

Eastern Michigan University
DigitalCommons@EMU

Master's Theses and Doctoral Dissertations

Master's Theses, and Doctoral Dissertations, and
Graduate Capstone Projects

2017

An examination of the psychometrics of the cognitive fusion questionnaire and its relationship to other constructs

Barry Eye

Follow this and additional works at: <http://commons.emich.edu/theses>

 Part of the [Psychology Commons](#)

Recommended Citation

Eye, Barry, "An examination of the psychometrics of the cognitive fusion questionnaire and its relationship to other constructs" (2017). *Master's Theses and Doctoral Dissertations*. 808.
<http://commons.emich.edu/theses/808>

This Open Access Dissertation is brought to you for free and open access by the Master's Theses, and Doctoral Dissertations, and Graduate Capstone Projects at DigitalCommons@EMU. It has been accepted for inclusion in Master's Theses and Doctoral Dissertations by an authorized administrator of DigitalCommons@EMU. For more information, please contact lib-ir@emich.edu.

Running head: CFQ PSYCHOMETRICS AND CONSTRUCT RELATIONS

An Examination of the Psychometrics of the Cognitive Fusion Questionnaire and Its
Relationship to Other Constructs

by

Barry Eye

Dissertation

Submitted to the Department of Psychology

Eastern Michigan University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Clinical Psychology

Dissertation Committee:

Ellen Koch, Ph.D., Chair

Tamara Loverich, Ph.D.

Thomas Waltz, Ph.D., Ph.D.

Katherine Porter, Ph.D.

March 10, 2017

Ypsilanti, Michigan

Acknowledgments

To my father for imparting his attention to detail and stubborn drive to do things “the right way”; for sharing his considerable font of knowledge—even when it was more than I needed—and helping create a sense of wonder about the world; and for his substantial (and almost successful) attempts to instill a sense of discipline.

To my mother for her tireless efforts at filling my younger days with vibrant experiences and inspiring a love of science; for always spoiling me with love, in spite of my impatient and sometimes acrimonious conduct; and for never doubting that I could do anything that I dreamed of doing.

Their boundless love and support are to blame for the most human and best parts of me. I hope I can provide as much, day in and day out, when it is my turn.

Abstract

More than one billion humans currently suffer from one or more mental health difficulties, the leading cause of disability in the world. Psychotherapy is well-established as efficacious and cost effective in the treatment of mental health difficulties, particularly the widely-used family of cognitive behavioral therapies (CBT). The most prominent, new CBT–acceptance and commitment therapy (ACT)–has shown efficacy equal to or better than traditional CBT across a range of such difficulties. ACT’s novel approach to language, *defusion* (the opposite of fusion), can help improve mental health by changing one’s relationship with their thoughts. Efforts to better understand this mechanism of ACT has been hampered by the absence of a robust fusion measure until recently, with the creation of the Cognitive Fusion Questionnaire (CFQ). The present study sought to confirm the psychometrics of the CFQ with a large adult undergraduate student sample from the U.S. and to further expand our empirical understanding of the relationship between fusion and other important clinical constructs. Results showed that the CFQ exhibited strong internal consistency reliability; a unidimensional factor structure; and construct, concurrent criterion, and incremental validity in relation to a number of other important clinical scales as predicted. However, the results also showed that the factor structure of the CFQ was shared with the predominant measure of the central ACT construct of psychological flexibility, implying the two are measuring the same underlying construct. These results, limitations of the present study and the CFQ, and future research directions are discussed.

Table of Contents

Acknowledgments.....	ii
Abstract.....	iii
List of Tables	viii
List of Figures	ix
An Examination of the Cognitive Fusion Questionnaire and Its Relationship to Other Constructs	1
Mental Health.....	2
Prevalence of difficulties.	2
Burden.....	3
Costs and expenditures.	4
Interventions.	6
ACT intervention research.....	6
Evidence of ACT efficacy and comparison to established therapies.....	8
Acceptance and Commitment Therapy.....	9
Philosophical and theoretical foundations.	9
Functional contextualism.....	10
Relational frame theory.....	10
Rule-governed behavior.....	11
Clinical model.....	12
Acceptance.....	15
Values.	15
Committed action.....	16

Contact with the present moment.	16
Self as context.	17
Defusion.	17
Toward an Understanding of Treatment Process.	19
Trends in psychotherapy research.	19
Empirical investigations of ACT.	20
Component processes.	20
Mechanisms of change.	21
Fusion and related constructs.	22
Related constructs.	22
Other measures of fusion.	24
The Cognitive Fusion Questionnaire.	25
The Present Study.	29
Hypotheses.	30
Method.	33
Participants.	33
A priori power analysis.	34
Procedure.	36
Measures.	36
Cognitive Fusion Questionnaire (CFQ).	36
Acceptance and Action Questionnaire-II (AAQ-II).	37
Behavioral Activation for Depression Scale-Short Form (BADSF).	38
Schwartz Outcome Scale-10 (SOS-10).	39

Depression Anxiety Stress Scales 21 (DASS21).....	39
Anxiety Sensitivity Index-3 (ASI-3).....	40
Results.....	41
Supporting Analyses	41
Detection and management of response bias.	42
Confirmation of assumptions of primary analytic approaches.	47
Primary Analyses	49
Internal consistency reliability.	49
Factor structure of the CFQ.	50
Relationship with the AAQ-II.....	52
Construct validity.....	54
Concurrent validity.	56
Incremental validity.	57
Discussion.....	60
Summary of Results	60
Relationship between fusion and psychological flexibility.	61
Limitations of the present study.....	66
Limitations of the CFQ and Future Directions	67
Construct validity.....	67
Generalizability.....	68
Mono-method assessment.....	68
Discriminant/divergent validity.	70
Requirements for clinical application.	72

Conclusion	73
References.....	74
Appendix A: IRB Approval Letter	103
Appendix B: Informed Consent Form	105
Appendix C: Demographics Questionnaire	108
Appendix D: AAQ-II	109
Appendix E: ASI-3	110
Appendix F: CFQ.....	111
Appendix G: BADS-SF	112
Appendix H: DASS21.....	113

List of Tables

<u>Table</u>	<u>Page</u>
<i>Table 1.</i> Summary of Replication and Extension Elements of the Present Study	28
<i>Table 2.</i> Racial/Ethnic Sample Demographics	34
<i>Table 3.</i> Scale/Subscale Statistics	50
<i>Table 4.</i> Factor Loadings for the EFA of the CFQ	52
<i>Table 5.</i> Factor Loadings for the EFA of the CFQ and AAQ-II Combined	53
<i>Table 6.</i> Bivariate Correlation Matrix of All Scales/Subscales	55
<i>Table 7.</i> Summary of Group Statistics and Comparison Analyses	57
<i>Table 8.</i> Summary of Regression Analysis Predicting SOS-10 Scores	58
<i>Table 9.</i> Summary of Regression Analysis Predicting DASS21 Scores	58
<i>Table 10.</i> Summary of Regression Analysis Predicting DASS21 Depression Scores	59
<i>Table 11.</i> Summary of Regression Analysis Predicting DASS21 Anxiety Scores.....	60

List of Figures

<u>Figure</u>	<u>Page</u>
<i>Figure 1.</i> The hexaflex: The ACT treatment model (Hayes et al., 2006).....	13
<i>Figure 2.</i> The ACT model of psychopathology (Hayes et al., 2006).	14

An Examination of the Cognitive Fusion Questionnaire and Its Relationship to Other Constructs

Mental health difficulties are an extraordinarily prevalent and costly problem, with more than one billion humans on the planet currently suffering from a mental health condition. It is the leading cause of disability, with global cost estimates of \$2.5 trillion dollars a year, and therefore, it requires effective and efficient treatments that can be implemented on a larger scale. Psychotherapy is well-established as efficacious in the treatment of mental health difficulties and shown to be more cost effective than alternatives, such as psychopharmacological treatments (American Psychological Association [APA], 2012). Within psychotherapy, the family of cognitive behavioral therapies (CBT) has the most empirical support (APA, 1995; see the Division 12's website on "Research-Supported Psychological Treatments" for updated information about empirical support), including the relatively newer acceptance and commitment therapy (ACT). ACT has shown equal or better efficacy than traditional CBT across a range of mental and physical health difficulties, with improvements lasting well after intervention. Perhaps the most novel contribution of ACT is its unique approach to language, a change process called *defusion* (the opposite of fusion), to help improve mental health by changing one's relationship with their thoughts. ACT and its community of researchers are representative of positive trends in the science of psychotherapy toward a more precise understanding of the mechanisms of improvement, but until recently, there was no robust measure of the important construct of fusion to properly facilitate such research. Recently, the Cognitive Fusion Questionnaire (CFQ) was developed in the U.K. by established ACT researchers as such a tool (Gillanders et

al., 2014). However, it is still very new, having only been used in seven peer-reviewed articles since its development. The present study therefore sought to confirm the psychometric properties of the CFQ with a very large, adult, undergraduate student sample from the U.S., and further expand an empirical understanding of the relationship between fusion and other important clinical constructs.

