An exploration of the association between overgeneralisation and depression in a novel and neutral valence rule-learning behavioural task

Abstract

Background: Beck defined seven cognitive distortions which sub serve the cognitive triad in depression. Inconsistent findings regarding their operation stem from the added complexity of valence and weaknesses in questionnaire measures. Research indicates the merits of further elucidating the specific and differing mechanisms of individual distortions to improve treatment. The relationship between one important distortion, overgeneralisation, and depressive symptoms is explored in a paradigm previously unexplored: a rule-learning behavioural task.

Methods: 272 participants completed measures of positive, negative generalisation, dysfunctional attitudes, anxiety and depression. The correlations between these and the behavioural task were investigated in an exploratory cross-sectional design.

Results: Hypothesis that generalisation per se would be positively associated with level of depressive symptoms was both partially supported and rejected. Behavioural task measures of generalisation did not correlate with either positive or negative overgeneralisation questionnaire measures.

Conclusions: Findings are inconclusive but promising that depressive symptoms are associated with the degree of generalisation. It is not clear if the no finding of an association between generalization measures relates to low level of depression in sample or carrying constructs of overgeneralisation. Several arising methodological issues are discussed to further improve this important direction of research.

Introduction

According to Beck’s (1967, 1976) cognitive model, the negative triad function as the central and common feature of depression. This triad of negative automatic thinking about the self, world and future are driven by underlying negative self-schemata or beliefs. These maladaptive schemata manifest as what Beck (1967) describes as seven cognitive distortions: arbitrary inference, absolutistic dichotomous thinking, selective abstraction minimization or magnification, personalization and overgeneralization. In this diathesis model, a stressful life event will activate this negatively biased information processing system and ultimately give rise to depression by distorting thought. Miranda, Persons and Byers (1990) found that amongst those remitted from a prior episode, dysfunctional beliefs fluctuated in tandem with the level of depressive symptoms, but that this pattern did not exist amongst those who had never been depressed. Carver (1998) found that overgeneralisation from negative events predicted later depressive symptoms, whilst accounting for intervening life events.

Blake, Dobson, Sheptycki and Drapeau (2016) examined the level of cognitive errors and depression in a sample receiving cognitive therapy, before cognitive restructuring had begun. They used the observer rating system Cognitive Errors Rating System (CERS; Drapeau, Perry and Dunkley, 2008) to measure 15 types of cognitive errors. Level of depressive symptoms was not correlated with positive, negative or neutral total level of errors. However, a different pattern emerged when separating participants according to overall level of errors. For those with a high level, there was a significant correlation with level of depressive symptoms and total cognitive errors, as well as specifically with negative selective abstraction. There was an inverse correlation between level of depression and positive overgeneralisation. For those with a low level of distortions, positive overgeneralisation and total errors was associated with lower levels of depressive symptoms. These findings demonstrate that the pattern of associations with depression is complex, and that investigating the cognitive errors individually and in a variety of contexts is further warranted.

Cognitive therapy operates on the premise that interrupting and changing these negative self-beliefs and thoughts will also reduce other symptoms of depression (Beck, 1976). With a relapse rate of up to 85% (Sim et al, 2016) and the lack of relative paucity of research into these specific distortions reflects a clear need to better understand them to ultimately improve treatment efficacy.

Fennell and Campbell (1984) found that overgeneralization was the cognitive distortion most ‘consistently’ associated with depression. Others have argued likewise (Carver and Ganellen, 1983, Ganellen 1988). Beck et al (1979) defined overgeneralization as *“the pattern of drawing a general rule or conclusion on the basis of one or more isolated incidents and applying the concept across the board to related and unrelated situations”.*

Overgeneralization has been studied most widely in the context of autobiographical memory. King et al (2010) review the evidence that overgeneralization in depression is typically associated with a greater recall of categorical or semantic relative to episodic or specific memory recall. A similar pattern has also been observed in future planning, with depression predicting less specificity in planning goal-attainment (Dickson and Moberly (2013). However, the picture is also mixed here with findings showing that depression is associated with greater over-general recall in response to negative cues, positive ones or that valence is unrelated to this process: suggesting further that there are a variety of factors moderating this relationship (King et al, 2010).

Epstein (1992) found that poor constructive thinkers made more negative generalisations to the self, than stronger constructive thinkers. There was no difference between the degree of constructive thinking and positive generalisations, but generalisation was greatest in response to positive outcomes. Epstein (1992) concluded that OG is a selective cognitive distortion, inherently motivational, as opposed to indiscriminately affecting information processing. Other findings dispute this, as Thew, Gregory and Roberts (2017) tentatively suggest that overgeneralization may only be the vehicle through which self-criticism leads to mood reduction. They found that both eating disorder and depressed patients, relative to non-patient controls, following a verbal ability task designed to elicit self-criticism reported a significantly lower mood. Similarly, Kernis, Brockner and Frankel (1989) found that self-esteem and overgeneralization are inversely correlated, and that overgeneralization mediates the extent of a reduction in self-esteem following negative feedback. Beck’s (1967, 1976) theory suggests that those with depression differ from non-depressed in operating a negative information processing bias, but the evidence suggests that the valence and the bias are not the same per se. Whilst OG is theorized to spread failure particular task to the greater self (Beck, 1976), the extent to which generalizing is harmful must surely depend on the effects. Indeed, van den Heuvel et al (2012) argue that the effects of positive overgeneralization act as an adaptive buffer against negative overgeneralization.

Van den Heuvel, Derksen, Eling and van den Staak (2012) found that there was a greater degree of both positive and negative overgeneralizations in a sample of participants with Major Depressive Disorder, relative to never depressed controls. Furthermore, those with MDD had a greater proportion of negative than positive generalization in relation to self-attributions but that for attributions across situations they showed a higher proportion of positive overgeneralizations. However, it is worth highlighting a further point which confounds results in the literature. There is a question about whether the tasks used in the aforementioned studies address the same construct. However, despite the seeming synonymous nature of attributions and overgeneralisation, they are distinct concepts. Measures such as the Attributional Style Questionnaire (ASQ; Abramson, Seligman and Teasdale, 1978) operate on the premise that overgeneralisation can be inferred from overly broad attributions of the cause in response to a situation, whilst overgeneralisation may more accurately be described as making a broad generalisation from the result of a situation (Kernis, Brockner and Frankel, 1989). Ganellen (1987) found the ASQ and Attitudes to Self ggeneralization subscale (ATS; Carver and Ganellen, 1983) are not highly correlated.

Another confounding factor to the effects of the OG is mood at the time of testing. Segal, Gemar and Williams (1999) argue that cognitive distortions are only activated during a negative mood. They found that when comparing patients treated with either CBT or pharmacotherapy there was no difference in their level of dysfunctional beliefs, unless sad mood is induced.

The affective inference hypothesis suggests a greater attentional bias for emotional material, meaning that performance will be selectively impaired or enhanced (Kircanski, Joormann and Gotlib, 2012). Whilst the extent to which depression is associated with impairments in working memory and executive function is still debated, relatively consistent findings show that depression is associated with a poorer ability to attend to the task at hand and inhibit extraneous processing (Kircanski, Joormann and Gotlib, 2012).

