all be combined into a *General-Difficulty appraising feelings* composite (G-DAF, 16 items). Lastly, to produce an overall marker of alexithymia, all five subscales can be combined into an *Alexithymia* composite (ALEXI, 24 items).

Table 2
A list of the subscale and composite scores that can be derived from the PAQ.

Name	No. of items	Content measured
Subscales		
N-DIF	4	Difficulty identifying, understanding, and differentiating between one's own *negative* feelings.
P-DIF	4	Difficulty identifying, understanding, and differentiating between one's own *positive* feelings.
N-DDF	4	Difficulty describing and communicating one's own *negative* feelings.
P-DDF	4	Difficulty describing and communicating one's own *positive* feelings.
G-EOT	8	Tendency to not focus attention on one's own emotions (negative and positive).
Composites		
G-DIF	8	Difficulty identifying, understanding, and differentiating between one's own feelings (negative and positive).
G-DDF	8	Difficulty describing and communicating one's own feelings (negative and positive).
N-DAF	8	Difficulty identifying and describing (i.e., appraising) one's own *negative* feelings.
P-DAF	8	Difficulty identifying and describing (i.e., appraising) one's own *positive* feelings.
G-DAF	16	Difficulty identifying and describing (i.e., appraising) one's own feelings (negative and positive).
ALEXI	24	Overall alexithymia; difficulty focusing attention on and appraising one's own feelings (negative and positive).

Note. PAQ = Perth Alexithymia Questionnaire, N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking, G-DIF = General-Difficulty identifying feelings, G-DDF = General-Difficulty describing feelings, N-DAF = Negative-Difficulty appraising feelings, P-DAF = Positive-Difficulty appraising feelings, G-DAF = General-Difficulty appraising feelings, ALEXI = alexithymia.

1.6. Psychometric studies of the Perth Alexithymia Questionnaire

We report the results of two psychometric studies of the PAQ here. Study 1 describes the item selection process and examines the factor structure and internal consistency reliability of the measure. Study 2 replicates this examination of factor structure and internal consistency reliability in another sample, and then examines concurrent and discriminant validity.

2. Study 1

2.1. Method

2.1.1. Participants and procedure

Study 1's sample was comprised of 231 adults (65.4% female) with a mean age of 41.52 years (*SD* = 16.93, range = 18–85).[3] All participants were English speaking residents of Australia, with 73.6% reporting Australia as their country of birth. For 32.9% their highest level of completed education was high school, for 29.9% it was a technical diploma, and for 36.8% it was a university degree. Participants were recruited via three avenues: an online survey recruiting company (Qualtrics panels), an advertisement on a social media website, or an advertisement on the unit website of an undergraduate psychology unit. About one quarter (25.5%) of the sample were current university students.

The PAQ was administered as part of an anonymous online survey. We administered it in an over-inclusive 66-item "development" form in Study 1 to provide us with a large pool of items to select from. We wrote 22 items designed to measure DIF, 22 items designed to measure DDF, and 22 items designed to measure EOT (see supplementary material). For the DIF and DDF components, there were negatively and positively valenced versions of each item (e.g., "When I'm feeling bad, I can't make sense of those feelings" and "When I'm feeling good, I can't make sense of those feelings"). Based on some *preliminary* exploratory factor analyses (EFAs) and confirmatory factor analyses (CFAs) on different clusters of these items,[4] we selected 24 to form the final PAQ (results

from these preliminary analyses are not reported in this paper, but some are provided in the supplementary material). Item selection was based on three criteria. Firstly, to properly capture the breadth of the alexithymia construct, we wanted the DIF, DDF and EOT components to be evenly represented in the measure (i.e., 8 items per component). Secondly, so that valence-specific comparisons could be made, for the DIF and DDF components, we wanted an equal number of negative and positive items. We also wanted the negative and positive versions of each subscale to have the same item content except for their valence-specific terms. In other words, if we selected an item for the P-DIF subscale (e.g., "When I'm feeling good, I get confused about what emotion it is") its equivalently worded item for the N-DIF subscale (e.g., "When I'm feeling bad, I get confused about what emotion it is") would also need to be selected. Thirdly, we required that all retained items load strongly (factor loading ≥ 0.40) on their intended factor and not cross-load over multiple factors (Gerbing & Anderson, 1988; Raubenheimer, 2004). Whilst 66 PAQ items were administered, in this paper we report the results of analyses that include *only* the 24 retained items.

2.1.2. Materials

2.1.2.1. Perth Alexithymia Questionnaire. The PAQ is a 24-item self-report measure of alexithymia. It is freely available for use and is provided in the supplementary material.

2.1.3. Analytic strategy

CFAs were conducted using AMOS 24. All other analyses were conducted using SPSS 24. All 24 PAQ items were reasonably normally distributed (max skewness = 0.86, max kurtosis = −1.24).

2.1.3.1. Descriptive statistics. Means and standard deviations were calculated for the 24-item PAQ's subscale and composite scores. To examine whether emotional valence (negative or positive) influenced the extent of people's appraisal difficulties, a paired *t*-test was conducted to compare N-DAF and P-DAF scores. Because people are generally driven by hedonistic motivations to obtain pleasure and avoid pain (Gross, 2014), and some variance in alexithymia is accounted for by avoidant defences (Preece, Becerra, Allan, et al., 2017), these

---

[3] Some additional participants also completed the online survey. However, their data was excluded during quality screening because they failed an attention check question and/or completed the survey impossibly quickly (i.e., at a rate of < 2 s per question, suggesting inattentive responding).

[4] Because at least five participants per variable in the analysis are usually required to conduct a robust factor analysis (e.g., Gorsuch, 1983; Kline, 1979), and we had 231 participants, in our preliminary analyses we analysed smaller *clusters* of items rather than analysing all 66 items together (Gerbing & Anderson, 1988; Raubenheimer, 2004). Our first EFA included the first 33 of these 66 items, and our second EFA included the last 33

(*footnote continued*)