Mental Health

Prevalence of difficulties. Mental health difficulty is present in a significant portion of the world's population. Recent estimates from the Global Burden of Disease Study show that of the 6.9 billion humans on the planet, 15.3%—more than one billion people—currently suffer from a mental or behavioral disorder (Whiteford et al., 2013). The study utilized data gathered based on diagnostic criteria for 20 major disorders under multiple versions of either the *Diagnostic and Statistical Manual of Mental Disorders* or the *International Classification of Diseases*, thereby providing an underestimate of the true global prevalence, had other disorders (e.g., personality disorders) been included. Statistical analyses revealed no significant differences between classification systems or versions, and of the 20 disorders captured by the estimates, the highest prevalence rates were for anxiety disorders (274 million, 4.0%), major depressive disorder (MDD; 405 million or 5.9% of the world's population), and substance use disorders (SUDs; 147.6 million, 2.1%). Other, related estimates also show that more than 800,000 people die each year by suicide (World Health Organization [WHO], 2014).

In the U.S., prevalence rates are even higher. Annual estimates in the U.S. show that 26.2% of adult Americans suffer from a mental illness, including anxiety disorders (18.1%), MDD (6.7%), alcohol use disorders (AUDs; 4.4%), and other SUDs (1.8%;

Kessler, Chiu, Demler, & Walters, 2005). More recent estimates of SUDs show a higher annual prevalence of 8.5% across all substances (Substance Abuse and Mental Health Services Administration [SAMHSA], 2013b), with 3.3% estimated to be suffering from both a SUD and a mental health diagnosis (SAMHSA, 2013a). Adult lifetime prevalence is near half of the population (46.4%), with relatively prevalent comorbidity including anxiety disorders (28.8%), MDD (16.6%), AUDs (18.6%), and other SUDs (10.9%; Kessler, Berglund et al., 2005). Estimates for children in the U.S. are similar (though child-specific diagnoses makeup much of the overall prevalence) with previous year prevalence around 13.1%, including anxiety disorders (0.7%) and MDD (2.7%; for children aged 8 to 15 years old; Merikangas et al., 2009); and lifetime prevalence around 46.3%, including anxiety disorders (25.1%) and MDD (11.2%; for children aged 13 to 18 years old; Merikangas et al., 2010). Suicide is the 10th leading cause of death in the U.S. for all individuals over 10 years old, accounting for 38,364 deaths a year. This is in addition to a reported 487,700 individuals seeking treatment in hospital emergency rooms for self-inflicted injuries (Centers for Disease Control and Prevention [CDC], 2012).

Burden. The overall burden of mental health concerns is estimated to be enormous. The most comprehensive and widely cited data uses a method of estimation for the burden of mental illness called disability-adjusted life years (DALYs), which are the sum of years of life lost (YLLs) and years lived with disability (YLDs)—with disability being defined as any short-term or long-term health loss (other than death; Murray et al., 2012). Recent estimates show that mental and behavioral disorders accounted for 185,190,000 DALYs (or 7.4% of all DALYs globally), including anxiety disorders (1.1% of all DALYs globally), MDD (2.5%), AUDs (0.7%), and other SUDs

(0.8%). This estimate, however, is likely a significant underestimate of the total burden of mental health due to artificially low YLLs (zero for almost all mental health concerns) because the cause of death in the calculation is attributed to the specific physical cause, rather than the underlying cause(s) of death—one of which is frequently mental health (e.g., almost a million suicide deaths annually, worldwide; WHO, 2014).

The YLDs component of the DALYs shows that mental health is the leading cause of disability worldwide, accounting for 175,000,000 YLDs (22.9%; Whiteford et al., 2013). The major types of disorders contributing to disability within mental health included anxiety disorders, which accounted for 3.5% of all YLDs; MDD (9.7%); AUDs (1.8%); and other SUDs (2.2%). Additionally, between 1990 and 2010 the burden of mental health increased by 37.6% (though mainly due to population growth and aging).

Parallel to prevalence, the estimated proportion of burden of mental health in the U.S. is higher than the global burden. Mental and behavioral disorders were estimated to account for 11,139,100 DALYs (13.6%; Murray et al., 2013), including anxiety disorders (2.3% of all U.S. DALYs), MDD (3.7%), AUDs (1.4%), and other SUDs (2.6%; Institute for Health Metrics and Evaluation [IHME], 2014). Like global estimates, estimates for the U.S. also showed mental health as the largest cause of disability, contributing 9,945,260 YLDs (27.1% of all U.S. YLDs). This burden was primarily due to anxiety disorders (5.1% of all U.S. YLDs), MDD (8.3%), AUDs (2.3%), and other SUDs (3.5%).

Costs and expenditures. In addition to the more intangible estimates of burden due to disability and life lost, quantifiable assessments have been made estimating the cost of mental health problems. The annual cost of mental health globally is estimated to be around \$2.5 trillion (3.5% of the world's gross domestic product [GDP]), with roughly

one third of that cost consisting of direct treatment and the other two thirds consisting of indirect costs such as lost income and productivity (Bloom et al., 2011).

Assessments of the cost of mental health in the U.S. are actually proportionally lower than global assessments, with estimates of \$317.6 billion (2.0% of U.S. GDP), including \$100.1 billion in direct treatment costs and the rest in disability benefits and lost wages (Insel, 2008). However, such assessments are certain to be underestimates as they do not account for lost productivity due to premature death; comorbid conditions; institutionalization and incarceration; homelessness; and aid from friends, family, and other social support sources that assist with the financial burdens of those with mental health difficulties—among other costs. For example, more comprehensive assessments of the cost of SUDs in the U.S. estimate the cost of AUDs at \$235 billion per year with only \$30 billion in health care—the rest is attributable to lost productivity, crime, and other indirect costs (Rehm et al., 2009). Similarly, other SUDs are estimated to cost the U.S. \$386 billion per year, with only \$107 billion in direct treatment costs (U.S. Department of Justice, 2011). Suicide, another less common specific concern, is estimated to cost the U.S. \$34.6 billion a year and self-inflicted injuries another \$6.5 billion (CDC, 2012). Taken together, the more detailed estimated costs of just suicide and SUDs are already more than double the (clearly incomplete) overall estimates.

The U.S. government alone spends almost \$1 trillion a year on health (16.2% of all government spending; WHO, 2011). Only 6.2% of this health spending (1.0% of total spending) is on mental health—though this is more than double the world median (2.8%). Further, government spending is the majority (58%) of all spending on mental health. As

such, treatments are needed for a host of mental disorders and concerns, in formats that are both effective and cost efficient on a public health scale.

Interventions. Psychotherapy is one form of effective mental health treatment. The efficacy of psychotherapy is well-established, with significant and large effects across diagnostic conditions that last longer and are more cost effective than alternatives, like psychotropics (APA, 2012). The psychotherapeutic approach with by far the most empirical support for efficacy is the family of CBT. CBT has been shown to be very efficacious for a substantial range of problems and diagnoses, including anxiety, mood, trauma, interpersonal, and somatic problems across the lifespan (Butler, Chapman, Forman, & Beck, 2006). Not only is the empirical support for CBT substantial, the large majority of treatment comparison studies show higher response rates to CBT than to comparison treatments (Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012).

The newest group of CBT approaches has growing empirical support (Kahl, Winter, & Schweiger, 2012). The specific therapy with the most empirical support within the newest generation is ACT (A-Tjak et al., 2015). ACT was designed as a process-oriented, transdiagnostic treatment, and as such, it has been investigated for efficacy across an array of psychological problems.

ACT intervention research. ACT interventions have had significant variation in both format and duration. The most common modality is individual, including in-person (e.g., Arch et al., 2012) and internet-delivered interventions (e.g., Hesser et al., 2012). Other modalities include group therapy (e.g., Ossman, Wilson, Storaasli, & McNeill, 2006), workshop format (e.g., Blackledge & Hayes, 2006), and self-help interventions (e.g., Johnston, Foster, Shennan, Starkey, & Johnson, 2010). Protocols have ranged from

1 to 48 sessions ($M = 12.1$) consisting of 3 to 24 ($M = 6.6$) total treatment hours, across 1 to 16 weeks ($M = 5.2$; Hayes, Pankey, Gifford, Batten, & Quiñones, 2002; Öst, 2008). Notably, Powers and colleagues (2009) did not find a significant dose-response relationship in their meta-analysis of ACT, indicating that shorter interventions may be as effective as longer treatments. Multiple studies have shown improvements from ACT interventions being maintained up to three years post-treatment (e.g., Vowles, McCracken, & O'Brien, 2011) and with even as few as four sessions creating improvements that lasted over one year (e.g., Bach, Hayes, & Gallop, 2012).

ACT has been used for a diverse range of clinical and non-clinical issues. Within clinical psychology, ACT has been used to target depression (e.g., Zettle & Rains, 1989); anxiety disorders, including generalized anxiety disorder (e.g., Wetherell et al., 2011), obsessive-compulsive disorder (e.g., Twohig et al., 2010), social anxiety disorder (e.g., Dalrymple & Herbert, 2007), and other anxiety disorders (see Swain, Hancock, Hainsworth, & Bowman, 2013, for a review); psychosis (e.g., Bach & Hayes, 2002); substance use, including smoking (e.g., Gifford et al., 2004), methamphetamines (e.g., Smout et al., 2010), and polysubstance abuse (Hayes, Wilson et al., 2004); and personality disorders such as borderline personality disorder (Gratz et al., 2008). ACT has further been applied to numerous medical issues within health psychology, including chronic pain (e.g., Dahl, Wilson, & Nilsson, 2004), epilepsy (e.g., Lundgren, Dahl, Melin, & Kies, 2006), diabetes management (e.g., Gregg, Callaghan, Hayes, & Glenn-Lawson, 2007), and weight control (e.g., Lillis, Hayes, Bunting, & Masuda, 2009). ACT has also been used with at-risk samples to improve stress (e.g., Bond & Bunce, 2000), general distress (e.g., Lappalainen et al., 2007), sub-clinical depression (e.g., Forman,

Herbert, Moitra, Yeomans, & Geller, 2007), and anxiety symptomology (e.g., Muto, Hayes, & Jeffcoat, 2011). Finally, ACT has also been used with non-clinical samples to address issues such as racial prejudice (e.g., Lillis & Hayes, 2007) and stigma towards mental illness (e.g., Masuda et al., 2007) as well as performance, such as rowing (e.g., Fernández, Secades, Terrados, García, & García, 2004) and chess (e.g., Ruiz & Luciano, 2009).