In terms of Construal Level Theory (CLT) of cognitive processing, depression is associated with a higher level of construal (Watkins, 2008). Higher levels of construal are more general and abstract, whilst lower levels are more contextual and detail specific. A function of higher levels of construal is that it operates with better generalisation across situations. Therefore, whilst higher level of construal is occurring as a function of depression, situations in which elicit a negative response are likely to be more overgeneralised without the benefit of a lower level of construal. Watkins, Baeyens and Read (2009) argue therefore that this is what causes the deleterious effects of rumination that is associated with depression. It is in this light that concreteness training has shown promise to decrease depressive symptoms associated with OG but would be further enhanced with more accurate knowledge of the parameters within which it operates (Mogoase, Brailean and David (2013).

There is also a further question of whether OG is in fact a one dimensional and single construct. MacLeod and Williams (1990) compared two contemporary measures of overgeneralization in depression: the ATS and the Cognitions Questionnaire (CQ: Fennell and Campbell, 1984). Whilst both measures correlate highly with Beck Depression Inventory scores, they do not correlate with each other. The ATS overgeneralization subscale asks questions about the extent to which the respondent generalizes from a particular incident of failure. The CQ measures both positive and negative generalization in response to imagined hypothetical scenarios. MacLeod and Williams (1990) also make the point that the lack of correlation between measures may reflect that ATS relies more on respondents’ awareness of their own thinking.

The other important factor to consider is the task itself. Questionnaire measures asking about specific (hypothetical) situations suffer a limitation in that humans ‘vary idiosyncratically’ such that situations which activate cognitive distortions will also differ (Rectar, Segal and Gemar (1998). Questionnaires of cognitive distortions/errors ask about how much a respondent agrees with statement or how likely to have a certain thought, but this not a direct measure of overgeneralisation. Even those which require participant to imagine hypothetical event and respond accordingly have limitations, as Klar, Gabai, and Baron (1996) argue that they require a level of self-awareness as a separate and possibly confounding factor.

The observer rating system Blake, Dobson, Sheptycki and Drapeau (2016) has its shortcomings, not least that of likely bias in content within a therapy session. There is also the reporting bias in depression to consider. Hunt, Ashara and Cashaw, (2010) report an effect of differentially reporting depressive symptoms depending on the covert or overt nature of the measure. The present study seeks to address some gaps in the literature by investigating the degree to which depression is associated with generalization, without the influence of valence, in a seemingly never before studied behavioral paradigm: rule-learning.

Perhaps the most relevant is Dobson and Dobson (1981), who depressed and control participants were asked to identify three rules of varying difficulty based on the pattern of shapes and colours on a playing card. Participants were asked to select additional cards to gather more information about the rule in question. They found that for those with depression, relative to controls, were able solve the rule as easily but that they were more likely to continue to selecting cards which confirmed the information they had already learnt about the rule, rather than continue to gather additional information. They did not measure cognitive distortion, but this plausibly may have been a function of overgeneralization.

Implications of better understanding the processes have widespread implications for diagnostic criteria, case formulations, therapeutic targets of change and outcome measures (Dozois, Covin and Brinker, 2003). MacLeod, Koster and Fox (2009) argue that overgeneralizing would make an ideal candidate target in Cognitive Bias Modification but that the process and situations in which it occurs need to be better understood. It is also of interest to delineate the highly co-morbid anxiety and depression. Dunn et al (2009) argue that enhancing understanding of the unique cognitive profiles of disorders is beneficial in aiding differential diagnosis. There is a need to understand the specificity of the two disorders. Tairi, Adams and Zilikis (2016) found cognitive errors were associated with both anxiety and depression, but overgeneralisation was only associated with anxiety and not depression. However, Carver (1998) however found that OG was uniquely related to depression but not anxiety.

The primary objective is to explore the level correlation between depression and generalization in a novel behavioral task. The specific hypothesis is that higher levels of depressive symptoms will be associated with a greater level of generalization of rule learning. The secondary objective is to further assess how measures of generalization on this task relate to questionnaire measures of both positive and negative generalization. The tentative second hypothesis is that negative ATS generalization will be negatively correlated with the behavioural measures of generalization, with the opposite true for the positive generalization scales.

Methods

Participants

Three hundred self-selecting participants were initially recruited and paid, although the final sample consisted of 272 participants. See Figure 1 for the breakdown of how many participants did not complete or were excluded at each stage.

Figure 1: Participants who did not complete or were excluded at each stage of game.

Demographic information can be found in Table 1. The age of participants ranged from 19 to 67, with a mean age of 35 (SD=10.6). One hundred and sixty (60%) were male. Participants were mostly from the US with a minority from India. Most reported English as their native tongue (90%) and 18 (7%) were either Hindi or Tamil speakers. Twenty-five (9%) had high school as their highest level of education, whilst 151 had a Bachelors (46%) or postgraduate degree (11%).

Sixty-nine participants reported having received a diagnosis of either depression (10%), depression and anxiety (10%) or anxiety alone (6.3%). Of these, 32 had a current diagnosis from a mental health professional, with 1.1% of participants having a first time diagnosis and 10.8% were recurring. Thirty-seven had a previous but no current diagnosis (13.8%), whilst the remaining 200 (74.3%) had never been diagnosed. Four (1.5%) reported being colour blindness.

Ethical approved was sought and given by the appropriate UCL committee. Participants were recruited through Amazon Mechanical Turk website, where users complete Human Interaction Tasks for payment. They were told that the task would take roughly one hour, but that some participants were completing much faster and some slower. Payment for successfully completing the task was five US dollars irrespective.

The sole inclusion/exclusion criteria were that this was their first attempt and that they were not colour blind. Participants were made aware of the conditions and after agreeing were directed to an external website to complete the task.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | **Mean (SD)** | **Median (IQR)** |
| **Age** | 35.19(10.63) | 32(12.75) |
| **Gender** | **n** | **%** |
| Male | 160 | 58.8 |
| Female | 112 | 41.2 |
| **Colour Blind** | **n** | **%** |
| Yes | 4 | 1.5 |
| No | 268 | 98.5 |
| **Language** | **n** | **%** |
| English | 243 | 89.4 |
| Tamil or Hindi | 18 | 6.6 |
| Other | 11 | 4 |
| **Education** | **n** | **%** |
| High school | 25 | 9.2 |
| Some College | 82 | 30.1 |
| Vocational/technical school(2 year) | 10 | 3.7 |
| Bachelors degree | 124 | 45.6 |
| Masters degree | 27 | 9.9 |
| Professional degree (MD, JD, etc. | 3 | 1.1 |
| Doctoral degree | 1 | 0.4 |
| **MH diagnosis\*** | **n** | **%** |
| None | 200 | 74.3 |
| Depression | 26 | 9.7 |
| Depression and anxiety | 26 | 9.7 |
| Anxiety disorder | 17 | 6.3 |
| **When diagnosed\*** | **n** | **%** |
| Never | 200 | 74.3 |
| Current only | 3 | 1.1 |
| Current AND past | 29 | 10.8 |
| Past only | 37 | 13.8 |

\*Data missing for mental health (MH) diagnosis (n=3)

Table 1 – Participant demographics and reported Mental Health diagnosis by qualified professional.

Measures

Attitudes to Self

The Attitudes to Self (ATS; Carver and Ganellen, 1983) is a ten question self-report questionnaire designed to measure self-punitiveness in depression, expanding this into three constructs. Three questions relate to the extent to which one holds oneself to overly high standards. Three questions concern the tendency towards self-criticism when failing to meet these high standards and four questions relate to the generalization of a single failure to a broader sense of self-worth. Answers are scored on a five point Likert scale ranging from ‘I agree a lot’ to ‘I disagree a lot’. Carver and Ganellen (1983) found confirmatory factor analysis to show that the three constructs are distinct and also uniform across genders.