items. CFAs were also conducted on these item clusters, principally to maximise the goodness-of-fit of the factor solution by identifying and excluding those items that had pronounced covariances between their error terms. A group of 24 items that performed well in these analyses were then selected as candidates for retention and subjected to another EFA. Following this, our *main* (rather than *preliminary*) analyses were a series of CFAs on these 24 items, the results of which are presented in Study 1.
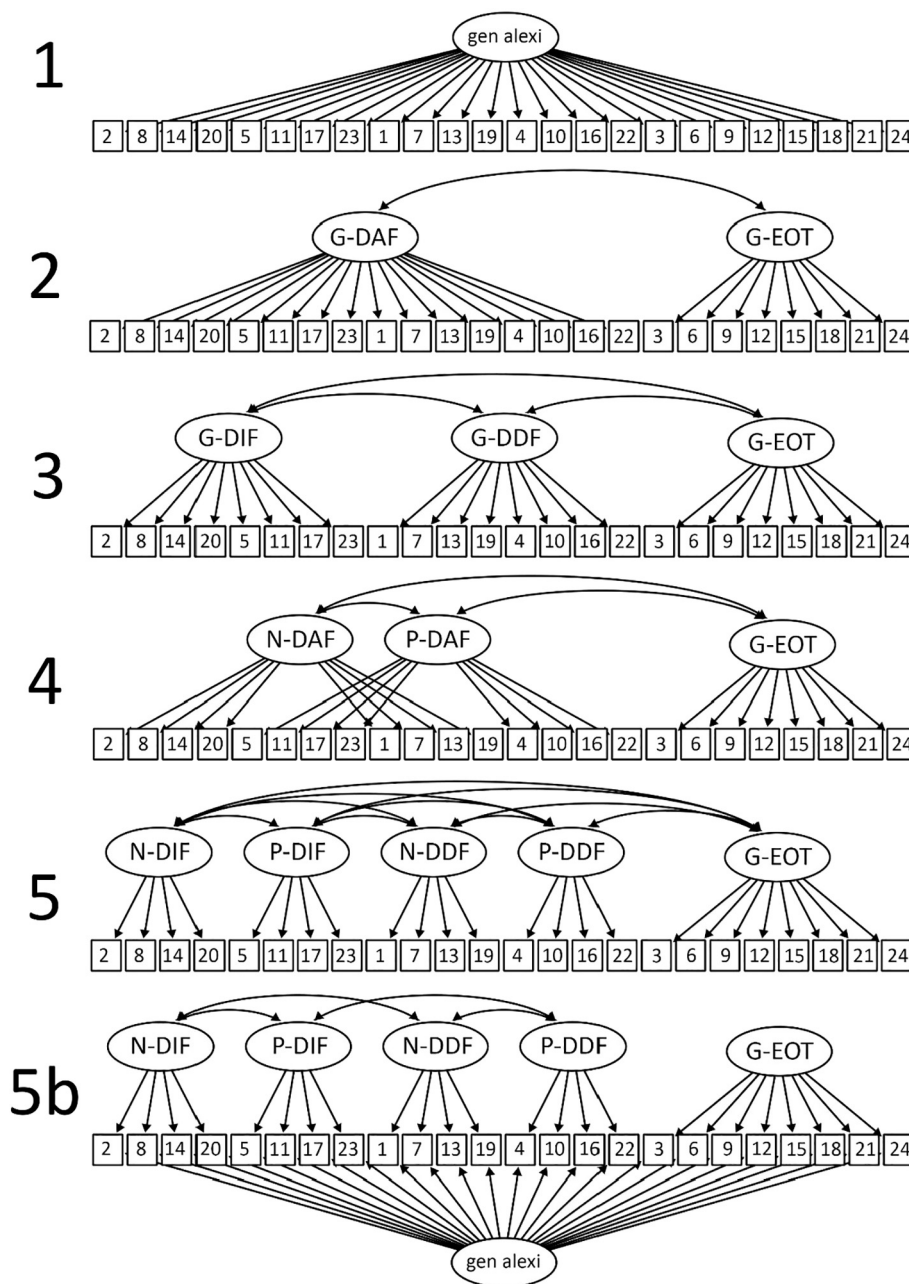
Fig. 1. The confirmatory factor analysis models assessed in Study 1 and Study 2; Models 1, 2, 3, 4, 5, and 5b. Squares indicate item numbers, ellipses indicate latent factors. Item error terms are not displayed. N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking, G-DIF = General-Difficulty identifying feelings, G-DDF = General-Difficulty describing feelings, N-DAF = Negative-Difficulty appraising feelings, P-DAF = Positive-Difficulty appraising feelings, G-DAF = General-Difficulty appraising feelings, gen alexi = general alexithymia.

avoidant defences are likely to be applied more so to negative feelings, so we expected that people would report more difficulties appraising negative feelings than positive feelings.

2.1.3.2. Factor structure. The factor structure of the 24-item PAQ was examined via a series of CFAs. CFAs were conducted using maximum likelihood estimation based on a Pearson covariance matrix. The 24 retained PAQ items were used as the observed variables. Six theoretically informed models of increasing complexity were examined (see Fig. 1).

(1) Model 1 was a 1-factor model, where all 24 items were specified to load on a "general alexithymia" factor. (2) Model 2 was a 2-factor correlated model, where a distinction between the *attention* and *appraisal* stages of emotion valuation was made; items were specified to load on "G-EOT" or "G-DAF" factors. These two factors were allowed to correlate. (3) Model 3 was a 3-factor correlated model, where a distinction was made between the DIF, DDF and EOT components of alexithymia, but no distinction was made based on valence; items were

specified to load on "G-EOT", "G-DIF" or "G-DDF" factors. All these factors were allowed to correlate. (4) Model 4 was an alternate 3-factor correlated model, where a distinction was made based on valence, but the DIF and DDF components of alexithymia were not separated; items were specified to load on "G-EOT", "N-DAF" or "P-DAF" factors. All these factors were allowed to correlate. (5) Model 5 was a 5-factor correlated model that reflected the intended subscale structure of the PAQ, whereby a distinction was made based on valence, and a distinction was made between the DIF, DDF and EOT components of alexithymia; items were specified to load on "G-EOT", "N-DIF", "P-DIF", "N-DDF" or "P-DDF" factors. All these factors were allowed to correlate. (6) Model 5b was a bifactor model version of Model 5 (Gignac et al., 2007; Reise, 2012). All items were specified to load on a broad "general alexithymia" factor, as well as load on one of five narrow factors ("G-EOT", "N-DIF", "P-DIF", "N-DDF" or "P-DDF") that corresponded to the five intended subscales. Those narrow factors designed to measure the same emotional valence were allowed to correlate (i.e., "N-DIF" and "N-DDF"; "P-DIF" and "P-DDF"), as were those narrow factors designed to

Table 3
Descriptive statistics and Cronbach's alpha reliability coefficients for the administered measures in Study 1 and Study 2.

Measure/subscale	Total				Females		Males	
	M	SD	range	α	M	SD	M	SD
Study 1								
PAQ								
Subscales								
N-DIF	13.51	5.87	4–28	.87	13.79	6.12	12.96	5.37
P-DIF	10.64	5.10	4–28	.88	10.55	5.37	10.81	4.57
N-DDF	15.78	6.67	4–28	.90	15.92	7.00	15.50	6.04
P-DDF	12.53	6.01	4–28	.91	12.15	6.03	13.25	5.94
G-EOT	26.43	10.65	8–56	.90	25.43	10.83	28.31	10.08
Composites								
G-DIF	24.15	9.93	8–56	.90	24.34	10.46	23.78	8.90
G-DDF	28.30	11.57	8–56	.92	28.07	11.99	28.75	10.78
N-DAF	29.28	11.72	8–56	.93	29.72	12.33	28.46	10.50
P-DAF	23.17	10.52	8–56	.93	22.70	10.98	24.06	9.58
G-DAF	52.45	20.29	16–112	.95	52.41	21.46	52.53	18.00
ALEXI	78.88	28.34	24–168	.95	77.84	29.96	80.84	25.06
Study 2								
PAQ								
Subscales								
N-DIF	13.38	6.41	4–28	.89	14.03	6.57	12.28	5.98
P-DIF	11.30	5.76	4–28	.89	11.38	6.08	11.18	5.20
N-DDF	15.35	6.89	4–28	.91	15.80	7.00	14.60	6.63
P-DDF	12.97	6.12	4–28	.90	12.72	6.24	13.38	5.90
G-EOT	28.97	11.19	8–56	.90	27.92	11.35	30.74	10.72
Composites								
G-DIF	23.68	11.28	8–56	.92	25.41	11.57	23.45	10.68
G-DDF	28.32	12.16	8–56	.93	28.52	12.30	27.98	11.92
N-DAF	28.73	12.71	8–56	.94	29.83	13.07	26.88	11.89
P-DAF	24.27	11.39	8–56	.94	24.10	11.95	24.55	10.41
G-DAF	52.99	22.58	16–112	.96	53.93	23.23	51.43	21.39
ALEXI	81.97	30.91	24–168	.96	81.84	31.92	82.17	29.15
DASS-21								
Depression	5.87	5.89	0–21	.94	6.63	6.14	4.59	5.21
Anxiety	4.40	4.85	0–21	.89	4.92	5.06	3.54	4.34
Stress	6.17	5.38	0–21	.92	6.99	5.55	4.80	4.79
Total scale	16.44	14.86	0–63	.96	18.54	15.37	12.93	13.28
ERQ								
Reappraisal	28.80	7.19	6–42	.89	28.95	7.60	28.56	6.46
Suppression	15.78	5.28	4–28	.78	15.33	5.57	16.54	4.68