Evidence of ACT efficacy and comparison to established therapies. Evidence for the efficacy of ACT has been accumulating across a growing number of clinical concerns. The first meta-analysis of ACT outcome studies showed a large effect size compared to control conditions ($d = 0.99$; wait list, placebo, and treatment as usual control), and a moderate effect size compared to other active treatments ($d = 0.48$; Hayes et al., 2006). The first independent meta-analysis was actually of all new generation CBTs but included ACT (Öst, 2008). It found a large effect size compared to wait list conditions ($g = 0.96$) and treatment as usual conditions ($g = 0.79$), and a moderate effect size compared to other active treatments ($g = 0.53$). Soon after, another meta-analysis found that ACT showed moderate-to-large effect sizes compared to wait list and psychological placebo conditions ($g = 0.68$) and moderate effect sizes compared to treatment as usual conditions ($g = 0.42$), but was not superior to established treatments ($g = 0.18, p = 0.13$; Powers, Zum Vörde Sive Vörding, & Emmelkamp, 2009). These conclusions were challenged and reanalyzed, finding that ACT was in fact superior to established treatments ($g = 0.27, p = 0.03$; Levin & Hayes, 2009).

More recently, another meta-analysis of ACT focused on comparisons with traditional CBT (Ruiz, 2012). While some have argued that ACT is no different from

traditional CBT as a technology (see Hofmann & Asmundson, 2008, for a discussion), analysis showed that ACT outperformed traditional CBT in 11 out of 15 studies on the primary outcome measure at post-treatment, with only 2 out of 15 in favor of CBT (Ruiz, 2012). Overall, ACT showed a small-to-moderate effect size compared to CBT ($g = 0.42$), with advantages at post-treatment ($g = 0.37$) and follow-up ($g = 0.42$). The studies included both individual and group treatment approaches ranging from brief to longer duration (1 to 17 sessions), with a range of populations (e.g., clinical and non-clinical, psychological and medical health) for a variety of concerns (e.g., anxiety, depression, smoking cessation, chronic pain, stress). Current evidence, therefore, appears to suggest that ACT is at least as efficacious as traditional CBT for many symptoms and concerns.

Acceptance and Commitment Therapy

ACT is a transdiagnostic modern behavioral therapy that shares some similarities with traditional CBT, but it also has some important differences (Hayes, Strosahl, & Wilson, 2012). Based on functional contextualism and an empirical theory of language and cognition, ACT has a unified theoretical model of six change processes that interrelate with the core idea of psychological flexibility. ACT has been garnering accelerated empirical support as a clinical intervention in the past two decades, with evidence of efficacy across a multitude of populations and conditions. Further, this supporting literature includes investigations into the processes and mechanisms of ACT, much of which has supported ACT's underlying model and philosophy.

Philosophical and theoretical foundations. ACT flows from the behavior analytic tradition and has the pragmatic philosophy of functional contextualism at its foundation. It is also based on a comprehensive, experimental accounting of human

language and cognition known as relational frame theory (RFT), which includes an analysis of rule-governed behavior.

Functional contextualism. The philosophy of ACT is grounded in functional contextualism (Hayes et al., 2012). Two key elements of functional contextualism that are germane to ACT and its clinical application are its chosen unit of analysis and its truth criterion. Functional contextualism focuses on the whole, ongoing event in context (Hayes, Hayes, & Reese, 1988). Inherent in this focus is the idea that behaviors of certain topographies may vary in function based on differing contexts, which also include things like the individual's learning history. The key to understanding a target behavior then becomes its function, as that is critical to intervention. Behavior also includes all internal experiences (i.e., thoughts, feelings, and sensations), and concomitant with the function-focused approach is the idea that none of these experiences are inherently good or bad; rather, the question is how they function for that individual within a given context. This focus is directed by the truth criterion of functional contextualism: effective action. The goal of ACT then is to predict and influence behavior in the direction of pragmatic workability, and it can be judged based on meaningful clinical improvement (Twohig, 2012).

Relational frame theory. ACT is also built on the idea that human language is a double-edged sword that can be responsible for both achievement and psychological pain, and it is connected to a contemporary behavior analytic accounting of human language and cognition: RFT (Hayes, Barnes-Holmes, & Roche, 2001). RFT is an empirically-supported theory that explains how symbolic thought contributes both to our evolutionary success as a species and unique ability to suffer covertly (for an introduction see Törneke,

2010). A key conclusion from RFT is that humans' cognitive ability allows us to relate events under *arbitrary* contextual control (e.g., we can learn that a dime is more valuable than a nickel), rather than simply responding to the formal properties of stimuli (in spite of a dime being physically smaller than a nickel; Hayes, 2004). This makes it possible for relational learning to transfer from relations based on direct experience and generalization, to indirect relations that are based on a learned ability to relate stimuli mutually and in combination. The result is that instead of attending to direct environmental contingencies, the function of behavior can be driven by a network of relations that are altered by contextual cues (Twohig, 2012). In this manner, derived stimulus relations are created, which in many cases are automatically reinforced (e.g., by confirming a held assumption) and can lead to the creation and entrenchment of behavior that is ultimately problematic for the individual. ACT emphasizes functional context in part because studies have shown that the functional context can be modified independently of the stimulus (Hooper, Saunders, & McHugh, 2010), while relational contexts cannot be unlearned once trained (Wilson & Hayes, 1996).

Rule-governed behavior. One implication of the relational framing of verbal behavior that ACT addresses is rule-governed behavior, or behavior that is directed by a particular verbal context (Drossel, Waltz, & Hayes, 2007). Within networks of cognitive contingencies created via relational framing, verbal rules can come to govern behavior independent from other direct-acting environmental contingencies. While this ability is beneficial in allowing for behavior without direct experience (e.g., planning), it can also lead to problematic patterns of behaving. There are three types of rule-following behavior that are frequently relevant in ACT which differ from one another on the

context of behavior maintenance: augmenting, tracking, and pliance. Augmentals are instructions that transform the stimulus function of an event to either establish events as reinforcing/punishing or alter the value of reinforcement/punishment of an event. In ACT, augmentals are manipulated in values clarification work (e.g., in order to shift control of behavior from immediate contingencies to more workable delayed consequences; Hayes et al., 2012). Tracking describes behavior that is driven by the environmental consequences (i.e., contingencies) of the behavior itself, while rules function only as guides (Drossel et al., 2007).

In contrast, pliance describes behavior that follows a rule(s) and generally ignores feedback from the environment. It is based on a learning history that rewards rule following (generally or for specific behaviors or classes of behaviors); that is, it is initially mediated by real social demands and can come under the control of perceived/constructed social demands including self-generated demands. Research has also shown that cognitive rule following makes people less sensitive to environmental contingencies (Hayes, 1989). ACT utilizes approaches such as metaphors and experiential exercises in part to minimize the risk of developing a repertoire of behavior in clients that is rigid and rule-bound, and instead cultivates the psychological flexibility required to follow rules when they are effective and choose different actions when they are shown to not be useful for them.

Clinical model. The ACT model is simultaneously a model of psychopathology and treatment (Hayes et al., 2012). It is centered on the idea of psychological flexibility, which is “the ability to contact the present moment more fully as a conscious human being, and to change or persist in behavior when doing so serves valued ends” (Hayes,

Luoma, Bond, Masuda, & Lillis, 2006, p. 7). In the model, six interdependent change processes are grouped into two overlapping sets: (a) commitment and behavior change processes, comprised of contact with the present moment, self as context, values, and committed action; and (b) mindfulness and acceptance processes, made up of contact with the present moment, self as context, acceptance, and defusion (see Figure 1).