Positive Generalisation

The Positive Generalisation (POG; Eisner, Johnson and Carver, 2008) is a 16-question self-report questionnaire. Answers are on a five point Likert scale of agreement, with possible scores ranging from 16 to 80.

It measures the tendency to generalize a positive experience to a wider sense of worth within three domains. Six questions concern the tendency to generalize success laterally from one area of life to another (‘If I succeed at something, it makes me feel I will succeed in other areas as well’). Five questions concern generalising upwards from a success in one area to a much larger and broader outcome (‘If I do well on my Psych course, it would make me think of being a famous Psychologist’). Five questions to address the tendency to generalise explicitly social outcomes (‘When I made my first friend here, I knew I’d be a big success socially’). The authors’ factor analysis confirmed that they are three distinct subscales. The POG was developed to mirror the ATS generalisation subscale, to explore to depression and mania in bipolar disorder. Some questions refer to specific events or situations and others are more broad. Eisner, Johnson and Carver (2008) found there was no association between either history of- or current depressive symptoms and any of the three POG subscales. All three subscales predicted risk for mania.

Dysfunctional Attitude Scale (form A)

The Dysfunctional Attitude Scale (DAS; Weissman, 1979) is two 40-item self-report questionnaires, forms A and B. Only the DAS-A is used here. Answers are on a seven point Likert scale of agreement, with possible scores ranging between 40 and 240. Based on Beck’s model of cognitive dysfunction, the DAS measures dysfunctional schemas or beliefs theorised to sub-serve negative automatic thoughts. There is also some debate about the factor structure of the DAS. For example, Power et al (2004) argue for a three factor solution whilst Moore et al (2014) found a one factor solution they called ‘perfectionist beliefs’, although more than half of the items on the DAS did not load onto any specific factor. They also argue that it has a heavy bias towards interpersonal over other domains. Nevertheless, they did find a significant correlation (r=0.35, p<0.001) with Beck Depression Inventory score, suggesting adequate construct validity. The sample used by Power et al (2004) also showed good internal internal consistency, with a Cronbach’s alpha = 0.898 for the DAS-A.

Zung Depression Scale

The Zung Depression Scale (Zung, 1965) is a widely used 20-item self-report measure of depressive symptoms, covering somatic, affective and psychological symptoms. Answers are rated on a four point Likert scale, with a possible score ranging between 20 and 80. Although a screening rather than diagnostic tool, a score between 50 to 59 is generally considered to indicate mild depression and 60-69 indicating moderate depression. A score above 70 indicates severe depression. Zung (1965) reported a mean score of 26 in a mixed US sample, whilst Knight, Waal-Manning and Spears (1983) found a mean(SD) of 32 (6.7) for men and 36 (7.4) for females in a random sample of 1173 non-clinical respondents. Knight, Waal-Manning and Spears (1983) report an alpha coefficient of 0.79. Lee et al (1994) found a high sensitivity of 92.3% and specificity of 87.5% in a sample of non-clinical geriatric adults. Campo-Arias et al (2006) found an internal consistency of Cronbach’s alpha of 0.803 in a similar yet younger sample, but found a much lower specificity of 70.3%. Faravelli, Albanesi and Poli (1986) found that the BDI and Zung had a strong correlation (r= 0.74) in sample of 100 patients with depression, although they suggest that the reason that the correlation is not higher reflects a slightly differing factor structure of the measures.

STAI (trait)

The State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch and Lushene, 1970) is a widely used two-part self-report measure of anxiety. There are 20 items on each of the scales, with the trait scale asking about how one *generally* feels, whilst the state scale refers to the *present moment*. Only the trait scale is used here. Answers are reported on a four point Likert scale, ranging from ‘almost never‘ to ‘almost always’, and a possible score range between 20 and 80. Knight, Waal-Manning and Spears (1983) found an internal consistency of Cronbach’s alpha= 0.87. They also found a high level of correlation between the STAI trait- form with Zung, reflecting the general covariance of anxiety and depressive symptoms (r=0.7, p<0.001).

Other questions relating to mental health

A questionnaire about whether they had received any diagnoses from a professional mental health worker, such as a psychiatrist. They were asked to elaborate onwhat and when this covers.

Cognitive Task

The game was developed by Professor Loewenstein at the Hebrew University of Jerusalem to measure generalisation in the context of rule learning, and was adapted for the present experiment.

Participants are presented with an image of a coloured ring. Within this outer ring are several smaller coloured circles evenly spaced along the inside edge of the ring. Participants have the choice of pressing two buttons in response to the image. Incorrect responses to trials are repeated to maximise the possibility of learning, and feedback (either smiley or sad face) is given for the previous trial. The objective is to work out the rule that dictates what button to press. The game has eight levels of increasing complexity, with new additional elements being introduced at each stage. An example screenshot of each stages 1 to 7 are in Figure 2, with an accompanying description of the changes from the previous stage. These new complexities require participants to reassess whether the rule they are using is in fact still appropriate. For example, on stage 3 the colours of the circles and outer ring change from only being red or white, to being a variety of different colours. If a participant has been basing their decisions on colours, they will have to revaluate the rule they are using. Participants are told that there is one correct rule common to all stages. The correct rule once is therefore one that can generalise to all stages of the game, in spite of the complexities added at each new level. Eight consecutive correct responses are required to reach the next level and being able to do so suggests that they have successfully generalised a rule.

The ‘correct’ rule is based on which group of circles match the colour of the outer ring. However, the way in which this is expressed may vary according to how a participant represents this. For example, one variant of this rule will be to say that “when the largest number of circles matches the outer ring, press the left button”. Alternatively, one might express this as “when the overall majority of circles matches the outer ring, press the left button”. Another possible expression of the rule is “when the smallest number of circles matches the outer right, press the right button”.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Stage 1:** |  | -5 circles  -2 colours  -3 of one colour at top and 2 of other colour at bottom | **Stage 2:** |  | Change from previous stage:  -matching circles can be in any spatial orientation |
| **Stage 3:** |  | Change from previous stage:  -New colours introduced | **Stage 4:** |  | Change from previous stage:  -Number of possible circles increases |
| **Stage 5:** |  | Change from previous stage:  -Smallest/ largest number of circles is no longer always 2 | **Stage 6:** |  | Change from previous stage:  -3 simultaneous colours |
| **Stage 7:** |  | Change from previous stage:  -Feedback in odd trials only |  |  |  |

Figure 2: Example screenshot from stages 1 to 7, with accompanying description of change from previous stage

On stage eight, there are three groups of different coloured circles within the outer ring. There are always 13 circles in total and is unique in having a fixed number of thirty trials. There are three types of “ambiguous” trials and three types of “unambiguous trials”, yielding a total of five trials on each of the six different types. Feedback on level eight is only provided for the unambiguous trials. The different trial types can be seen in Table 2 below, with corresponding example image patterns in Figure 3.