Note. Study 1 data includes 231 participants (151 females, 80 males) and Study 2 data includes 748 participants (468 females, 280 males). PAQ = Perth Alexithymia Questionnaire, DASS-21 = Depression Anxiety Stress Scales-21, ERQ = Emotion Regulation Questionnaire, N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking, G-DIF = General-Difficulty identifying feelings, G-DDF = General-Difficulty describing feelings, N-DAF = Negative-Difficulty appraising feelings, P-DAF = Positive-Difficulty appraising feelings, G-DAF = General-Difficulty appraising feelings, ALEXI = alexithymia.

measure the same component of appraisal (i.e., "N-DIF" and "P-DIF"; "N-DDF" and "P-DDF"). Bifactor models are appropriate when a measure is designed to "primarily reflect a strong common trait [e.g., alexithymia], but there is multidimensionality caused by well-defined clusters of items from diverse subdomains" (Reise, 2012, p. 692), which is the case for the PAQ, so we expected Model 5b would be the best fitting solution.

The goodness-of-fit of the CFA models was judged based on the pattern of factor loadings and factor intercorrelations within each model, and via four fit indexes: the comparative fit index (CFI), Tucker Lewis index (TLI), root mean square error of approximation (RMSEA), and standardised root mean residual (SRMR). CFI and TLI values ≥ 0.90 were judged to indicate acceptable fit, as were RMSEA and SRMR values ≤ 0.08 (Bentler & Bonett, 1980; Browne & Cudeck, 1992; Marsh, Hau, & Wen, 2004). To directly compare the fit of the models, the

Akaike Information Criterion (AIC) was also used; AIC penalises more complex models, and lower values indicate a better fitting model (Byrne, 2013). Factor loadings ≥ 0.40 were considered meaningful loadings (Stevens, 1992).

2.1.3.3. Internal consistency reliability. Cronbach's alpha reliability coefficients were calculated. Reliability coefficients ≥ 0.70 were considered acceptable, ≥ 0.80 were considered good, and ≥ 0.90 were considered excellent (Groth-Marnat, 2009).

2.2. Results and discussion

2.2.1. Descriptive statistics

Descriptive statistics are reported in Table 3. As predicted, participants reported significantly more difficulties appraising negative feelings (N-DAF; $M = 29.28$, $SD = 11.72$) than positive feelings (P-DAF; $M = 23.17$, $SD = 10.52$), $t(230) = 10.12$, $p < 0.001$.

2.2.2. Factor structure

All CFA fit indexes indicated that those models featuring the five intended subscales as factors (i.e., Model 5 and Model 5b) were a good fit to the data and the best models we tested. For CFA fit index values, factor loadings, and factor intercorrelations, see Tables 4, 5, and 6, respectively.

The one-factor model (Model 1) was a poor fit to the data according to all examined fit indexes, suggesting that the PAQ was measuring a multidimensional construct. The two-factor model (Model 2) displayed better fit than Model 1, indicating that making a broad distinction between the *attention* and *appraisal* stages of emotion valuation added value to the factor solution. In turn, Model 3, which further separated the appraisal stage into its DIF and DDF components, displayed better fit than Model 2. Thus, whilst the "G-DIF" and "G-DDF" factors in Model 3 were very highly correlated (estimated $r = 0.86$), there was some statistical value in separating them. Overall though, the models that did not account for emotional valence (i.e., Model 1, Model 2, Model 3) exhibited poor fit index values. Our CFAs therefore emphasised that it was statistically useful to distinguish between the appraisal of *negative* emotions and the appraisal *positive* emotions. Model 4, for example, produced much better fit indexes than Model 3, indicating that, with respect to the appraisal stage, it was statistically more important to account for valence than it was to distinguish between DIF and DDF. The superiority of Model 5 over Models 3 and 4, nonetheless, highlighted that making *all* these distinctions (i.e., distinguishing between the attention and appraisal stages of emotion valuation, distinguishing between the DIF and DDF components of appraisal, and distinguishing between the appraisal of negative and positive emotions) added value to the factor solution. All items loaded well on their intended factor in Model 5, and all five factors were strongly positively correlated (estimated rs = 54 to 0.88). In Model 5b, the addition of the broad "general alexithymia" factor into the factor solution improved fit further. Item variance was generally split between the broad factor and the narrow factors, with most items loading more so on the broad factor (see Table 5). There was therefore statistical support for the presence of a strong "general alexithymia" factor, as well as statistical support for five narrow factors reflecting the subdomains (i.e., the subscales) of the multidimensional alexithymia construct (Reise, 2012). The 24-item PAQ, hence, had a factor structure that was consistent with its theoretical basis in this data-set.

2.2.3. Internal consistency reliability

As displayed in Table 3, all subscale scores had good to excellent internal consistency reliability (α = 0.87 to 0.91) and all composite scores had excellent internal consistency reliability (α = 0.90 to 0.95).

Table 4
Goodness-of-fit index values from confirmatory factor analyses of the 24 PAQ items in Study 1 and Study 2.

Model	χ² (df)	CFI	TLI	RMSEA (90% CI)	SRMR	AIC
			Study 1			
Model 1	1611.636 (252)	.671	.639	.153 (.146-.160)	.1041	1707.636
Model 2	1191.319 (251)	.772	.750	.128 (.120-.135)	.0835	1289.319
Model 3	1059.520 (249)	.804	.782	.119 (.112-.126)	.0782	1161.520
Model 4	739.385 (249)	.881	.868	.093 (.085-.100)	.0618	841.385
Model 5	479.035 (242)	.943	.935	.065 (.057-.074)	.0504	595.035
Model 5b	441.586 (224)	.947	.935	.065 (.056-.074)	.0550	593.586
			Study 2			
Model 1	4067.055 (252)	.731	.705	.142 (.139-.146)	.0938	4163.055
Model 2	2672.916 (251)	.829	.812	.114 (.110-.118)	.0627	2770.916
Model 3	2452.332 (249)	.845	.828	.109 (.105-.113)	.0607	2554.332
Model 4	1584.897 (249)	.906	.896	.085 (.081-.089)	.0469	1686.897
Model 5	1164.626 (242)	.935	.926	.071 (.067-.076)	.0422	1280.626
Model 5b	1007.574 (224)	.945	.932	.068 (.064-.073)	.0381	1159.574

Note. For all examined models χ² $p < 0.001$. CFI = comparative fit index, TLI = Tucker Lewis index, RMSEA = root mean square error of approximation, SRMR = standardised root mean residual, AIC = Akaike information criterion, CI = confidence interval.

Table 5
Standardised item factor loadings from confirmatory factor analyses of the 24 PAQ items in Study 1 and Study 2; loadings are displayed for Model 5 and Model 5b.