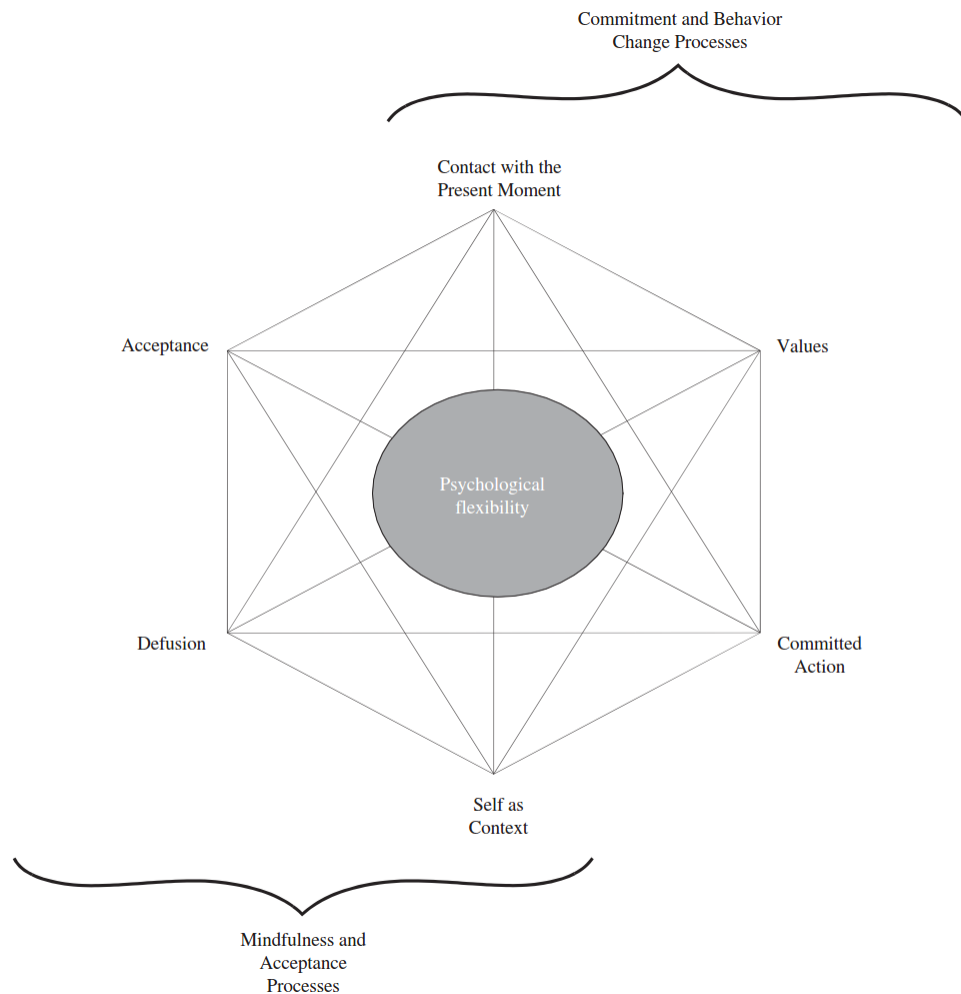


Figure 1. The hexaflex: The ACT treatment model (Hayes et al., 2006).

Correspondingly, the core of the ACT model of psychopathology is psychological *in*flexibility. Each of the six psychological processes of change supporting psychological

flexibility has a foil process. These six foil processes—dominance of the conceptualized past and feared future, lack of values clarity, inaction, impulsivity, or avoidant persistence, attachment to the conceptualized self, cognitive fusion, and experiential avoidance—are all interrelated and promote psychological inflexibility, which is frequently characterized by a narrowing of one’s behavioral repertoire resulting in distress (see Figure 2).

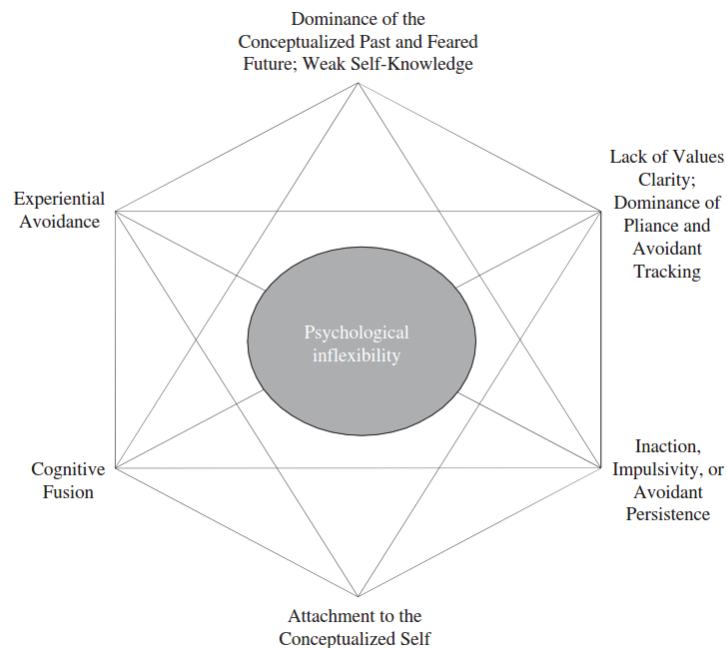


Figure 2. The ACT model of psychopathology (Hayes et al., 2006).

The paradigm of ACT is based on the premise that psychological flexibility represents a unified model of human functioning and adaptability (Hayes et al., 2012). It posits that pain is a natural part of life, but suffering is caused by some form of psychological inflexibility. Clinically, ACT employs the six change processes towards

the goal of increasing contextual control on verbal/cognitive processes such that a person can engage the positive consequences of their behavior in pursuit of valued living.

Acceptance. Acceptance is the action of allowing and engaging with experiences (both internal and external), as they are occurring (Twohig, 2012). Importantly, acceptance includes a nonjudgmental posture towards experience (i.e., it is not the same as tolerance of experiences). The problematic inverse of acceptance is experiential avoidance—a control process that reduces or eschews experiences expected to be distressing (based either on past direct experience or verbally established relations) via suppression, escape, or similar behaviors. Avoidance behaviors are more likely to become rigid, rule-governed behavior because they are commonly developed under aversive control (Folkman, Lazarus, Gruen, & DeLongis, 1986). Patterns of experiential avoidance grow, in part because of how easily verbal events become related to aversive events within mutually entailed relational frames (Hayes et al., 2012). This contributes to the resistance of avoidance to extinction, as they are then maintained by reductions in aversive private events (negative reinforcement; Ruiz, 2010). The goal of acceptance, then, is to reduce attempts to change the form or frequency of unwanted private events and to curtail behavioral avoidance when those efforts cause psychological harm or otherwise move away from valued directions.

Values. Values are qualities freely chosen (i.e., in a context free of aversive control) that motivate, guide, and fulfill life (Hayes et al., 2012). They are distinct from goals in that they are directions or ways of being (process) rather than destinations or things to achieve (outcome). The opposite condition of clear contact with chosen values is one where behavior is driven by reactivity or a desire to please others without regard

for self, and whose experience is characterized by a lack of direction and motivation and frequent feelings of emptiness. Functionally, values clarification is done to increase tracking of an individual's behavior, provide augmentals that transform the reinforcing/punishing level of values-relevant events, and (ultimately) increase psychological flexibility (Twohig, 2012). The process of constructing values is meant to put people back in touch with the things that give their life meaning and identify the desirable elements in order to have those elements motivate their behavior.

Committed action. Committed action is an intentional and continuously re-directing way of behaving (including private mental activity) that is values-based (Hayes et al., 2012). Opposing processes include impulsivity and apathy, avoidant persistence, and other forms of inaction. The goal of committed action is the development of ever-increasing patterns of effective action in valued directions, and this is accomplished via development and pursuit of short-, medium-, and long-term goals in the service of identified values. Efforts to live one's values will necessarily include the utilization of the other five ACT processes. ACT protocols frequently include more traditional behavior therapy components such as exposure and skills training as part of committed action work to help achieve these values-based goals.

Contact with the present moment. This ACT process seeks to develop the ability to intentionally (voluntarily) notice ongoing experiences (both internal and external) as they are occurring, in an open and nonjudgmental manner (Twohig, 2012). The goal is not to have attention firmly maintained in the present at all times, but rather that one is able to notice when awareness is not in the present and be flexible enough to be able to shift the focus back to the now when doing so would be helpful. The problematic inverse of

mindfulness is being directly focused on events from the past (e.g., severe guilt) or potential events in the future (e.g., impairing worry). This inflexibility of being stuck outside the present inhibits a person's ability to respond to changing demands within their environment by altering their behavior and acting in accordance with their own values (Hayes et al., 2012). In a more technical sense, mindfulness helps to transition control over one's behavior from verbally constructed reports of self to an increased awareness of present environmental contingencies.

Self as context. Self as context is the ability to have ongoing self-awareness as a context for verbal knowing (Hayes et al., 2006). The goal is to be able to be aware of one's reality without becoming invested in what events are experienced or how they are experienced. The opposite process of self as context is an over-attachment to a concept of self and the prioritization of the protection of this concept above and beyond effective action in the present moment (Twohig, 2012). Self as context is a perspective taking skill (a viewpoint one looks from) that improves flexibility in responding to one's experience and fosters an expanded social consciousness and sense of compassion. This is accomplished by training up deictic (by demonstration) verbal relations to enhance perspective taking and theory of mind skills (Weil, Hayes, & Capurro, 2011).

Defusion. Defusion is the process of disentangling from distressing private events in order to view them simply as ongoing mental activity (Hayes et al., 2012). The problematic opposite—cognitive fusion—is an over-attachment to the literal contents of thoughts that allows them to hold considerable power over responding. Clinically, fusion generally presents as selective sensitivity to environmental contingencies based upon pliance to verbal rules. That is, cognitive fusion is generally characterized by poor

tracking of punishing contingencies for behavior, excessive tracking of reinforcing contingencies for behavior, or the inverse for alternative behaviors (or both). This tracking bias skews behavior towards learned verbal relations and away from direct environmental feedback. Defusion attempts to alter the function (but not form, frequency, etc.) of private events by changing the way one relates to them, in order to decrease their believability (Hayes et al., 2006). The dominance of verbal events is reduced primarily by increasing psychological distance from thoughts by viewing the mind as a distinct part of experience rather than intrinsic to our perspective, generating skepticism for the content the mind produces, and shifting away from a literal understanding of the products of the mind. This is done with the goal of allowing for more flexible tracking of consequences based on action taken toward values.