Unambiguous trials have the same correct/incorrect answers regardless of the rule used (see above). However, ambiguous trials are those where a participant’s response will depend on the rule they have employed and it is thus possible to retrospectively deduce this rule for each participant according to their choices. This is done based on the pattern of responses on type 2 and type 5 ambiguous trials (see Table 2). Three out of five trials with a particular pattern of responding on these trials represents one of the aforementioned possible rule variants. The type 4 trials could be considered ambiguous (as the largest colour was not the overall majority, meaning that the “majority” rule would lead to a different answer than the “smallest” and “largest” rules) but in practice only one participant approached the game in this way. These trials were analysed separately. Finally, responses to the ambiguous trials serve as a measure of how participants respond in a situation where it is not necessarily clear-cut how their given rule may apply. Therefore, the error rates and reaction times on the differing trial types, and how they relate to each other, are the variables from which it is intended to infer information regarding a participant’s level of generalisation from one trial (type) to another.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | Correct button using: | |
|  | Number of circles of each colour | | | Trial type | "smallest" rule | "largest" rule |
| Matching outer ring (**M**) | **7 M** | 4 | 2 | 1)Unambiguous | Left | Left |
| Matching outer ring (**M**) | 7 | **4 M** | 2 | 2)Ambiguous | Left | Right |
| Matching outer ring (**M**) | 7 | 4 | **2 M** | 3)Unambiguous | Right | Right |
| Matching outer ring (**M**) | **6 M** | 4 | 3 | 4) Ambiguous | Left | Left |
| Matching outer ring (**M**) | 6 | **4 M** | 3 | 5)Ambiguous | Left | Right |
| Matching outer ring (**M**) | 6 | 4 | **3 M** | 6)Unambiguous | Right | Right |

Table 2: Stage 8 trial types and correct button for each. There were five trials of each type. *Note: correct button only varies according rule on trial types 2 and 5. See Table XX for corresponding example pattern of each trial type.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Trial type** | **Example of pattern displayed** |  | **Trial type** | **Example of pattern displayed** |
| 1.) Unambiguous |  |  | 4.) Ambiguous |  |
| 2.) Ambiguous |  |  | 5.) Ambiguous |  |
| 3.) Unambiguous |  |  | 6.) Unambiguous |  |

Figure 3: Example image pattern for the six stage 8 trial types.

Procedure

The study design was cross-sectional. Participants completed a standard consent form, guaranteeing anonymity, confidentiality, data protection and the right to withdraw without prejudice. They then completed the questionnaires. After this, participants filled out a brief feedback form and were provided with a completion code. This code was to be input on the Amazon website for their work to be reviewed and approved for payment.

Analysis plan

Several outcome variables were created to reflect generalisation within the game. The first is the highest level reached with three or more errors. This ought to reflect the level at which a participant learnt the correct underlying rule, as in any subsequent stages they will make fewer or no errors. A lower number indicates that they learnt the rule sooner and therefore a higher level of generalisation.

The next outcome variable is the median correct reaction time on stage seven divided by the median reaction time on stage one, which reflects speeding over the game – a lower ratio indicates greater speeding and therefore greater generalisation.

The percentage of errors made on stage 8 was analysed for all trial types, as well as the median correct reaction times, for the ambiguous (types 2 and 5) and unambiguous (types 1,3 and 6) trial types. As faster RTs and fewer errors should also suggest higher generalisation, this was done in order to check if there were any associations between the questionnaire measures and any of the different trial types.

Ambiguous minus unambiguous median correct reaction time was calculated. Unambiguous trials are supposed to be more straightforward and so comparing the reaction time on ambiguous trials relative to unambiguous ones suggests how well a participant is able to generalise the rule they have learnt to the new ambiguous trial. A lower score (smaller difference) should therefore indicate greater generalisation. Finally, the percentage of errors on stage 8 ambiguous trials minus unambiguous trials was calculated. Again a lower score (smaller difference) indicating greater generalisation.

Bivariate Pearson’s correlations were performed between each of these variables and the questionnaire measures. The correlation with Zung is the primary of interest and how this relationship differs from the correlation with the positive and negative generalisation questionnaire measures.

Results

Descriptive statistics

Participant final sample

Seven participants who exceed 1,000 trials were excluded (figure 1). This was substantially more than the remainder of participants who completed the game, whose mean (SD) number of trials was 172 (142). This could be a function of the use of “bots” (Stokel-Walker, 2018). The final analysis was based on the remaining sample of 272 participants.

Participants not in final sample

Using an independent samples t-test, the 170 participants who completed the Zung questionnaire but not the entire task had a higher mean Zung score than those who did complete the game, with a marginally significant difference of 2.3 points (SD: 11.40, t=2.10, df = 351, p= 0.038). However, when comparing the ZUNG score of those who went on to complete stage 1 (n=90), but not the entire task, with those who completed the entire task, this difference disappeared (p=0.792)

Questionnaire scores

Summary scores on the questionnaires are presented in table 3. The sample had a relatively low level of depressive symptoms with a mean (SD) Zung of 36 (11). A score between 50 and 59 indicates mild depression, with 35 participants falling into this category. Five participants scored above 60 and only 2 of these scored above 70, indicating severe depression.

Twenty-six reported having a current diagnosis of (see Table 4). Two of these were for first time diagnoses and 12 were for recurring co-morbid depression and anxiety disorders.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Mean (*SD*)** | **Median (*IQR*)** | **Range** | **Minimum /maximum possible** |
| **ZUNG** | 36.31 (*11.09*) | 34 (*18*) | 20 - 71 | 20 - 80 |
| **STAI** | 39.76 (*13.27*) | 38 (*20*) | 20 - 80 | 20 - 80 |
| **DAS\*** | 124.91 (*39.23*) | 121 (*59*) | 40 - 240 | 40 - 240 |
| **ATS self-criticism** | 9.53 (*3.51*) | 10 (*5*) | 3 – 15 | 3 – 15 |
| **ATS high standards** | 10.68 (*3.16*) | 11 (*4*) | 3 – 15 | 3 – 15 |
| **ATS generalization** | 11.20 (*4.68*) | 11 (*8)* | 4 - 20 | 4 - 20 |
| **POG lateral** | 14.89 (*5.63*) | 14 (*7*) | 6 - 30 | 6 - 30 |
| **POG upward** | 18.15 (*5.87*) | 20 (*11*) | 5 - 25 | 5 -25 |
| **POG social** | 19.04 (*5.67*) | 21 (*10)* | 5 - 25 | 5 - 25 |

\* Data missing (n=1)

Table 3 – Participant mean, median, interquartile range scores for outcome measures

|  |  |
| --- | --- |
|  | **n *(%)*** |
| **No diagnosis** | 204 *(75)* |
|  |  |
| **Depression** |  |
| current only | 2 *(0.7)* |
| past only | 16 *(5.9)* |
| current and past | 12 *(4.4)* |
|  |  |
| **Anxiety** |  |
| current only | 2 *(0.7)* |
| past only | 8 *(2.9)* |
| current and past | 8 *(2.9)* |
|  |  |
| **Depression and Anxiety** |  |
| current only | 0 |
| past only | 8 *(2.9)* |
| both | 12 *(4.4)* |

Data missing (n=3)

Table 4 – Participant reported depression and/or anxiety diagnoses

Game results

Rule category

The majority of participants were found to use either the ‘largest’ rule (19%), where the ‘largest’ number of matching circles would determine the button choice, or the ‘smallest’ rule (76%). Only 12 participants did not fall into either of these rule categories – they were excluded from the analysis of stage 8 data.

Evidence of learning/rule attainment

The level at which the rule was learnt was positively correlated with both the number of errors on each of the first 7 stages individually (0.46>r>0.12, 0.001<p<0.048) and the total number of errors (r=0.49, p<0.001). There was also a correlation between the stage of rule attainment and the number of trials per stage (r=0.493, p<0.001). This was also the case with all stages individually (0.51>r>0.27, all p<0.001), with the exception of stage one (r=0.11, p=0.072). This indicates that the behavioural outcome is measuring the level of generalisation. Intriguingly, there was a small negative correlation between rule attainment and stage 1 correct RT (r=-0.169, p=0.005), but no correlation with RT on any other stage (all p>0.207).