Factor/item	Study 1		Study 2	
	Model 5	Model 5b	Model 5	Model 5b
N-DIF				
2-When I'm feeling *bad*, I can't tell whether I'm sad, angry, or scared.	.72	.48(.53)	.78	.57(.55)
8-When I'm feeling *bad*, I can't make sense of those feelings.	.82	.55(.62)	.84	.58(.62)
14- When I'm feeling *bad*, I get confused about what emotion it is.	.88	.59(.66)	.88	.52(.70)
20-When I'm feeling *bad*, I'm puzzled by those feelings.	.77	.58(.52)	.80	.50(.62)
P-DIF				
5-When I'm feeling *good*, I can't tell whether I'm happy, excited, or amused.	.66	.56(.36)	.77	.60(.59)
11-When I'm feeling *good*, I can't make sense of those feelings.	.86	.68(.53)	.84	.38(.75)
17- When I'm feeling *good*, I get confused about what emotion it is.	.89	.67(.57)	.85	.37(.75)
23- When I'm feeling *good*, I'm puzzled by those feelings.	.79	.63(.45)	.83	.23(.80)
N-DDF				
1-When I'm feeling *bad* (feeling an unpleasant emotion), I can't find the right words to describe those feelings.	.79	.22(.76)	.82	.53(.64)
7-When I'm feeling *bad*, I can't talk about those feelings in much depth or detail.	.85	.14(.83)	.82	.43(.68)
13-When something *bad* happens, it's hard for me to put into words how I'm feeling.	.86	.02*(.87)	.88	.53(.70)
19-When I'm feeling *bad*, if I try to describe how I'm feeling I don't know what to say.	.88	.26(.84)	.89	.48(.74)
P-DDF				
4- When I'm feeling *good* (feeling a pleasant emotion), I can't find the right words to describe those feelings.	.79	.50(.60)	.78	.60(.63)
10- When I'm feeling *good*, I can't talk about those feelings in much depth or detail.	.86	.51(.69)	.81	.11(.82)
16- When something *good* happens, it's hard for me to put into words how I'm feeling.	.85	.54(.66)	.86	.31(.80)
22- When I'm feeling *good*, if I try to describe how I'm feeling I don't know what to say.	.90	.54(.72)	.87	.27(.82)
G-EOT				
3-I tend to ignore how I feel.	.68	.35(.59)	.71	.46(.53)
6-I prefer to just let my feelings happen in the background, rather than focus on them.	.75	.59(.48)	.67	.50(.44)
9-I don't pay attention to my emotions.	.56	.45(.35)	.80	.67(.50)
12-Usually, I try to avoid thinking about what I'm feeling.	.86	.61(.60)	.77	.44(.62)
15-I prefer to focus on things I can actually see or touch, rather than my emotions.	.78	.56(.55)	.69	.47(.50)
18-I don't try to be 'in touch' with my emotions.	.83	.67(.52)	.77	.57(.53)
21-It's not important for me to know what I'm feeling.	.59	.46(.38)	.64	.46(.45)
24-It's strange for me to think about my emotions.	.78	.47(.62)	.81	.50(.65)

Note. For the bifactor model (Model 5b), factor loadings inside the brackets are loadings on the broad "general alexithymia" factor, and factor loadings outside the brackets are loadings on the narrow "N-DIF", "N-DDF", "P-DIF", "P-DDF" or "G-EOT" factor. N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking.

* $p > .05$.

3. Study 2

3.1. Method

3.1.1. Participants and procedure

To replicate and extend the results of Study 1, the 24-item PAQ was subsequently administered to a new sample. This sample was comprised of 748 adults (62.6% females) with a mean age of 47.57 years ($SD = 17.29$, range = 18–88).[5] Participants were recruited via an online survey recruiting company (Qualtrics panels). All participants were English speaking residents of Australia, with 75% born in Australia. The

---

[5] Some additional participants also completed the online survey, however their data was excluded during quality screening because they failed an attention check question and/or completed the survey impossibly quickly (i.e., at a rate of < 2 s per question, suggesting inattentive responding).

Table 6

Estimated correlations between the PAQ factors in the examined confirmatory factor analysis models in Study 1 and Study 2.

	Factor				
	F1	F2	F3	F4	F5
Model 2					
F1 "G-DAF"	-	.68*	-	-	-
F2 "G-EOT"	.69*	-	-	-	-
Model 3					
F1 "G-DIF"	-	.92*	.64*	-	-
F2 "G-DDF"	.86*	-	.69*	-	-
F3 "G-EOT"	.62*	.70*	-	-	-
Model 4					
F1 "N-DAF"	-	.80*	.63*	-	-
F2 "P-DAF"	.71*	-	.66*	-	-
F3 "G-EOT"	.65*	.63*	-	-	-
Model 5					
F1 "N-DIF"	-	.80*	.92*	.74*	.58*
F2 "P-DIF"	.73*	-	.71*	.93*	.63*
F3 "N-DDF"	.84*	.56*	-	.82*	.65*
F4 "P-DDF"	.61*	.88*	.73*	-	.67*
F5 "G-EOT"	.54*	.59*	.68*	.62*	-
Model 5b					
F1 "N-DIF"	-	.36*	.78*	-	-
F2 "P-DIF"	.52*	-	-	.70*	-
F3 "N-DDF"	.95*	-	-	.26*	-
F4 "P-DDF"	-	.80*	-.28	-	-
F5 "G-EOT"	-	-	-	-	-
F6 "gen alexi"	-	-	-	-	-

Note. Values below the diagonal are for Study 1, those above the diagonal are for Study 2. Model 5b was a bifactor model, correlations were not specified between some of the factors. PAQ = Perth Alexithymia Questionnaire, N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking, G-DIF = General-Difficulty identifying feelings, G-DDF = General-Difficulty describing feelings, N-DAF = Negative-Difficulty appraising feelings, P-DAF = Positive-Difficulty appraising feelings, G-DAF = General-Difficulty appraising feelings, gen alexi = general alexithymia.

  * $p < .05$.

highest level of completed education for 37% of the sample was high school, for 34.6% it was a technical diploma, and for 26.8% it was a university degree. About one tenth (9.4%) of the sample were current university students. The PAQ was administered as part of a battery of psychological questionnaires in an anonymous online survey.

3.1.2. Materials

The battery included the PAQ, the ERQ (Gross & John, 2003), and the Depression Anxiety Stress Scales-21 (DASS-21; Lovibond & Lovibond, 1995).

3.1.2.1. Emotion Regulation Questionnaire. The ERQ (Gross & John, 2003) is a 10-item self-report measure of people's usage of two emotion regulation strategies; *Cognitive reappraisal* (6 items, e.g., "When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm") and *Expressive suppression* (4 items, e.g., "I control my emotions by not expressing them"). Items are answered on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) with higher scores indicating greater use of that strategy. Cognitive reappraisal is usually associated with adaptive outcomes, and expressive suppression is associated with maladaptive outcomes; hence, low cognitive reappraisal scores and high expressive suppression scores are indicative of emotion regulation difficulties (Gross & John, 2003). The ERQ has demonstrated good validity and reliability (Gross & John, 2003).

3.1.2.2. Depression Anxiety Stress Scales-21. The DASS-21 (Lovibond & Lovibond, 1995) is a 21-item self-report measure of *Depression* (7 items, e.g., "I felt that life was meaningless"), *Anxiety* (7 items, e.g., "I felt I was close to panic"), and *Stress* (7 items, e.g., "I found it hard to wind down") symptoms experienced in the past week. All items can be summed into a *Total scale* score representing overall levels of psychological distress (Kia-Keating et al., 2017). Items are answered on a 4-point Likert scale ranging from 0 to 3, with higher scores indicating more severe symptomatology. The DASS-21 has demonstrated good validity and reliability (Lovibond & Lovibond, 1995).

3.1.3. Analytic strategy

Descriptive statistics were calculated, and N-DAF and P-DAF scores were compared in the same manner as Study 1. All 24 PAQ items were reasonably normally distributed (max skewness = 0.82, max kurtosis = −1.23).

3.1.3.1. Factor structure and internal consistency reliability. The factor structure and internal consistency reliability of the PAQ was examined in the same manner as Study 1.

3.1.3.2. Concurrent validity. Pearson correlations were calculated between PAQ scores and ERQ/DASS-21 scores. Because alexithymia is a deficit in the emotion valuation process, and this valuation process is responsible for activating emotion regulation attempts, alexithymia is a "crucial rate-limiting factor" for successful emotion regulation (Gross, 2014, p. 13) so we expected that high PAQ scores would be associated with a more maladaptive emotion regulation profile on the ERQ (i.e., lower use of cognitive reappraisal, higher use of expressive suppression) and higher levels of depression, anxiety, and stress on the DASS-21.