The approach towards language within ACT and RFT, with its shift from the content of thoughts to our relationship with our thoughts, is arguably its largest novel contribution to the intervention literature and CBT family of treatments; much of the new generation of CBTs is characterized by mindfulness- and acceptance-based approaches, and committed action can be viewed as an application of behavioral activation from the perspective of values. Because of this, the central construct relevant to psychologically problematic language-based behavior—fusion—was the focus of the present study. And the goal of the present study was additional validation of an instrument able to measure fusion with precision and fidelity. This is hoped to contribute to progress towards an increasingly refined understanding of the key processes of intervention, ultimately resulting in enhanced efficiency and effectiveness of treatment for psychological difficulty.

Toward an Understanding of Treatment Process

ACT development, like the behavioral traditions in which it was based, has rejected a categorical model of psychopathology in favor of a dimensional model that is inherently transdiagnostic and focused on process, rather than content, of treatment (Hayes, 2004). In this way, it has been an important part of general trends in cognitive-behavioral intervention research.

Trends in psychotherapy research. In the 1960s, psychologists began noting a low practical yield from the theory-heavy earliest decades of the field (e.g., Lazarus, 1967). In the 1980s and into the 1990s, researchers observed a trend in psychotherapy research towards a more pragmatic or applied tendency (Omer & Dar, 1992). This led to the identification of multiple treatment packages from the CBT family that were efficacious for a number of disorders. Research then proceeded to evaluate the relative efficacy of treatments against their competitor packages, coinciding with the rise of the prominence of randomized controlled trials (RCTs) within psychotherapy research. Intervention research then also began to direct efforts at a better understanding of efficacy at a more elemental level, beginning with so-called dismantling studies, which compared treatment components or packages with and without various ingredients (as many of the efficacious packages had common or overlapping parts). Such studies sought to understand what precisely were the active ingredients resulting in efficaciousness (e.g., Cahill, Carrigan, & Frueh, 1999) as well as improve cost-effectiveness and efficiency (e.g., Van Brunt, 2000).

Slightly lagging behind the trend in psychotherapy research from overall outcomes to specific mechanisms, there has also been a parallel shift from treatments

being developed and investigated to target specific diagnoses and their idiosyncrasies to an emphasis on transdiagnostic constructs representing more universal mechanisms of intervention. Among the most commonly studied are ideas such as avoidance (e.g., experiential avoidance; Chawla & Ostafin, 2007), tolerance/intolerance (e.g., distress tolerance; Zvolensky, Bernstein, & Vujanovic, 2010), sensitivity/reactivity (e.g., anxiety sensitivity; Reiss, Peterson, Gursky, & McNally, 1986), or difficulties in emotion regulation skillfulness (e.g., emotion regulation; Gross, 2009). Along with theoretical investigations of universal factors, transdiagnostic treatments were developed to target such constructs, cutting across diagnostic lines (e.g., Universal Protocol for Transdiagnostic Treatment of Emotional Disorders; Barlow et al., 2010). ACT, and work done by the ACT research community, embodies both of these trends.

Empirical investigations of ACT. As previously mentioned, ACT is a transdiagnostic treatment whose efficacy has been investigated across a broad array of psychological problems. Its creators have made a concerted effort to research the active processes throughout development, via methods such as moderation and mediation in order to test the theoretical model and understand how ACT is impacting recipients.

Component processes. Research by the ACT community has examined whether and how the theoretical model of ACT operates clinically via correlational studies, component examinations, and investigations of mechanisms of change. More than 30 studies have found that the central element of the ACT model, psychological flexibility, negatively correlates with an extensive list of established measures of common symptom clusters (e.g., depression, anxiety, worry, trauma symptoms, and pain) and positively relates to quality of life, behavioral effectiveness (e.g., job performance), and general

health measures across a range of populations (for reviews see Hayes et al., 2006; Hayes, Levin, Plumb-Villardaga, Villatte, & Pistorello, 2013; Ruiz, 2010).

The developers of ACT have also made a concerted effort to investigate the six change processes within the ACT model to test the theory underlying the model. To date, all have at least some support (see Hayes et al., 2013, for a review). Veterans with PTSD showed greater reductions in symptoms when exposed to a full ACT protocol versus one without a self as context component (Williams, 2006). Defusion exercises have been shown to reduce believability of and subjective distress from negative self-relevant thoughts (Masuda, Hayes, Sackett, & Twohig, 2004). Acceptance has been shown to reduce reported distress (Gutiérrez, Luciano, Rodríguez, & Fnk, 2004). Values interventions have been shown to decrease negative physiological arousal during distressing tasks (Creswell et al., 2005) and increase task persistence (Gutiérrez et al., 2004). Numerous studies have shown mindfulness-based interventions to be efficacious for improving mood, anxiety, and other symptoms (see Hofmann, Sawyer, Witt, & Oh, 2010, for a review). Because committed action is based on more basic behavioral methods it has extensive support in the literature, however, dismantling studies have supported the cohesiveness of the ACT model by showing that committed action operates through changes in acceptance and values (Lundgren, Dahl, & Hayes, 2008).

Mechanisms of change. Beyond active component testing (i.e., identifying what is producing change), researchers have also examined the mechanisms of ACT interventions that can help explain *how* change is being produced. Currently, roughly 30 mediational and moderational analyses of ACT have been performed, including a few analyses done within a treatment comparison study (e.g., Zettle & Hayes, 1986, as

reanalyzed in Hayes et al., 2006; see Hayes et al., 2013, for a review). Results have consistently shown that measures of psychological flexibility (e.g., Bond & Bunce, 2000, as reanalyzed in Hayes et al., 2006) as well as measures of component processes such as defusion (e.g., Gaudiano & Herbert, 2006), acceptance (e.g., Gifford et al., 2004), and values (e.g., Lundgren et al., 2008) mediate improvements in target constructs (e.g., quality of life) and do so better than alternative mediators (Hayes et al., 2013). In most cases these studies have assessed mediators concurrently with outcomes (i.e., statistical mediators; e.g., Hayes, Bissett et al., 2004), but in a few cases temporal precedence was established giving stronger support for potential causal mediation (e.g., Gifford et al., 2004; Lundgren et al., 2008; Zettle & Hayes, 1986, as reanalyzed in Hayes et al., 2006). Again, these process examinations have cut across clinical and non-clinical populations (e.g., college students, treatment-seeking community members, hospitalized inpatients) and concerns (e.g., anxiety, depression, psychosis, stigma/prejudice, diabetes management), as well as modalities (e.g., group and individual therapy, workshops) showcasing the broad applicability and flexibility in applying the ACT model.

Fusion and related constructs. Unfortunately, a better empirical understanding of the process of fusion and its role in treatment efficacy has been hampered somewhat by a lack of a precise measurement tool of the construct itself. Despite its unique contribution to the CBT tradition, the above-mentioned studies and others have either measured only some aspects of fusion, or actually measured constructs that while overlapping to some degree, do not fully and solely equate to fusion.

Related constructs. The most common of these surrogate constructs include thought-action fusion, believability of thoughts, mentalization, and decentering.

Decentering, also referred to as “meta-cognitive awareness,” has been a construct of interest in cognitive therapy since before ACT’s fusion concept and is still prominent with approaches such as mindfulness-based cognitive therapy (MBCT). It overlaps with fusion in that both involve the relationship with thoughts shifting to a more detached stance; however, decentering also includes aspects of acceptance and self-compassion (Fresco, Segal, Buis, & Kennedy, 2007). These elements are explicitly represented in the common measure of decentering, the Experiences Questionnaire (Fresco, Moore et al., 2007), which when used to measure fusion complicates the empirical picture and detracts from the precision of any understanding gleaned.

Similarly, *mentalization*, or “reflective function,” (Fonagy & Target, 2002) has overlap with fusion but also goes beyond it. Mentalization includes the understanding of ones’ mental states *and* aspects of attachment and emotion regulation (Gumley, 2010) that are not a part of the process of defusion. These additional aspects again introduce additional variables into measurement that confound interpretation of results.

Believability of thoughts is another common proxy for fusion sometimes used which, inversely to the constructs already mentioned, fails to represent fusion fully. Although fusion encompasses believability, it also includes the literal, evaluative, emotion-eliciting, and overly analytical aspects of thought as well as additional elements such as the dominance of thoughts over behavior (Gillanders et al., 2014). Simple believability is neither necessary nor sufficient in the measurement of fusion.

Lastly, although the name of *thought-action fusion* (TAF; Shafran, Thordarson, & Rachman, 1996) would appear to label/describe ACT’s process, it instead represents an aspect of belief. TAF simply describes a metacognitive bias about the power of thoughts

and their influence on behavior. This type of bias is one potential problem of fused thinking—and somewhat common in those diagnosed with obsessive-compulsive disorder (Rachman, Thordarson, Shafran, & Woody, 1995)—but is by no means representative of the totality of fusion as a construct and is incongruent with the fundamental approach of ACT as it is focused on thought *content*.