For stage 8, there was a significant difference between ambiguous and unambiguous trial types, both in terms of reaction times and error rates. A paired sample t-test showed that the mean median correct RT for ambiguous trials was 0.33s slower than for unambiguous trials (SD=1.30, t=4.14, df=258, p<0.001). The mean type 4 ambiguous median RT was also slower than unambiguous trials by 0.32s (SD=2.03, t=2.57, df=258, p<0.001). Mean percentage of errors was 2.74% higher for ambiguous versus unambiguous trials (SD=13.36, t=3.31, df=258, p=0.001). The mean percentage of errors for type 4 trials was higher than unambiguous ones by 4.14% (SD=13.33, t=5.00, df=258, p<0.001).

There was no significant difference in reaction times between those using the ‘smallest’ and ‘largest’ rules, on any of the different stage eight trial types. There was also no significant difference between these rule groups in percent errors on either ambiguous or type 4 ambiguous trials (p>0.584). However, those using the ‘largest’ rule made significantly fewer errors than those using the ‘smallest’ rule on unambiguous, with a mean difference of 6.56% (SD: 1.88, t=3.50, df= 257 p<0.001).

Associations with depressive symptoms and cognition

Associations within depressive symptom and cognition questionnaires

As expected, there was a high level of correlation between the various questionnaires, see Table 5. Zung and STAI were highly positively correlated (r=0.88, p<0.001). Zung also correlated strongly with DAS (r=0.53, p<0.001), ATS self-criticism (r=0.53, p<0.001) and ATS generalisation (r=0.66, p<0.001). Zung correlated very weakly and negatively with both POG upward (r=-0.11, p=0.063) and POG social (r=-0.12, p=0.047). Surprisingly, there was also a weak positive correlation with POG lateral (r=0.24, p<0.001).

Associations between depressive symptoms, cognition questionnaires and demographics

The demographic variables were largely categorical, so simple linear regressions were instead used, in turn, with Zung as the outcome. Age was strongly and negatively associated with ZUNG. With every one year of age increase, the predicted Zung score decreased by 0.26 (95%CIs: -0.34 to -0.14, p<0.001). There was no significant difference between genders (p=0.265) or language (English vs Tamil/Hindi, p=0.12). Neither did any individual level of education predict Zung score, when compared with high school education as base reference category (global p=0.812).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | ZUNG | POG lateral | POG upward | POG social | STAI | DAS | ATS high-standards | ATS self-criticism | ATS generalization |
| ZUNG | Pearson Correlation | 1 | 0.24 | -0.11 | -0.12 | 0.88 | 0.523 | -0.05 | 0.53 | 0.66 |
|  | Sig. (2-tailed) |  | <0.001 | 0.063 | 0.047 | <0.001 | <0.001 | 0.392 | <0.001 | <0.001 |
| POG lateral | Pearson Correlation | 0.24 | 1 | 0.48 | 0.46 | 0.31 | -0.003 | -0.37 | 0.04 | 0.14 |
|  | Sig. (2-tailed) | <0.001 |  | <0.001 | <0.001 | <0.001 | 0.966 | 0<0.001 | 0.48 | 0.019 |
| POG upward | Pearson Correlation | -0.11 | 0.48 | 1 | 0.89 | -0.003 | -0.50 | -0.26 | -0.09 | -0.16 |
|  | Sig. (2-tailed) | 0.063 | <0.001 |  | <0.001 | 0.96 | <0.001 | <0.001 | 0.135 | 0.009 |
| POG social | Pearson Correlation | -0.12 | 0.46 | 0.89 | 1 | -0.01 | -0.47 | -0.22 | -0.09 | -0.13 |
|  | Sig. (2-tailed) | 0.047 | <0.001 | <0.001 |  | 0.837 | 0<0.001 | 0<0.001 | 0.125 | 0.027 |
| STAI | Pearson Correlation | 0.88 | 0.31 | -0.003 | -0.01 | 1 | 0.54 | -0.08 | 0.58 | 0.761 |
|  | Sig. (2-tailed) | 0<0.001 | <0.001 | 0.96 | 0.837 |  | 0<0.001 | 0.189 | <0.001 | <0.001 |
| DAS | Pearson Correlation | 0.53 | -0.003 | -0.50 | -0.47 | 0.54 | 1 | 0.14 | 0.50 | 0.6 |
|  | Sig. (2-tailed) | 0<0.001 | 0.966 | <0.001 | 0<0.001 | <0.001 |  | 0.023 | <0.001 | <0.001 |
| ATS high-standards | Pearson Correlation | -0.05 | -0.37 | -0.26 | -0.22 | -0.08 | 0.14 | 1 | 0.32 | 0.06 |
|  | Sig. (2-tailed) | 0.392 | <0.001 | <0.001 | 0<0.001 | 0.189 | 0.023 |  | <0.001 | 0.301 |
| ATS self-criticism | Pearson Correlation | 0.53 | 0.04 | -0.09 | -0.09 | 0.58 | 0.50 | 0.32 | 1 | 0.69 |
|  | Sig. (2-tailed) | <0.001 | 0.48 | 0.135 | 0.125 | 0<0.001 | <0.001 | <0.001 |  | <0.001 |
| ATS generalization | Pearson Correlation | 0.66 | 0.14 | -0.16 | -0.13 | 0.76 | 0.61 | 0.06 | 0.69 | 1 |
|  | Sig. (2-tailed) | <0.001 | 0.019 | 0.009 | 0.027 | 0<0.001 | <0.001 | 0.301 | <0.001 |  |

Data missing for DAS (n=1)

Table 5 – Pearson correlations for questionnaire measures vs themselves

Correlations between depression symptom and cognition questionnaires and behavioural outcome variables

Results for the following behavioural outcomes can be found in Table 6.

Highest level 3 or more incorrect

There was a very weak correlation between the highest level reached with three or more incorrect responses and ATS generalisation, which narrowly missed significance (r=0.11, p=0.068). This is in the opposite direction to the hypothesis. There was no correlation between this measure of rule attainment and any of the other questionnaire measures (p>0.097 for all).

Stage 7 correct RT divided by stage 1 correct RT

For the reaction time of correct trials on stage 7 as a proportion of stage 1 reaction times, there was a small negative correlation with Zung (r=-0.14, p=0.02), STAI (r=-0.13, p=0.03 and ATS self-criticism (r=-0.13, p=0.038) scores only. This is consistent with the hypothesis that overgeneralisation should lead to greater speed as the game progresses.

Stage 8 percent errors, all trial types

There was no evidence of a correlation between the percentage of errors on stage 8 trials, of any type, and any of the questionnaire measures.

Stage 8 median correct reaction time, all trial types

Stage 8 median correct reaction time on unambiguous trials was weakly and positively correlated with Zung score (r=0.17, p=0.006). Zung was not correlated with any of the other measures. Both POG lateral and social negatively correlated with RT on all three trial types, whilst DAS had a positive correlation with RT on all three trial types (see Table 6).

Median correct reaction times on stage 8 for ambiguous minus unambiguous trials

For correct reaction times on stage eight for ambiguous minus unambiguous trials, there was no evidence of a correlation with any of the questionnaires (p>.136 for all).

Percentage of errors on stage 8 ambiguous minus unambiguous trials

For the percentage of errors on stage eight ambiguous minus unambiguous trials, there was no evidence of a correlation with any of the questionnaires (p>0.244 for all).