3.1.3.3. Discriminant validity. Discriminant validity was examined via a second-order EFA (principal axis factoring using direct oblimin rotation) using the PAQ and DASS-21 subscale scores as the observed variables. Factors were extracted based on the scree-plot method (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Our purpose here was to establish whether the PAQ was measuring a latent construct (i.e., alexithymia) that was statistically separable from people's current level of psychological distress. As abovementioned, people's overall level of alexithymia should correlate with their current level of distress, but conceptually alexithymia and distress are still separable constructs (Preece, Becerra, Allan, et al., 2017; Taylor et al., 1999). We therefore expected all five PAQ subscales to load on a "general alexithymia" factor and all three DASS-21 subscales to load on a separate "psychological distress" factor. This analysis was motivated by results from several recent studies (e.g., Leising, Grande, & Faber, 2009; Marchesi, Ossola, Tonna, & De Panfilis, 2014), which found that some existing alexithymia measures have poor discriminant validity with measures of distress.

3.2. Results and discussion

3.2.1. Descriptive statistics

Descriptive statistics are reported in Table 3. Like Study 1, consistent with our expectations, participants reported significantly more difficulties appraising negative feelings (N-DAF; $M = 28.73$, $SD = 12.71$) than positive feelings (P-DAF; $M = 24.27$, $SD = 11.39$), t (747) = 14.26, $p < 0.001$.

3.2.2. Factor structure

The pattern of CFA findings was the same as Study 1. CFAs of the 24 PAQ items indicated that Model 5 and Model 5b were, again, good fits to the data and the best of the models we tested. For CFA fit index values, factor loadings, and factor intercorrelations, see Tables 4, 5, and 6, respectively. In Model 5, all items loaded strongly on their intended

Table 7

Pearson correlations between the PAQ and DASS-21/ERQ in Study 2.

Measure/ subscale	DASS-21				ERQ	
	Depression	Anxiety	Stress	Total	Reappraisal	Suppression
PAQ						
Subscales						
N-DIF	.51*	.52*	.54*	.57*	-.17*	.33*
P-DIF	.44*	.46*	.43*	.48*	-.17*	.38*
N-DDF	.50*	.44*	.48*	.51*	-.14*	.41*
P-DDF	.43*	.40*	.38*	.44*	-.12*	.46*
G-EOT	.28*	.24*	.21*	.26*	-.06	.55*
Composites						
G-DIF	.51*	.53*	.53*	.57*	-.18*	.38*
G-DDF	.50*	.45*	.47*	.51*	-.14*	.46*
N-DAF	.53*	.50*	.53*	.56*	-.16*	.39*
P-DAF	.45*	.45*	.43*	.48*	-.15*	.44*
G-DAF	.52*	.51*	.52*	.56*	-.17*	.44*
ALEXI	.48*	.45*	.45*	.50*	-.14*	.52*

Note. PAQ = Perth Alexithymia Questionnaire, DASS-21 = Depression Anxiety Stress Scales-21, ERQ = Emotion Regulation Questionnaire, N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking, G-DIF = General-Difficulty identifying feelings, G-DDF = General-Difficulty describing feelings, N-DAF = Negative-Difficulty appraising feelings, P-DAF = Positive-Difficulty appraising feelings, G-DAF = General-Difficulty appraising feelings, ALEXI = alexithymia.

* $p < .05$.

factor and all five factors were significantly positively correlated. Model 5b produced the best fit index values overall, suggesting that the factor structure of the PAQ was well represented by a broad "general alexithymia" factor, and five narrow factors corresponding to the intended subscales of the PAQ. The PAQ therefore had a factor structure that was consistent with its theoretical basis in this data-set.

3.2.3. Internal consistency reliability

Like Study 1, all subscale scores had good to excellent levels of internal consistency reliability (α = 0.89 to 0.91) and all composite scores had excellent internal consistency reliability (α = 0.92 to 0.96; see Table 3).

Table 8

Factor loadings from a second-order exploratory factor analysis of the PAQ and DASS-21 subscale scores in Study 2 to examine discriminant validity.

Measure/subscale	Factor 1 "general alexithymia"	Factor 2 "psychological distress"
PAQ		
N-DIF	**.72**	.22
P-DIF	**.82**	.06
N-DDF	**.80**	.11
P-DDF	**.90**	-.03
G-EOT	**.74**	-.12
DASS-21		
Depression	.08	**.80**
Anxiety	.01	**.88**
Stress	-.03	**.94**

Note. Principal axis factoring with direct oblimin rotation was used. Factor loadings ≥0.40 are in boldface. The correlation between factor 1 and factor 2 was 0.52. PAQ = Perth Alexithymia Questionnaire, DASS-21 = Depression Anxiety Stress Scales-21, N-DIF = Negative-Difficulty identifying feelings, P-DIF = Positive-Difficulty identifying feelings, N-DDF = Negative-Difficulty describing feelings, P-DDF = Positive-Difficulty describing feelings, G-EOT = General-Externally orientated thinking.

3.2.4. Concurrent validity

Supporting the concurrent validity of the PAQ, people reporting higher levels of alexithymia on the PAQ also tended to report more emotion regulation problems on the ERQ (i.e., higher usage of expressive suppression, lower usage of cognitive reappraisal) and a higher level of depression, anxiety, and stress symptoms on the DASS-21 (see Table 7).

3.2.5. Discriminant validity

Consistent with our expectations, the second-order EFA of the PAQ and DASS-21 subscale scores extracted two factors, accounting for 78.91% of the variance in subscale scores (see Table 8). Factor 1 (eigenvalue = 4.86), which we call "general alexithymia", consisted of all five PAQ subscales. Factor 2 (eigenvalue = 1.46), which we call "psychological distress", consisted of all three DASS-21 subscales. All subscale scores loaded strongly on their primary factor with no crossloadings, suggesting that the PAQ subscale scores were successfully measuring a latent construct that was statistically separable from people's current level of psychological distress.

4. General discussion

Our purpose in this paper was to document the development of the PAQ and examine its psychometric properties across two studies. In both studies, the PAQ performed well on every marker of validity and reliability that we tested.

The factor structure of the PAQ was replicable and consistent with its theoretical basis (Preece, Becerra, Allan, et al., 2017). All 24 items loaded cleanly onto one of five narrow factors, which corresponded to the five subscales we intended the measure to have (N-DIF, P-DIF, N-DDF, P-DDF, G-EOT), and there was also strong evidence in our bifactor CFA models for a broader "general alexithymia" factor. The PAQ therefore appeared to assess a coherent multidimensional construct, and could distinguish between the DIF, DDF and EOT components of alexithymia. The DIF and DDF items were, moreover, particularly closely linked in our factor analyses, and thus a broader distinction between the attention (EOT) and appraisal (DIF, DDF) stages of emotion valuation was evident (see also, Erni, Lötscher, & Modestin, 1997; Kooiman et al., 2002; Loas, Otmani, Verrier, Fremaux, & Marchand, 1996). Our factor analyses suggested, in fact, with respect to measuring the appraisal stage, that it was statistically more important to distinguish between negative and positive valence, than it was to distinguish between DIF and DDF (see also, Feldman Barrett et al., 2001; van der Velde et al., 2013). Indeed, in both our data-sets, people tended to report significantly more difficulties appraising negative emotions than positive emotions. The PAQ therefore appears to provide a more nuanced assessment of alexithymia than existing measures that do not account for emotional valence.