Other measures of fusion. In addition to measuring different constructs as proxies for fusion, there have also been a few attempts at developing measures for fusion as it is defined within the ACT model. However, these attempts have generally suffered from the same problems of either narrowly representing believability of thoughts or including additional elements outside the concept of fusion. There is a scale originally developed for children but also psychometrically evaluated for adults that includes aspects of fusion: The Avoidance and Fusion Questionnaire for Youth (AFQ-Y; Greco, Lambert, & Baer, 2008). Unfortunately, the AFQ-Y also measures experiential avoidance, an overarching ACT concept. As the fusion elements of the measure have not been evaluated independently, it is problematic as a precise fusion instrument.

In line with the conceptual issues mentioned above, two instruments sometimes used to measure fusion only address the believability aspect of thoughts. The Believability of Anxious Feelings and Thoughts Scale (BAFT; Herzberg et al., 2012) focuses on believability and does so in a way that is specific to anxiety. The Automatic Thoughts Questionnaire (ATQ; Hollon & Kendall, 1980) originally only included a frequency scale; however, a believability scale was added in an ACT study in order to better measure the construct of fusion (Zettle & Hayes, 1987). Unfortunately, it still does

not capture other important aspects of fusion and, like the BAFT, is specific to content common with a particular clinical problem (depression).

Easily the most comprehensive and accurate fusion measure prior to the creation of the CFQ was the Drexel Defusion Scale (DDS; Forman et al., 2012). Although it does more holistically measure fusion, it does have a few limitations as an instrument. It involves the rating of vignettes to judge defusion tendency rather than a direct report of respondents' experiences of defusion and questions are assessed around these specific situations, making them (again) content specific. Additionally, the measure begins with a definition of the concept of defusion in the instructions, which may itself prime defused responding and therefore bias scores.

The Cognitive Fusion Questionnaire. Because much of the previous research around fusion has been confounded by numerous measurement issues due to instruments that are often narrowly defined, content specific measures with confounding elements separate from the construct of fusion, a group of researchers developed the CFQ (Gillanders et al., 2014). Its development was thorough, consisting of a series of five studies that will be covered here in some detail, as the present study is a partial replication and extension of this work.

Prior to initiating formal research, a group of four researchers with ACT knowledge and clinical practice experience generated 44 potential scale items that were concrete and behaviorally operationalized, with the intent of developing a unidimensional fusion scale. The researchers then requested feedback from an independent group of nine ACT experts, resulting in a partial rewording and paring of the original pool of questions down to 42 potential items.

The researchers then recruited six diverse adult samples from the U.K.: 242 people seeking ACT for stress management training, 215 people with a range of mental health difficulties, 133 people diagnosed with multiple sclerosis, 219 dementia caregivers, a non-clinical young adult sample of 592, and a community sample of 447 people (there was also a seventh sample used in many of the latter studies that consisted of 74 people, most having a history of a major depressive disorder diagnosis). In a series of two studies using generally robust statistical methods, they explored their pool of potential items and confirmed a robust unidimensional scale of seven fusion questions. They performed a parallel analysis, a sequence of multiple exploratory factor analyses (EFA), and additional correlational, regression, and confirmatory factor analyses (CFA) to initially identify a single factor with robust psychometrics. They then followed with a series of CFA analyses to confirm the model fit across a range of indices, resulting in a comprehensive confirmation of their original factor analytic model.

After identifying and confirming a fusion scale with robust structural psychometrics, the researchers then performed a series of three further studies using the seven samples to confirm other psychometric properties of the scale. Investigations included confirmations of test-retest reliability, sensitivity to treatment, and construct, divergent, incremental, criterion, and internal consistency validity. Across all studies and samples, the CFQ was compared to 24 measures of relevant constructs, finding acceptable or better psychometric properties.

Overall, Gillanders and colleagues (2014) developed an empirically derived fusion measure using generally robust statistical methods. The resulting CFQ is the first measure to directly assess fusion as it is conceptualized within the ACT theoretical

model and does so in a well-specified, behaviorally operationalized way. It is also a very brief, self-report measure, enhancing its utility via research efficiency. Further, it was developed by well-known researchers within the ACT community, and its item development process included consultation with other ACT “experts” separate from the research teams involved, with the goal of cementing a unidimensional measure of fusion that, while fully representing important aspects of the construct, also did not include related but distinct concepts. Lastly, the development studies included a diverse set of samples totaling over 1,800 participants, with the measure showing good psychometric properties across those examined.

In the discussion of the development of the CFQ, the researchers noted a number of limitations of their series of studies, particularly in the realm of clinical application. They noted deficits related to the study of clinical samples (e.g., ones diagnosed via standardized diagnostic interviews), the CFQ’s ability to measure change in fusion over longer periods of time or courses of treatment, and response to interventions specifically targeting fusion. They also noted a lack of investigation of the CFQ’s relationship to measures of related constructs such as decentering.

These are important limitations to note; however, the CFQ is very new. It has only been used in studies published in seven peer-reviewed articles since its original development. Only one of those articles was performed with U.S. samples (Fergus, 2015), and they—like the other applied studies using the CFQ—minimally reported psychometric properties. Before the CFQ is used as a confirmed measure of fusion in clinical studies, as well as theoretical studies expanding beyond the ACT model into other important constructs, it could benefit from additional investigations that confirm

and expand upon its psychometric foundations. The present study was proposed to do precisely that. Table 1 summarizes the elements of replication and extension of the present study, including specification of relevant measures utilized.

Table 1

Summary of Replication and Extension Elements of the Present Study

Psychometric Property	Original Development	Present Study
Internal Consistency Reliability	✓	✓
Factor Structure	✓ (x)	✓ combined with AAQ-II
Construct Validity	AAQ-II SMS FFMQ KIMS TCQ PBRS ATQ RSQ HADS Anxiety subscale HADS Depression subscale BDI-II SCL90-GSI CESD CORE-OM PDQ-4 MBI GHQ-12 WDAS WDDS DLSS VLQ (<i>behavior rating only</i>) WHOQOL-BREF GJSS	AAQ-II (✓) ASI-3 ASI-3 Cognitive subscale ASI-3 Social subscale ASI-3 Physical subscale BADSF BADSF Activation subscale BADSF Avoidance subscale DASS21 DASS21 Anxiety subscale DASS21 Depression subscale DASS21 Stress subscale SOS-10
Incremental Validity	predict GHQ-12 beyond AAQ-II predict CESD beyond PBRS, RSQ predict HADS beyond ICQ-H, AAQ-II	predict SOS-10 beyond AAQ-II predict DASS21 beyond AAQ-II predict DASS21-D beyond BADSF predict DASS21-A beyond ASI-3
Divergent Validity	BIDR-IM	(x)
Concurrent Validity	(x)	DASS21 ASI-3
Test-Retest Reliability	✓	(x)
Sensitivity to Treatment	✓	(x)

Note: Full measure names can be found elsewhere in this document or in Gillanders et al., 2014

Examination of the CFQ allowed for an additional investigation that is particularly important for the ACT model and its application. As discussed earlier, the ACT clinical model conceptualizes six related processes contributing to the central concept of psychological flexibility. The Acceptance and Action Questionnaire-II (AAQ-II; Bond et al., 2011) is an instrument that was designed to measure psychological flexibility, to which defusion is conceptualized as a distinct and contributing construct. In the series of development studies for the CFQ, the researchers examined the CFQ with the AAQ-II, finding high correlations as anticipated by the ACT model (Gillanders et al., 2014). However, due to concerns of interdependence between the two constructs, they also informally (i.e., statistical results were not reported) considered the factor structure of the two measures together. In three of the five samples, they found the psychometrically desirable result of two separate factors corresponding to the two scales of origin, but in the other two samples, they found that both scales loaded onto the same factor, implying they were measuring the same construct. This mixed result demands further investigation, which was also undertaken by the present study.

The Present Study

The present study sought to contribute to the trend of psychotherapy intervention research examining change at a process level by further exploring the psychometrics of the CFQ, the newest and arguably most robust measure of fusion. Fusion is a concept and process that represents a unique contribution of ACT, a prominent and growing approach within the CBT tradition—the most empirically supported type of therapy. The present investigation of the CFQ had two primary purposes using a very large, adult,

undergraduate student sample from the U.S.: (a) to confirm the basic psychometric properties of internal consistency reliability and factor structure (including an examination of the CFQ's structure in relation to that of the AAQ-II); and (b) to expand the usefulness of the CFQ by considering clinically important relationships via analyses of construct validity (related to both ACT process constructs and outcome measures), concurrent criterion validity (related to symptom measures), and incremental validity (to a range of relevant measures) with a number of commonly used clinical instruments that had not yet been examined with the CFQ.

Hypotheses

For convenience, each hypothesis is followed by a note about any similar/relevant information examined primarily in the original series of development studies (i.e., Gillanders et al., 2014), with other citations noted:

1. The CFQ will have strong internal consistency reliability in the sample.
 - a. Multiple studies have reported strong alpha values, including the initial development samples ($0.88 < \alpha < 0.93$ across five adult U.K. samples), a sample of U.K. adults with psychosis ($\alpha=0.90$; Johns et al., 2016), a sample of U.K. psychology undergraduates ($\alpha=0.93$; Marshall & Brockman, 2016), and two samples of U.S. adults ($\alpha=0.95$ in both samples; Fergus, 2015)
2. The CFQ will show a single factor representing a unidimensional construct.
 - a. The original development showed a single factor based on six U.K. samples of adults.