Regressions

No further analysis using regressions was justified. None of the demographic variables were significantly correlated with both Zung and any of the main outcome variables. In an effort to control for those rushing and not reading the questionnaires, the time spent completing the questionnaires was analysed, but this did not correlate with any of the questionnaire measures or main outcome variables, so it was not included in any regressions. Several questionnaires were highly correlated with each other, such that the high proportion of shared variance would render a multiple linear regression meaningless. To avoid this issue of multicollinearity, an exploratory factor analysis could be appropriate. However, this is beyond the exploratory scope of this study and will be carried out in subsequent research.

Effects of diagnosis

Using the same main variables as above for the outcome, a simple linear regression was performed with the categorical variable of reported diagnosis as the predictor. No diagnosis was used as the base reference category. None of of the other combinations of diagnoses, listed in Table 4, showed evidence of an association with the main outcome variables (all p>0.1).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **ZUNG** | **POG lateral** | **POG upward** | **POG social** | **STAI** | **DAS\*\*** | **ATS high-standards** | **ATS self-criticism** | **ATS generalisation** |
| Correct RT stage 7 as a proportion of stage one | Pearson Correlation, r= | -0.14 | -0.04 | 0.05 | 0.04 | -0.13 | -0.08 | 0.08 | -0.13 | -0.08 |
|  | (2-tailed) p= | 0.02 | 0.554 | 0.465 | 0.538 | 0.03 | 0.198 | 0.177 | 0.038 | 0.187 |
| Stage 8 percent errors: Ambiguous minus unambiguous trials\* | Pearson Correlation | 0.01 | 0.03 | -0.006 | 0.01 | 0.01 | 0.03 | 0.01 | 0.07 | 0.004 |
|  | (2-tailed) p= | 0.859 | 0.694 | 0.919 | 0.873 | 0.87 | 0.668 | 0.884 | 0.244 | 0.953 |
| Stage 8 median correct RT: Ambiguous minus unambiguous trials\* | Pearson Correlation | 0.002 | 0.01 | -0.07 | -0.06 | -0.02 | 0.09 | -0.05 | -0.04 | -0.03 |
|  | (2-tailed) p= | 0.97 | 0.831 | 0.27 | 0.371 | 0.813 | 0.136 | 0.402 | 0.555 | 0.687 |
| Highest level with 3 or more incorrect | Pearson Correlation | -0.04 | -0.02 | -0.05 | -0.03 | -0.008 | 0.07 | 0.10 | 0.10 | 0.11 |
|  | (2-tailed) p= | 0.556 | 0.709 | 0.456 | 0.599 | 0.895 | 0.232 | 0.097 | 0.101 | 0.068 |
| Stage 8 percent errors: Unambiguous trials\* | Pearson Correlation | 0.01 | 0.02 | -0.04 | -0.05 | 0.06 | 0.06 | 0.01 | 0.09 | 0.03 |
|  | (2-tailed) p= | 0.847 | 0.814 | 0.577 | 0.431 | 0.371 | 0.374 | 0.82 | 0.169 | 0.592 |
| Stage 8 percent errors: Type 4 ambiguous trials\* | Pearson Correlation | -0.03 | 0.09 | 0.02 | -0.03 | -0.05 | -0.001 | 0.03 | 0.04 | -0.01 |
|  | (2-tailed) p= | 0.666 | 0.15 | 0.757 | 0.664 | 0.433 | 0.993 | 0.587 | 0.489 | 0.829 |
| Stage 8 percent errors: Ambiguous trials\* | Pearson Correlation | 0.03 | 0.04 | -0.01 | -0.04 | 0.02 | 0.03 | -0.007 | 0.04 | 0.04 |
|  | (2-tailed) p= | 0.586 | 0.562 | 0.822 | 0.542 | 0.79 | 0.693 | 0.907 | 0.477 | 0.535 |
| Stage 8 median correct RT: Unambiguous trials\* | Pearson Correlation | 0.17 | 0.01 | -0.19 | -0.23 | 0.12 | 0.24 | 0.01 | 0.05 | 0.08 |
|  | (2-tailed) p= | 0.006 | 0.87 | 0.002 | 0 | 0.05 | 0 | 0.834 | 0.392 | 0.192 |
| Stage 8 median correct RT: Type 4 ambiguous trials\* | Pearson Correlation | 0.11 | 0.02 | -0.13 | -0.14 | 0.03 | 0.15 | -0.02 | 0.006 | -0.001 |
|  | (2-tailed) p= | 0.081 | 0.698 | 0.045 | 0.023 | 0.642 | 0.016 | 0.742 | 0.922 | 0.989 |
| Stage 8 median correct RT: Ambiguous trials\* | Pearson Correlation | 0.09 | 0 | -0.16 | -0.17 | 0.03 | 0.18 | -0.02 | 0.01 | 0.03 |
|  | (2-tailed) p= | 0.159 | 0.998 | 0.011 | 0.005 | 0.615 | 0.003 | 0.733 | 0.841 | 0.684 |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

Note – \*Missing data (n=33), \*\*missing data (n=1)

Table 6 – Pearson correlations between main outcome variables and questionnaire measures

Discussion

In terms of the main hypothesis that depressive symptoms would be associated with greater generalisation, the results were mixed. Contrary to the hypothesis, there was no evidence that depressive symptoms were correlated with the stage at which the rule was learnt. If generalisation were to occur to a greater extent alongside depressive symptoms, a positive correlation would have been expected. That there was no association between Zung score and the relative reaction speed across ambiguous to unambiguous trials, also refuting the hypothesis and is surprising. This is because if greater generalisation were associated with depression as hypothesised, a negative correlation indicating greater generalisation to the new ambiguous trial would be expected. Likewise, that the relative percentage of errors between ambiguous and unambiguous trial types was not correlated with Zung scores equally fails to support the hypothesis.

However, there was some support in favour of the first hypothesis. The speed of learning across the first seven stages was correlated with ZUNG in the expected direction. Higher levels of depressive symptoms were associated with a faster rate of generalising as the game progressed.

That there was no correlation between the percentage of errors on any stage 8 trial types and Zung score is also interesting. The fact that stage 8 errors was unrelated to any of other questionnaire measures supports that the behavioural measure is likely not at issue here.

This indicates that there may be no association between how well the rule is retained once learnt and the level of depression. It may also indicate something about the appropriate level of difficulty of the task, in so much as depression is associated with working memory. Whilst this may be a function of the relatively low level of depression in the sample, the deficits in WM typical in depression would plausibly have lead to a significant correlation in this regard (Kircanski, Joormann and Gotlib, 2012).

That Zung score was associated with a greater reaction time on unambiguous trials is surprising as these trials should not be the ones to give pause for thought. It would be interesting to further explore how this may relate to working memory, but if this were a significant factor, it is not clear why there was no evidence of an association with ambiguous trials as well.

Ambiguous trials showed a slower median reaction time compared to unambiguous trials, as would be expected in new situation. The same pattern was found for the percentage of errors between trial types.

The evidence did not support the second hypothesis. The stage at which the rule was learnt was not associated with either negative generalisation nor any of the three constructs of positive generalisation. Whether the reason for this relates to issues with measures or differing underlying constructs of OG is not clear (MacLeod and Williams (1990).

Lateral and social positive overgeneralisation subscales were negatively correlated with reaction times on all stage 8 trial types, whilst dysfunctional attitudes score was positively correlated. There is no clear explanation for these findings.