The PAQ also correlated with established measures of emotion regulation and psychopathology in expected ways. High PAQ scores were associated with higher usage of expressive suppression, lower usage of cognitive reappraisal, and higher levels of depression, anxiety and stress symptoms. The PAQ subscales, moreover, loaded on a different higher-order factor to the DASS-21 subscales when factor analysed, and thereby demonstrated a form of discriminant validity (i.e., the PAQ is not a measure of people's current level of distress; see Leising et al., 2009). Whilst our conclusions about the PAQ's clinical relevance must be tentative given that we did not use a clinical sample, these results are consistent with contemporary models of psychopathology (e.g., Ellard, Fairholme, Boisseau, Farchione, & Barlow, 2010; Rottenberg & Johnson, 2007; Taylor et al., 1999), which consider alexithymia as a key risk factor for the development of psychiatric symptomatology. High levels of internal consistency reliability were also evident for all PAQ subscale and composite scores, and the

excellent internal consistency reliability of the PAQ G-EOT subscale is of particular note given that some previous attempts to measure EOT via self-report have had low reliability. In a similar Australian sample to the two analysed here, for example, Preece, Becerra, Robinson, and Dandy (2017) found that the TAS-20 EOT subscale had a low Cronbach's alpha reliability coefficient of 0.64, and half the TAS-20 EOT items did not load meaningfully on the EOT factor in factor analysis (see also, Gignac et al., 2007; Thorberg et al., 2010).

We think future research with the PAQ could consequently help enhance theoretical understanding of the alexithymia construct and its relationship to other variables. Clinical researchers are, for example, increasingly trying to examine alexithymia at the component level and establish whether different psychopathology categories (e.g., Bankier, Aigner, & Bach, 2001; Leweke et al., 2012; Lyvers et al., 2018; Marchesi et al., 2014) or brain injury categories (e.g., Williams & Wood, 2010) have characteristic DIF, DDF and EOT profiles. To date, however, such work has primarily used the TAS-20 and thus confident inferences about EOT, or valence-specific deficits in DIF and DDF, have not been possible. Our data suggest that by using the PAQ more detailed alexithymia profiles may now be obtained.[6] Pending validation in clinical populations, such profiles could, in effect, help researchers and clinicians to better conceptualise and treat those clinical presentations characterised by emotion processing deficits. Patients reporting high levels of alexithymia on the PAQ are, for instance, likely to be good candidates for psychotherapeutic interventions (e.g., Edwards, Shivaji, & Wupperman, 2018; Neumann, Malec, & Hammond, 2017) that are designed to develop emotion schemas and reduce usage of experiential avoidance as an emotion regulation strategy (Preece, Becerra, Allan, et al., 2017).

4.1. Limitations and some future directions

We think the introduction of the PAQ makes a useful contribution, but some limitations of our studies should be noted that will require further research. Firstly, the size of our sample in Study 1 was large enough for the factor analyses to be robust according to commonly accepted criteria (e.g., a minimum of 100 participants and at least 5 participants per variable in the analysis; Gorsuch, 1983; Kline, 1979), but it was a modest sample size. Secondly, we designed the PAQ with adult, adolescent, nonclinical and clinical respondents in mind, but our results here only apply to adults from the general community, so the measure's performance in clinical and adolescent samples remains to be determined. Thirdly, we did not examine the test-retest reliability of the PAQ, so more research is required to determine the extent to which PAQ scores are consistent over time. Fourthly, the concurrent validity of the PAQ was explored only in relation to other self-report measures and we did not administer any other measures of alexithymia. We ideally would have included the TAS-20 and BVAQ in our Study 2 questionnaire battery, but extending the length of the battery further was not feasible for this study. Future work should evaluate the PAQ directly against other measures of alexithymia, which would allow for more detailed analyses of convergent and discriminant validity. We think factor analyses of the PAQ, TAS-20 and BVAQ items together, for example, could be particularly useful for determining how the non-valenced DIF and DDF items of the TAS-20 and BVAQ are interpreted by participants; that is, whether these items are interpreted as corresponding more so to negative valence, positive valence, or some aggregate of both valence types. Such data

would help researchers to better frame and understand the existing body of alexithymia work that has so far mostly used the TAS-20 and BVAQ.

4.2. Conclusions

Our data suggest that the PAQ has good validity and reliability as a self-report measure of alexithymia. Strengths of the measure include its capacity to reliably measure each component of the construct as it is defined by the attention-appraisal model, and its capacity to do so across negative and positive emotions. Whilst further research is needed to confirm these findings across different population types, on present evidence, the PAQ is appears to be a useful alexithymia assessment tool.

Ethical statement

Ethics approval for this project was granted by the Edith Cowan University Human Research Ethics Committee. The guidelines of this committee were followed. All participants provided informed consent for their data to be used.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2018.05.011.

References

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*, 23–32.

Bagby, R. M., Taylor, G. J., Parker, J. D., & Dickens, S. E. (2006). The development of the Toronto structured interview for alexithymia: Item selection, factor structure, reliability and concurrent validity. *Psychotherapy and Psychosomatics, 75*, 25–39.

Bagby, R. M., Taylor, G. J., Quilty, L. C., & Parker, J. D. (2007). Reexamining the factor structure of the 20-item Toronto Alexithymia Scale: Commentary on Gignac, palmer, and Stough. *Journal of Personality Assessment, 89*, 258–264.

Bankier, B., Aigner, M., & Bach, M. (2001). Alexithymia in DSM-IV disorder: Comparative evaluation of somatoform disorder, panic disorder, obsessive-compulsive disorder, and depression. *Psychosomatics, 42*, 235–240.

Bardeen, J. R., Fergus, T. A., Hannan, S. M., & Orcutt, H. K. (2016). Addressing psychometric limitations of the difficulties in emotion regulation scale through item modification. *Journal of Personality Assessment, 98*, 298–309.

Becerra, R., Amos, A., & Jongenelis, S. (2002). Organic alexithymia: A study of acquired emotional blindness. *Brain Injury, 16*, 633–645.

Becerra, R., Preece, D., Campitelli, G., & Scott-Pillow, G. (2017). The assessment of emotional reactivity across negative and positive emotions: Development and validation of the Perth emotional reactivity scale (PERS). *Assessment,* 1–13 (doi: 1073191117694455).

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606.

Berenbaum, H. (1996). Childhood abuse, alexithymia and personality disorder. *Journal of Psychosomatic Research, 41*, 585–595.

Bermond, B., Clayton, K., Liberova, A., Luminet, O., Maruszewski, T., Ricci Bitti, P. E., ... Wicherts, J. (2007). A cognitive and an affective dimension of alexithymia in six languages and seven populations. *Cognition and Emotion, 21*, 1125–1136.

Bermond, B., Vorst, H. C., Vingerhoets, A. J., & Gerritsen, W. (1999). The Amsterdam alexithymia scale: Its psychometric values and correlations with other personality traits. *Psychotherapy and Psychosomatics, 68*, 241–251.

Bilotta, E., Giacomantonio, M., Leone, L., Mancini, F., & Coriale, G. (2015). Being alexithymic: Necessity or convenience. Negative emotionality × avoidant coping interactions and alexithymia. *Psychology and Psychotherapy: Theory, Research and Practice, 89*, 261–275.

Bradley, M. M., & Lang, P. J. (2007). The international affective picture system (IAPS) in the study of emotion and attention. In J. A. Coan, & J. J. B. Allen (Eds.). *Handbook of emotion elicitation and assessment* (pp. 29–46). New York, NY: Oxford University Press.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*, 230–258.

Byrne, B. M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* New York, NY: Psychology Press.

Coriale, G., Bilotta, E., Leone, L., Cosimi, F., Porrari, R., De Rosa, F., & Ceccanti, M. (2012). Avoidance coping strategies, alexithymia and alcohol abuse: A mediation analysis. *Addictive Behaviors, 37*, 1224–1229.