3. The CFQ and AAQ-II will be significantly, positively correlated, evidencing appropriate construct validity based on the ACT model.
 - a. The original development showed consistently significant correlations ($0.72 < r < 0.87$) across four adult U.K. samples.
4. The CFQ and AAQ-II items will load onto two separate factors according to their scales of origin.
 - a. Refer to discussion above
5. The CFQ and Behavioral Activation for Depression Scale-Short Form (BADS-SF; Manos, Kanter, & Luo, 2011) will be significantly, negatively correlated, evidencing appropriate construct validity based on the ACT model.
 - a. The original development showed significant correlation ($r = -0.21$, $p=0.03$) with the Valued Living Questionnaire in one adult U.K. sample.
6. The CFQ and Depression Anxiety Stress Scales 21 (DASS21; Lovibond & Lovibond, 1995) depression, anxiety, and stress subscale scores, as well as total score, will be significantly, positively correlated, (separately) evidencing construct validity.
 - a. The original development showed significant correlations ($0.45 < r < 0.85$) with multiple different anxiety, depression, and general symptom measures across different adult U.K. samples.
7. The CFQ and Schwartz Outcome Scale-10 (SOS-10; Blais et al., 1999) total score will be significantly, negatively correlated, evidencing construct validity.

- a. The original development showed significant correlations ($-0.39 < r < -0.45$) with multiple different quality of life/life satisfaction measures across different adult U.K. samples.
8. The CFQ will show significant differences between groups with mild versus severe reported symptom levels on the DASS21 depression, anxiety, and stress subscale scores, as well as total score, (separately) evidencing concurrent criterion validity.
 - a. Concurrent criterion validity was not examined in the development studies or any other published, peer-reviewed literature for concurrent criterion validity; this is a unique contribution of the present study.
9. The CFQ will show significant differences between groups with mild versus severe reported symptom levels on the Anxiety Sensitivity Index -3 (ASI-3; Taylor et al., 2007) physical, cognitive, and social subscale scores, as well as total score, (separately) evidencing concurrent criterion validity.
 - a. Concurrent criterion validity was not examined in the development studies or any other published, peer-reviewed literature for concurrent criterion validity; this is a unique contribution of the present study.
10. The CFQ will significantly, incrementally predict the SOS-10 beyond the AAQ-II, evidencing relevant incremental validity.
 - a. The CFQ significantly predicted scores on the General Health Questionnaire 12 (GHQ-12) incrementally beyond the AAQ-II in one adult U.K. sample.

11. The CFQ will significantly, incrementally predict the DASS21 total score beyond the AAQ-II, evidencing relevant incremental validity.
 - a. Same as #10 above.
12. The CFQ will significantly, incrementally predict the DASS21 depression subscale score beyond the BADS-SF, evidencing relevant incremental validity.
 - a. Neither the BADS-SF/DASS21 nor any similar measure combination was examined in the development studies or any other published, peer-reviewed literature for incremental validity.
13. The CFQ will significantly, incrementally predict the DASS21 anxiety subscale score beyond the ASI-3, evidencing relevant incremental validity.
 - a. Neither the ASI-3/DASS21 nor any similar measure combination was examined in the development studies or any other published, peer-reviewed literature for incremental validity.

Method

Participants

The sample analyzed was ($n = 1717$) undergraduate students from Eastern Michigan University in the U.S. (after statistical exclusions from a starting sample of $n = 1782$; see details in the “Supporting Analyses” section below). The age of the sample ranged from 18 to 56 years old ($M = 20.9$, $SD = 4.2$). The sample was 69.7% female, 29.4% male, and 1.0% other gender. The sample was predominantly White (63.1%), Black or African American (19.6%), and multiracial/multiethnic (9.4%); see Table 2 for full racial/ethnic demographics of the sample according to the current U.S. Census categories.

Table 2

Racial/Ethnic Sample Demographics

	Frequency	%
White	1083	63.1
Black or African American	337	19.6
Multiracial/Multiethnic	161	9.4
Mexican, Mexican American, or Chicano	22	1.3
Puerto Rican	1	0.1
Cuban	1	0.1
Other Hispanic, Latino, or Spanish	18	1.0
Asian Indian	15	0.9
Chinese	15	0.9
Filipino	3	0.2
Japanese	2	0.1
Korean	17	1.0
Vietnamese	1	0.1
Other Asian	8	0.5
American Indian or Alaskan Native	2	0.1
Native Hawaiian	1	0.1
Samoan	1	0.1
Other Pacific Islander	1	0.1
Other Race/Ethnicity	28	1.6

A priori power analysis. Unfortunately, there appears to be little agreement within the statistical community about the necessary sample size to perform properly powered EFA. Many follow general guidelines independent of investigated conditions, but some argue any rule of thumb does not account for differences across models, estimation methods, and other parameters of investigation—not to mention relationships between variables and methods for handling missing data—all of which can impact

necessary sample size for power (e.g., Schmitt, 2011). However, the same statisticians note that there are at least three different methods proposed within the statistics community that more robustly capture necessary power based on idiographic factors of any proposed investigation, and that all three also suffer from criticisms. These criticisms include needing information about the sample to be analyzed *prior to* data gathering and dependence on consensus in other debates around key factors such as fit indices where *no* such consensus exists. Fortunately, consensus only appears necessary when the goal is to minimize the necessary sample size while ensuring proper power. In the case of the present study, the sample size was known and large, and far exceeded even the most conservative guidelines for necessary sample size. For reference, various general guidelines range from recommendations of 2 to 10 participants per item, to 10 or more participants per expected factor, to 20 to 300 subjects total (for a review, see Sapanas & Zeller, 2002). The largest EFA of the present study included 14 items and was expected to predict two factors, resulting in the present sample having five to ten times the necessary size compared to all referenced guidelines.

Similarly, the sample appeared sufficient for all planned analyses with values of $\alpha = 0.05$ and $\beta = 0.8$, with one-tailed analyses based on hypothesized relationships between variables. Based on these assumptions and even assuming only small effect sizes, independent *t*-tests used to support criterion concurrent validity required $n = 698$ (even assuming a 2:1 ratio in the mild versus severe groups), multiple linear regression used to support incremental validity only required $n = 395$, and all correlational analyses required $n = 614$ (Faul, Erdfelder, Lang, & Buchner, 2007). Far smaller sample sizes would be necessary to detect the moderate or larger effect sizes expected.

Procedure

Data for the present study were gathered in an online survey as the screening phase for an intervention study to help manage anxiety and stress (see Appendix A for the original ethics review board approval letter). Participants were recruited via EMU's psychology research sign-up system portal (Sona-Systems.com) to the online survey on the study's specific web address at Qualtrics.com. All students can have access to the research sign-up system, though the majority with access are those taking psychology courses. The survey consisted of (in the following order): an informed consent form (see Appendix B), demographics questionnaire (see Appendix C), AAQ-II (Bond et al., 2011; see Appendix D), ASI-3 (see Appendix E), CFQ (Gillanders et al., 2014; see Appendix F), BADS-SF (Manos et al., 2011; see Appendix G), SOS-10 (Blais et al., 1999), and the DASS21 (Lovibond & Lovibond, 1995; see Appendix H). The research portal was also used to grant extra credit for participants taking part in the survey based on individual course instructor approval upon completion of the measures.

Measures

Cognitive Fusion Questionnaire (CFQ). The CFQ is a brief self-report survey designed to measure client fusion (Gillanders et al., 2014). The CFQ consists of seven items about respondents' cognitive fusion rated from 1 (*never true*) to 7 (*almost always true*), with no reverse scored items. The seven items are summed to a total score (7 to 49), with higher scores indicating more cognitive fusion.

Development of the CFQ included the use of seven varying and sometimes mixed samples of adults from the U.K. that included students, dementia caregivers, individuals with multiple sclerosis, prison service workers, and individuals with depression. These

studies showed that it loaded onto a single factor, was sensitive to the effects of an ACT intervention targeting cognitive fusion across three time points, and had acceptable psychometric properties. Internal consistency reliability was reported for five of the development samples ($0.88 < \alpha < 0.93$). Construct validity was confirmed across a number of measures in multiple domains including those assessing the ACT model, related constructs, important outcomes, and broader outcomes like quality of life. Divergent validity was confirmed by comparing the CFQ to a measure of socially desirable responding. Incremental validity was confirmed above other appropriate measures in the prediction of depressed mood and two different measures of distress. Criterion validity was confirmed by comparing distressed and non-distressed samples. Test-retest reliability was confirmed in a small subsample ($n = 82$) and showed good temporal stability across a 4-week period.

Only a few other studies have also reported internal consistency reliability values, ranging from $0.90 < \alpha < 0.95$ across four samples, including two of U.S. adults, one of U.K. adults diagnosed with various psychotic disorders, and one of U.K. undergraduates (Fergus, 2015; Johns et al., 2016; Marshall & Brockman, 2016). Some psychometric properties were also reported for translated versions of the CFQ in three studies. These included a Korean translation (Kim & Cho, 2015), and a Spanish (Catalan) translation (an adult sample: Romero-Moreno, Márquez-González, Losada, Gillanders, & Fernández-Fernández, 2014; and an adolescent sample: Solé et al., 2015).