The relationship between rule category and trial type war largely not noteworthy besides to say that the reason for a significantly great level of errors on unambiguous trials for smallest rule has no clear or expected reason. The lack of an observed association between any questionnaire measure and the reaction time on ambiguous relative to unambiguous trials has two plausible explanations. The first is that this is not a fit for purpose outcome measure. The second is that overgeneralisation is indeed a multi factor construct and the behavioural measure here is distinct from the ones measured by the questionnaires. The same can be said for of the proportion of errors across trial types.

The correlations between questionnaire were largely as expected and in support of the literature. The correlation found between Zung and the lateral positive overgeneralisation subscale conflicts with Eisner, Johnson and Carver (2008). It is also worth highlighting that the correlations between STAI and the behavioural outcomes were the same as those observed between the latter and Zung. In terms of delineating between depression and anxiety disorders, this measure is not able to elucidate further.

Another surprising observation was the lack of evidence for an association between gender and depressive symptoms. Whilst age was significantly correlated as the previous findings indicate, gender is also widely recognised as a strong predictor (eg Piccinelli and Wilkinson (2000).

It would be premature to speculate about the impact these findings may have on treatment but they certainly do show that this is an area in which future research has the potential to enhance understanding of this cognitive distortions and the ramifications of this for treatment and formulation may be to widen the parameters in which they currently sit. This study had a number of strengths. As an initial exploration of generalisation in a neutral context and a novel paradigm, this study has laid an important foundation in advancing towards some previously unanswered questions. The findings may not show major support for the hypothesis but it has served to highlight both that there may be an association worth following up and also some ways in which to enhance this work in future. Whilst the cross-sectional design does not allow speculation for causation, it did facilitate a much larger sample size than may otherwise have been feasible. This was major strength, allowing a high level of power to detect statistical differences.

There were several limitations. Notably, the level of depression of the sample was low and thus reducing sensitivity. One unanswered question is whether the degree of generalisation in this paradigm is related to the degree of symptom severity, so it would clearly be advantageous in future iterations of this study to use a sample of clinically depressed individuals. In would be beneficial to maximise the validity of the sample in order better support the objective of discovering how this generalisation in this context may be best used to enhance treatment targets, planning and efficacy. Multiple comparisons were not also corrected for and the study would have benefited from a factor analysis to distinguish between the overlap questionnaire measures. However, the background research in this paper was not felt suited to justification of what such factors may represent.

There was a high proportion of participants who began the task but did not finish. There are several possible explanations but there is likely some bias as a result. It is likely that for some of those dropping out it was purely a function of the time left available to them when they began the task, after the use of automated software to accept the task begins the timer. However, those who did complete the Zung questionnaire did have a small but statistically significant greater Zung score. Whilst the overall level of depressive symptoms was so low to begin with, it is seemingly unlikely that this slightly greater Zung score could be directly causal, but there is again some bias that would be best mitigated against in future studies. That this difference in Zung scores disappeared when comparing those who completed the first stage with the final sample is also interesting. The question of working memory would again be of interest to control for in future, particularly if a sample with greater depressive symptoms is used. The effects of this may be twofold. Firstly, to reduce the bias which may serve as a reason for a greater score of depressive symptoms in those who were unable to to complete the the start of the game (Kircanski, Joormann and Gotlib, 2012). Secondly, the associated difficultly with executive function would plausibly counteract in the opposite direction the expected pattern of greater generalisation.

The complete mismatch between self-reported diagnoses and behavioural outcomes, and Zung scores and these outcomes is also striking. It is likely there is a degree of bias introduced in the form of the depression reporting bias (Hunt, Ashara and Cashaw, 2010). The question asking about diagnoses may also have been too time period vague or in definition of what qualifies as a diagnosis. If participants report a current diagnosis they may be in remission. The cognitive vulnerability to depression that remains in remittance could account for the significant correlation between Zung and diagnoses but is then unclear why this ‘depressive scar’ (Lewinsohn, Steinmetz, Larson and Frankel, 1981) does not account for an association with the behavioral outcome measures.

In light of this, it would also be prudent to include a measure of current mood. This is also sensible because of the idiosyncrasy of participants (Rectar, Segal and Gemar, 1998), where the task itself may induce a negative mood in some participants, especially if they are completing tasks on this platform as a source of income as many do. This in turn is important in that the diathesis model suggests such a stressor may selectively have activated overgeneralization in some but not others (Beck, 1976). This would be of further interest to possibly have a between-subjects design comparing those with induced negative mood versus those not (Segal, Gemar and Williams (1999).

In summary, this study was the first to employ a behavioural rule learning task to measure generalisation in relation to depressive symptoms, free from task valence. The findings suggest that in this neutral context, there may be an association between depressive symptoms and the extent of generalisation. Results were likely hampered by the low level of depression in the sample but the study has fulfilled its mission to shed the first light on the path of a necessary direction of research.

**References**

**Abramson, L.Y., Seligman, M.E.P., & Teasedale, J.D. (1978**). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87, 49–74.*

**Beck, A. T.** (1967). Depression: Clinical, experimental, and theoretical aspects. *New York: Harper & Row*.

**Beck, A.T. (1976).** Cognitive therapy and the emotional disorders. *New York: International Universities Press.*

**Blake, E., Dobson, K.S., Sheptycki, A.R. & Drapeau, M. (2016).** The Relationship between Depression Severity and Cognitive Errors. *American Journal of Psychotherapy, 70(2), 203-221.*

**Campo-Arias, A., Diaz-Martinez, L., Rueda-Jaimes, G., Jaimes, R., Cadena, L. & Leonor Hernández, N. (2006).**Validation of Zung Self-Rating Depression Scale among the Colombian general population. *Social Behavior and Personality: an international journal, 34, 87-94.*

**Carver, C.S. (1998).** Generalization, Adverse Events, and Development of Depressive Symptoms. *Journal of Personality, 66(4), 607-619.*

**Carver, C.S. & Johnson, S.L. (2009).** Tendencies Toward Mania and Tendencies Toward Depression Have Distinct Motivational, Affective, and Cognitive Correlates. *Cognitive Therapy Research, 33(6), 552-569.*

**Carver, C.S. & Ganellen, R.J. (1983).** Depression and Components of Self-Punitiveness: High Standards, Self-Criticism and Overgeneralization*. Journal of Abnormal Psychology, 92(3), 330-337.*

**Dickson, J.M. & Moberly, N.J. (2013).** Reduced Specificity of Personal Goals and Explanations for Goal Attainment in Major Depression. *Plos One.*

**Dobson, D.J.G. & Dobson, K.S. (1981).** Problem Solving Strategies in Depressed and Nondepressed College Students. *Cognitive Therapy and Research, 5(3), 237-249.*

**Dozois, D. J. A., Covin, R., & Brinker, J. K. (2003**). Normative data on cognitive measures of depression. *Journal of Consulting and Clinical Psychology,* 71(1), 71-80.

**Drapeau, M., Perry, J.C. & Dunkley, D. (2008).** The Cognitive Error Rating System. *Montreal: McGill University.:*  Cited in **Blake, E., Dobson, K.S., Sheptycki, A.R. & Drapeau, M. (2016).** The Relationship between Depression Severity and Cognitive Errors. *American Journal of Psychotherapy, 70(2), 203-221.*

**Dunn, B.D., Stefanovitch, I., Buchan, K., Lawrence, A.D. & Dalgelish, T. (2009**). A reduction in positive self-judgment bias is uniquely related to the anhedonic symptoms of depression. *Behaviour Research and Therapy, 47(5), 374-381.*

**Eisner, L.R. Johnson, S.L. & Carver, C.S. (2008)**. Cognitive Responses to Failure and Success Relate Uniquely to Bipolar Depression Versus Mania. *Journal of Abnormal Psychology, 117(*1), 154-163.