Duddu, V., Isaac, M. K., & Chaturvedi, S. K. (2003). Alexithymia in somatoform and depressive disorders. *Journal of Psychosomatic Research, 54*, 435–438.

Edwards, E., Shivaji, S., & Wupperman, P. (2018). The emotion mapping activity: Preliminary evaluation of a mindfulness-informed exercise to improve emotion labeling in alexithymic persons. *Scandinavian Journal of Psychology,* 1–9. http://dx.doi.

---

[6] Because alexithymia is, statistically, a dimensional trait rather than a categorical syndrome (e.g., Mattila et al., 2010; Parker, Keefer, Taylor, & Bagby, 2008), to interpret PAQ profiles we think it is most appropriate for a respondent's scores be compared to those of an appropriate normative sample. Following common assessment practices (e.g., Lezak, Howieson, & Loring, 2004), scores within one standard deviation of the population mean could be seen to represent an "average" level of alexithymia, whilst scores below or above this range could be seen to represent a "low" or "high" level of alexithymia, respectively.

org/10.1111/sjop.12438.

Ellard, K. K., Fairholme, C. P., Boisseau, C. L., Farchione, T. J., & Barlow, D. H. (2010). Unified protocol for the transdiagnostic treatment of emotional disorders: Protocol development and initial outcome data. *Cognitive and Behavioral Practice, 17*, 88–101.

Erni, T., Lötscher, K., & Modestin, J. (1997). Two-factor solution of the 20-ltem Toronto Alexithymia Scale confirmed. *Psychopathology, 30*, 335–340.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299.

Feldman Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion, 15*, 713–724.

Fukunishi, I., Yoshida, H., & Wogan, J. (1998). Development of the alexithymia scale for children: A preliminary study. *Psychological Reports, 82*, 43–49.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research, 25*, 186–192.

Gignac, G. E., Palmer, B. R., & Stough, C. (2007). A confirmatory factor analytic investigation of the TAS–20: Corroboration of a five-factor model and suggestions for improvement. *Journal of Personality Assessment, 89*, 247–257.

Gori, A., Giannini, M., Palmieri, G., Salvini, R., & Schuldberg, D. (2012). Assessment of alexithymia: Pychometric properties of the psychological treatment inventory-alexithymia scale (PTI-AS). *Psychology, 3*, 231.

Gorsuch, R. L. (1983). *Factor analysis.* Hillsdale, NJ: Erlbaum.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41–54.

Gross, J. J. (2014). Emotion regulation: Conceptual and empirical foundations. In J. J. Gross (Ed.). *Handbook of emotion regulation* (pp. 3–20). New York, NY: Guilford Press.

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*, 1–26.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348.

Groth-Marnat, G. (2009). *Handbook of psychological assessment.* John Wiley & Sons.

Haviland, M. G., & Reise, S. P. (1996). A California Q-set alexithymia prototype and its relationship to ego-control and ego-resiliency. *Journal of Psychosomatic Research, 41*, 597–607.

Haviland, M. G., Warren, W. L., & Riggs, M. L. (2000). An observer scale to measure alexithymia. *Psychosomatics, 41*, 385–392.

Honkalampi, K., Hintikka, J., Laukkanen, E., & Viinamäki, J. L. H. (2001). Alexithymia and depression: A prospective study of patients with major depressive disorder. *Psychosomatics, 42*, 229–234.

John, O. P., & Eng, J. (2014). Three approaches to individual differences in affect regulation: Conceptualizations, measures, and findings. In J. J. Gross (Ed.). *Handbook of emotion regulation* (pp. 321–345). (2nd ed.). New York, NY: Guilford.

Kia-Keating, M., No, U., Moore, S., Furlong, M. J., Liu, S., & You, S. (2017). Structural validity of the Depression, Anxiety, and Stress Scales-21 adapted for US undergraduates. *Emerging Adulthood, 1*–7. http://dx.doi.org/10.1177/2167696817745407.

Kline, P. (1979). *Psychometrics and psychology.* London: Academic Press.

Kline, P. (2013). *Handbook of psychological testing.* Routledge.

Kooiman, C. G., Spinhoven, P., & Trijsburg, R. W. (2002). The assessment of alexithymia: A critical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20. *Journal of Psychosomatic Research, 53*, 1083–1090.

Lane, R. D., Lee, S., Reidel, R., Weldon, V., Kaszniak, A., & Schwartz, G. E. (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosomatic Medicine, 58*, 203–210.

Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The levels of emotional awareness scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment, 55*, 124–134.

Lane, R. D., & Schwartz, G. E. (1987). Levels of emotional awareness: A cognitive-developmental theory and its application to psychopathology. *The American Journal of Psychiatry, 144*, 133–143.

Lane, R. D., Weihs, K. L., Herring, A., Hishaw, A., & Smith, R. (2015). Affective agnosia: Expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. *Neuroscience & Biobehavioral Reviews, 55*, 594–611.

Leising, D., Grande, T., & Faber, R. (2009). The Toronto Alexithymia Scale (TAS-20): A measure of general psychological distress. *Journal of Research in Personality, 43*, 707–710.

Leweke, F., Bausch, S., Leichsenring, F., Walter, B., & Stingl, M. (2009). Alexithymia as a predictor of outcome of psychodynamically oriented inpatient treatment. *Psychotherapy Research, 19*, 323–331.

Leweke, F., Leichsenring, F., Kruse, J., & Hermes, S. (2012). Is alexithymia associated with specific mental disorders? *Psychopathology, 45*, 22–28.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment.* USA: Oxford University Press.

Loas, G., Otmani, O., Verrier, A., Fremaux, D., & Marchand, M. P. (1996). Factor analysis of the French version of the 20-ltem Toronto Alexithymia Scale (TAS-20). *Psychopathology, 29*, 139–144.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression Anxiety Stress Scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*, 335–343.

Luminet, O., Vermeulen, N., Demaret, C., Taylor, G. J., & Bagby, R. M. (2006). Alexithymia and levels of processing: Evidence for an overall deficit in remembering emotion words. *Journal of Research in Personality, 40*, 713–733.

Lundh, L. G., Johnsson, A., Sundqvist, K., & Olsson, H. (2002). Alexithymia, memory of emotion, emotional awareness, and perfectionism. *Emotion, 2*, 361–379.

Lyvers, M., McCann, K., Coundouris, S., Edwards, M. S., & Thorberg, F. A. (2018). Alexithymia in relation to alcohol use, emotion recognition, and empathy: The role of externally oriented thinking. *American Journal of Psychology, 131*, 41–51.

Marchesi, C., Ossola, P., Tonna, M., & De Panfilis, C. (2014). The TAS-20 more likely measures negative affects rather than alexithymia itself in patients with major depression, panic disorder, eating disorders and substance use disorders. *Comprehensive Psychiatry, 55*, 972–978.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320–341.

Mattila, A. K., Keefer, K. V., Taylor, G. J., Joukamaa, M., Jula, A., Parker, J. D., & Bagby, R. M. (2010). Taxometric analysis of alexithymia in a general population sample from Finland. *Personality and Individual Differences, 49*, 216–221.

McDougall, J. (1974). The psychosomata and the psychoanalytic process. *Internal Review of Psychoanalysis, 1*, 437–459.

Meganck, R., Vanheule, S., & Desmet, M. (2008). Factorial validity and measurement invariance of the 20-item Toronto Alexithymia Scale in clinical and nonclinical samples. *Assessment, 15*, 36–47.

Müller, J., Bühner, M., & Ellgring, H. (2004). The assessment of alexithymia: Psychometric properties and validity of the Bermond–Vorst alexithymia questionnaire. *Personality and Individual Differences, 37*, 373–391.

Nemiah, J. C. (1977). Alexithymia: Theoretical considerations. *Psychotherapy and Psychosomatics, 28*, 199–206.

Nemiah, J. C. (1984). The psychodynamic view of anxiety. In R. O. Pasnau (Ed.). *Diagnosis and treatment of anxiety disorders* (pp. 117–137). Washington, DC: American Psychiatric Press.

Nemiah, J. C., & Sifneos, P. E. (1970). Psychosomatic illness: A problem in communication. *Psychotherapy and Psychosomatics, 18*, 154–160.

Neumann, D., Malec, J. F., & Hammond, F. M. (2017). Reductions in alexithymia and emotion dysregulation after training emotional self-awareness following traumatic brain injury: A phase I trial. *The Journal of Head Trauma Rehabilitation.* http://dx.doi.org/10.1097/HTR.0000000000000277.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory.* New York: McGraw-Hill.

Ochsner, K. N., & Gross, J. J. (2014). The neural bases of emotion and emotion regulation: A valuation perspective. In J. J. Gross (Ed.). *Handbook of emotion regulation* (pp. 23–42). New York, NY: Guilford Press.

Panayiotou, G., Leonidou, C., Constantinou, E., Hart, J., Rinehart, K. L., Sy, J. T., & Björgvinsson, T. (2015). Do alexithymic individuals avoid their feelings? Experiential avoidance mediates the association between alexithymia, psychosomatic, and depressive symptoms in a community and a clinical sample. *Comprehensive Psychiatry, 56*, 206–216.

Parker, J. D., Keefer, K. V., Taylor, G. J., & Bagby, R. M. (2008). Latent structure of the alexithymia construct: A taxometric investigation. *Psychological Assessment, 20*, 385–396.

Porcelli, P., & Mihura, J. L. (2010). Assessment of alexithymia with the Rorschach comprehensive system: The Rorschach alexithymia scale (RAS). *Journal of Personality Assessment, 92*, 128–136.

Preece, D., Becerra, R., Allan, A., Robinson, K., & Dandy, J. (2017). Establishing the theoretical components of alexithymia via factor analysis: Introduction and validation of the attention-appraisal model of alexithymia. *Personality and Individual Differences, 119*, 341–352.

Preece, D., Becerra, R., Robinson, K., & Dandy, J. (2017). Assessing alexithymia: Psychometric properties and factorial invariance of the 20-item Toronto Alexithymia Scale (TAS-20) in nonclinical and psychiatric samples. *Journal of Psychopathology and Behavioral Assessment, 1*–12. http://dx.doi.org/10.1007/s10862-017-9634-6.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*, 1–15.

Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology, 30*, 59–64.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667–696.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*, 544–559.

Rieffe, C., Oosterveld, P., Miers, A. C., Terwogt, M. M., & Ly, V. (2008). Emotion awareness and internalising symptoms in children and adolescents: The emotion awareness questionnaire revised. *Personality and Individual Differences, 45*, 756–761.

Ripper, C. A., Boyes, M. E., Clarke, P. J., & Hasking, P. A. (2018). Emotional reactivity, intensity, and perseveration: Independent dimensions of trait affect and associations with depression, anxiety, and stress symptoms. *Personality and Individual Differences, 121*, 93–99.

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reverse of social anxiety is not always the opposite: The reverse-scored items of the social interaction anxiety scale do not belong. *Behavior Therapy, 38*, 192–206.

Rottenberg, J. E., & Johnson, S. L. (2007). *Emotion and psychopathology: Bridging affective and clinical science.* American Psychological Association.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. F. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait mood scale. In J. W. Pennebaker (Ed.). *Emotion, disclosure, and health* (pp. 125–154). Washington, DC: American Psychological Association.

Sekely, A., Taylor, G. J., & Bagby, R. M. (2018). Developing a short version of the Toronto structured interview for alexithymia using item response theory. *Psychiatry Research.*

http://dx.doi.org/10.1016/j.psychres.2018.03.002.

Sifneos, P. E. (1973). The prevalence of 'alexithymic' characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22*, 255–262.

Sifneos, P. E. (1996). Alexithymia: Past and present. *The American Journal of Psychiatry, 153*, 137–142.

Stevens, J. P. (1992). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Erlbaum.

Suslow, T., & Junghanns, K. (2002). Impairments of emotion situation priming in alexithymia. *Personality and Individual Differences, 32*, 541–550.

Taylor, G. J., Bagby, R. M., & Parker, J. D. (1999). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness.* Cambridge University Press.

Taylor, G. J., Bagby, R. M., & Parker, J. D. (2003). The 20-item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychosomatic Research, 55*, 277–283.

Taylor, G. J., Parker, J. D., Bagby, R. M., & Bourke, M. P. (1996). Relationships between alexithymia and psychological characteristics associated with eating disorders. *Journal of Psychosomatic Research, 41*, 561–568.

Taylor, G. J., Ryan, D., & Bagby, R. M. (1985). Toward the development of a new self-report alexithymia scale. *Psychotherapy and Psychosomatics, 44*, 191–199.

Thorberg, F. A., Young, R. M., Sullivan, K. A., & Lyvers, M. (2009). Alexithymia and alcohol use disorders: A critical review. *Addictive Behaviors, 34*, 237–245.

Thorberg, F. A., Young, R. M., Sullivan, K. A., Lyvers, M., Hurst, C., Connor, J. P., & Feeney, G. F. (2010). A confirmatory factor analysis of the Toronto Alexithymia Scale (TAS-20) in an alcohol-dependent sample. *Psychiatry Research, 178*, 565–567.

van der Velde, J., Servaas, M. N., Goerlich, K. S., Bruggeman, R., Horton, P., Costafreda, S. G., & Aleman, A. (2013). Neural correlates of alexithymia: A meta-analysis of emotion processing studies. *Neuroscience & Biobehavioral Reviews, 37*, 1774–1785.

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS One, 8*, e68967.

Vermeulen, N., Luminet, O., & Corneille, O. (2006). Alexithymia and the automatic processing of affective information: Evidence from the affective priming paradigm. *Cognition & Emotion, 20*, 64–91.

Vorst, H. C., & Bermond, B. (2001). Validity and reliability of the Bermond–Vorst alexithymia questionnaire. *Personality and Individual Differences, 30*(3), 413–434.

Watters, C. A., Taylor, G. J., & Bagby, R. M. (2016). Illuminating the theoretical components of alexithymia using bifactor modeling and network analysis. *Psychological Assessment, 28*, 627–638.

Way, I. F., Applegate, B., Cai, X., Franck, L. K., Black-Pond, C., Yelsma, P., ... Muliett, M. (2010). Children's alexithymia measure (CAM): A new instrument for screening difficulties with emotional expression. *Journal of Child & Adolescent Trauma, 3*, 303–318.

Weiss, N. H., Gratz, K. L., & Lavender, J. M. (2015). Factor structure and initial validation of a multidimensional measure of difficulties in the regulation of positive emotions: The DERS-positive. *Behavior Modification, 39*, 431–453.

Williams, C., & Wood, R. L. (2010). Alexithymia and emotional empathy following traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 32*, 259–267.

Zeitlin, S. B., & McNally, R. J. (1993). Alexithymia and anxiety sensitivity in panic disorder and obsessive-compulsive disorder. *The American Journal of Psychiatry, 150*, 658–660.

Zou, C., Plaks, J. E., & Peterson, J. B. (2017). Don't get too excited: Assessing individual differences in the down-regulation of positive emotions. *Journal of Personality Assessment,* 1–11. http://dx.doi.org/10.1080/00223891.2017.1339711.