**Epstein, S. (1992).** Coping Ability, Negative Self-Evaluation and Overgeneralization: Experiment and Theory. *Journal of Personality and Social Psychology, 62(5), 826-836.*

**Faravelli, C., Albanesi, G. & Poli, E. (1986).** Assessment of depression: A comparison of rating scales. *Journal of Affective Disorders, 11(3) , 245 -253.*

**Fennell, M.J.V. & Campbell, E.A. (1984**). The Cognitions Questionnaire: Specific Thinking Errors in Depression. *British Journal of Clinical Psychology, 23, 81-92.*

**Ganellen, R.J. (1988).** Specificity of Attributions and Overgeneralization in Depression and Anxiety*. Journal of Abnormal Psychology, 91(1), 83-86*.

**Hunt, M., Auriemma, J. & Cashaw, A.C.A (2003).** Self-Report Bias and Underreporting of Depression on the BDI-II. *Journal of Personality Assessment, 80:1, 26-30,*

**Kernis, M. H., Brockner, J., & Frankel, B. S. (1989**). Self-esteem and reactions to failure: The mediating role of overgeneralization. *Journal of Personality and Social Psychology, 57, 707–714.*

**King, M.J., MacDougall, G.A., Ferris, M.S., Levine, B., MacQueen, G.M. & McKinnon, M.C. (2010).**A review of factors that moderate autobiographical memory performance in patients with major depressive disorder.  *Journal of Clinical and Experimental Neuropsychology, 32(10), 1122-1144.*

**Kircanski, K., Joormann, J., & Gotlib, I. H. (2012).** Cognitive Aspects of Depression. *Wiley Interdisciplinary Reviews. Cognitive Science, 3(3), 301–313*.

**Knight, R.G., Waal-Manning, H.J., Spears, G.F. (1983**). Some norms and reliability data for the State-Trait Anxiety Inventory and the Zung Self-Rating Depression scale. *British Journal of Clinical Psychology, 22(4), 245-249.*

**Lee, H.C., Chiu, H.F.K., Wing, Y.K., Leung, C.M., Kwong, P.K. & Chung, D.W.S. (1994).** The Zung Self-Rating Depression Scale: Screening for Depression Among the Hong Kong Chinese Elderly. *Journal of Geriatric Psychiatry and Neurology, 7(4), 216-220.*

**Lefebvre, M. F. (1981**). Cognitive distortion and cognitive errors in depressed psychiatric and low back pain patients. *Journal of Consulting and Clinical Psychology*, *49*, 517–525.

**Lewinsohn, P. M., Steinmetz, J. L., Larson, D. W., Franklin, J. (1981).**Depression-related cognitions: Antecedent or consequence? J*ournal of Abnormal Psychology, 90, 213–219.*

**Klar, Y., Gabai, T. & Baron, Y. (1996).** Depression and generalizations about the future: Who overgeneralizes what? *Personal and Individual Differences, 22(4), 575-584.*

**MacLeod, C., Koster, E.H.W. & Fox, E. (2009).** Whither cognitive bias modification research? Commentary on the special section articles. *Journal of Abnormal Psychology, 118, 89-99.*

**MacLeod, A.K. & Williams, J.M.G. (1990).** Overgeneralization: An important but non-homogenous construct. *British Journal of Clinical Psychology, 29, 443-444.*

**Miranda, J., Persons, J. B., & Byers, C. N. (1990).** Endorsement of dysfunctional beliefs depends on current mood state. *Journal of Abnormal Psychology, 99*(3*), 237-241*.

**Mogoaşe, C., Brăilean, A. & David, D. (2013).** Can Concreteness Training Alone Reduce Depressive Symptoms? A Randomized Pilot Study Using an Internet-Delivered Protocol.

*Cognitive Therapy Research, 37(4), 704-712.*

**Piccinelli, M., & Wilkinson, G. (2000).** Gender differences in depression: Critical review. *British Journal of Psychiatry, 177(6), 486-492.*

**Power, M. J., Katz, R., McGuffin, P., Duggan, C. F., Lam, D., Beck, A. T. (1994**). The Dysfunctional Attitude Scale (DAS): A comparison of Forms A and B and proposals for a new subscaled version. *Journal of Research in Personality, 28, 263-276.*

**Rectar, N.A., Segal, Z.V. & Gemar, M. (1998).** Schema Research in Depression: A Canadian Perspective. *Canadian Journal of Behavioural Science, 30(4), 213-224.*

**Segal, Z. V., Gemar, M., & Williams, S. (1999).** Differential cognitive response to a mood challenge following successful cognitive therapy or pharmacotherapy for unipolar depression. *Journal of Abnormal Psychology, 108(1), 3-10.*

**Sim, K., Lau, K.W., Sim, J., Sum, M.Y. & Baldessarini, R.J. (2016).** Prevention of Relapse and Recurrence in Adults with Major Depressive Disorder: Systematic Review and Meta-Analyses of Controlled Trials, *International Journal of Neuropsychopharmacology*, 19(2), 1-13.

**Spielberger, C.D., Gorsuch, R.L. & Lushene, R.E. (1970).** Manual for the State-Trait Anxiety Inventory. *Palo Alto, CA: Consulting Psychologists Press.*

**Stokel-Walker, C. (2018).** Bots on Amazon’s Mechanical Turk are ruining psychology studies. New Scientist Magazine. Available online: <https://www.newscientist.com/article/2176436-bots-on-amazons-mechanical-turk-are-ruining-psychology-studies/>

Accessed 02/09/18.

**Tairi, T., Adams, B., & Zilikis, N. (2016**). Cognitive Errors in Greek Adolescents: The Linkages Between Negative Cognitive Errors and Anxious and Depressive Symptoms. *International Journal of Cognitive Therapy, 9(3), 261-278.*

**Thew, G.R., Gregory, J.D., Roberts, K. & Rimes, K.A. (2017).** The phenomenology of self‐critical thinking in people with depression, eating disorders, and in healthy individuals. *Psychology and Psychotherapy: Theory, Research and Practice, 90, 751-769.*

**van den Heuvel, T.J., Derksen, J.J., Eling, P.A. & van der Staak, C.P. (2012**). An investigation of different aspects of overgeneralization in patients with major depressive disorder and borderline personality disorder. *British Journal of Clinical Psychology, 51, 376-395.*

**Watkins, E.R. (2008).** Constructive and Unconstructive Repetitive Thought. *Psychological Bulletin, 134(2), 163-206.*

**Watkins, E. R., Baeyens, C. B., & Read, R. (2009).** Concreteness training reduces dysphoria: Proof-of-principle for repeated cognitive bias modification in depression. *Journal of Abnormal Psychology, 118*(*1), 55-64.*

**Weissman, A.N. & Beck, A.T. (1978).** Development and validation of the dysfunctional attitude scale. Paper presented at the Annual meeting of the Association for the Advanced Behavior Therapy, Chicago.

Available online at: <https://files.eric.ed.gov/fulltext/ED167619.pdf>

Accessed 08/2018

**Zung, W.W.K. (1965).** A self-rating depression scale. *Archives of General Psychiatry*, *12*, *63–70.*