Acceptance and Action Questionnaire-II (AAQ-II). The AAQ-II is a measure of the core element of ACT's theoretical model, psychological flexibility (Bond et al., 2011). The AAQ-II is a 7-item self-report questionnaire with items rated from 1 (*never*

true) to 7 (*always true*), with no reverse scored items. The sum of all items (totaling 7 to 49) represents a single factor measuring psychological flexibility, with higher scores indicating less flexibility. Psychological flexibility is defined as “the ability to fully contact the present moment and the thoughts and feelings it contains without needless defense, and, depending upon what the situation affords, persisting in or changing behavior in the pursuit of goals and values” (Bond et al., 2011, p. 678).

The AAQ-II was developed to improve upon the original’s comprehension, reliability, and factor structure stability. The original development showed an improved factor structure evidenced by a robust mean value of internal consistency reliability of $\alpha = 0.84$ —the reliability values varied minimally from 0.78 to 0.88 across four samples: two of U.S. university students, one of U.S. substance misusers receiving treatment, and one of U.K. financial services workers (Bond et al., 2011). The original studies of the AAQ-II also found good test-retest reliability for both a 3-month (0.81) and 12-month (0.79) test-retest interval and—using established measures for comparison—supported concurrent and predictive validity with expected outcomes (e.g., psychological distress), convergent validity with similar constructs (e.g., thought suppression), and discriminant validity with dissimilar constructs (e.g., social desirability).

Behavioral Activation for Depression Scale-Short Form (BADS-SF). The BADS-SF is a brief version of the original designed to measure client activation (Manos et al., 2011). The BADS-SF consists of nine items about respondents’ activation and avoidance experiences in the past week rated from 0 (*not at all*) to 6 (*completely*). The scale includes some reverse score items and produces two subscale scores: activation (six items) and avoidance (three items) that sum to a total score (0 to 54), with higher scores

indicating more activation and less avoidance. Development research of the BADS-SF has shown it to have acceptable psychometric properties, including internal consistency reliability ($\alpha = 0.82$), construct validity, test-retest reliability, predictive validity, and the ability to track changes over the course of treatment.

Schwartz Outcome Scale-10 (SOS-10). The SOS-10 is a self-report survey of quality of life and psychological well-being (Blais et al., 1999). The SOS-10 consists of ten items about how the respondent feels over the past seven days that are rated on a 7-point scale from 0 (*never*) to 6 (*all of the time or nearly all of the time*) with no reverse scored items. The total score (from 0 to 60) was structured to represent psychological health and well-being—with higher scores indicating greater psychological health—and was designed to be used in measuring the effectiveness of mental health treatments broadly across the mental health population.

The initial validation study of the SOS-10 showed a high internal consistency reliability ($\alpha = 0.96$) and sensitivity to change based on intervention in a mixed inpatient/outpatient sample (Blais et al., 1999). Subsequent studies have independently shown strong test-retest reliability (0.86) in a sample of undergraduate students (Young, Waehler, Laux, McDaniel, & Hilsenroth, 2003) and independently confirmed the internal consistency reliability ($\alpha = 0.96$) in inpatient, outpatient, and non-patient samples (Hilsenroth, Ackerman, & Blagys, 2001) as well as sensitivity to intervention-induced change with an outpatient sample (Young et al., 2003).

Depression Anxiety Stress Scales 21 (DASS21). The DASS21 is a set of self-report scales that are designed to measure the negative emotional states of depression, anxiety, and stress (Lovibond & Lovibond, 1995). The DASS21 consists of 21 questions

divided into three groups of seven, each measuring one of three emotional states, with items responses based on the respondent's experience over the past week rated from 0 (*did not apply to me at all*) to 3 (*applied to me very much, or most of the time*), with higher scores indicating greater levels of distress. Scores on individual items are summed for each subscale (totaling 0 to 21) and for the full scale total (totaling 0 to 63).

The DASS21 has shown solid internal consistency reliability across the depression ($\alpha = 0.94$), anxiety ($\alpha = 0.87$), and stress ($\alpha = 0.91$) scales for a mixed clinical and community sample (Antony, Bieling, Cox, Enns, & Swinson, 1998). Independent studies have confirmed internal consistency reliability in non-clinical samples for the total ($\alpha = 0.93$), depression ($\alpha = 0.88$), anxiety ($\alpha = 0.82$), and stress ($\alpha = 0.90$) scales (Henry & Crawford, 2005). Additional studies have shown the DASS21 to be a valid measure for use in measuring the clinical status and change in clinical populations (Ng, Trauer, Dodd, Callaly, Campbell, & Berk, 2007).

Anxiety Sensitivity Index-3 (ASI-3). The ASI-3 is an 18-item self-report scale designed to measure the respondent's fear arising from beliefs that anxiety-related sensations will have negative consequences like death, insanity, or social rejection (Taylor et al., 2007). Participants respond to items on a 5-point scale from 0 (*very little*) to 4 (*very much*), with no reverse scored items. It is composed of three 6-item subscales: physical concerns, cognitive concerns, and social concerns—with a higher score on a given scale indicating a greater concern of that type. Scores on the physical concerns subscale are associated with expectations that heart/chest, throat, or stomach sensations will lead to serious illness or death. The cognitive concerns subscale assesses the extent to which the participant believes difficulties in thinking or concentrating will become

mental illness or insanity. The social concerns subscale indicates the extent to which the individual thinks observable anxiety reactions, such as trembling, blushing, or appearing nervous, will elicit social ridicule or rejection. Depending on the application, the total score of the three subscales combined represents an overall level of AS.

Internal consistency reliability of the ASI-3 subscales showed robust alpha values, with ranges of $\alpha = .76$ to $.86$, $.79$ to $.91$, and $.73$ to $.86$, for the physical, cognitive and social concerns subscales, respectively (Taylor et al., 2007). Many studies utilizing the ASI-3 since its creation demonstrate strong internal consistency reliability: for instance, McDermott, Tull, Gratz, Daughters, and Lejuez (2009) obtained subscale values of $\alpha = .93$, $.91$, and $.86$ (for physical, cognitive and social subscales, respectively) for crack/cocaine users with PTSD. The creators of the ASI-3 did not, however, confirm its test-retest reliability—the only available study to date confirming this showed a satisfactory value of $r = .64$ with a Turkish sample utilizing the Turkish translation of the ASI-3 (Mantar, Yemez, Alkin, & Dergisi, 2010). These Turkish researchers also confirmed the internal consistency of the ASI-3 with the Turkish sample, finding an overall alpha value of $\alpha = .93$; with alpha values of $\alpha = .89$, $\alpha = .88$, and $\alpha = .82$ for the physical, cognitive, and social subscales, respectively.

Results

Supporting Analyses

Prior to examining each of the primary analyses, the data set was investigated for biased responding and then analyzed to confirm the appropriateness of the data for each planned analysis. Due to the procedure for gathering data, there were no missing data.

All applicable analyses used a one-tailed $\alpha = .05$ and determined statistical significance using the guideline of $p < .05$.

Detection and management of response bias. Response bias in self-report measurement has been a concern in psychology for almost a century. Although it is often significantly overestimated according to the empirical literature on the subject (for a discussion, see McGrath, Mitchell, Kim, & Hough, 2010), there is some literature supporting elevated prevalence of biased responding in a few situations. One relevant type is so-called *careless responding* in situations where anonymity is facilitated and/or when there are secondary incentives to participate (for a brief review, see Godinho, Kushnir, & Cunningham, 2016). The present study was at risk for this form of responding, given that it was web-based and included undergraduate students, some of whom were compensated with extra credit for completing the survey.

The various approaches to detecting and deterring response bias can be broadly categorized as (a) within measure efforts or (b) post-hoc screening (Meade & Craig, 2012). Because the present study utilized existing measures, within measure methods—such as adding duplicative or “bogus” items as checks—were not available options to investigate potential response bias. A number of post-hoc screening methods, however, were possible in the present study.

Overall, there is no clear consensus on how to handle response bias in general or careless responding in particular—likely due to the paucity of rigorously designed studies (McGrath et al., 2010). Disagreement notwithstanding, there has recently been some initial concordance on recommendations that state robust post-hoc investigations should consider participants’ response times and patterns—including *intra*-individual

consistency—as well as *inter*-individual responding via examination of outliers (e.g., Huang, Curran, Keeney, Poposki, & DeShon, 2012; Meade & Craig, 2012). The most supported specific methods for each include using Mahalanobis distance (Mahalanobis, 1936) to detect individuals who differ problematically in their responding (i.e., outliers in a broader sense of response pattern); setting a minimum response time (though this is done arbitrarily rather than empirically, unfortunately) to consider overall lack of response effort; examining maximum long string (the highest number of identical consecutive responses) to detect *overly* consistent responding; and an even-odd consistency index to detect *intra*-individual problems of consistency (most commonly attributed to inattention while taking the survey).

The even-odd consistency index approach requires that the data being analyzed consist of a significant number of subscales, each needing a substantial number of items, in order for the resulting correlations to be statistically meaningful (i.e., have small enough confidence intervals; Niessen, Meijer, & Tendeiro, 2016). This is also true of related approaches such as psychometric antonyms. The measures utilized in the present study are generally brief (by design) and far below the reference studies, which both primarily used a 300-item questionnaire with high item-count subscales (Huang et al., 2012; Meade & Craig, 2012).

Similarly, the maximum long string method also requires moderate or longer length measures in order to properly differentiate conscientious from careless responders (Niessen et al., 2016). Although the choice of brief measures is a barrier to such analyses, brevity of surveys also reduces the risk of this type of responding by preventing problems such as inattentiveness by reducing the possibility of response fatigue.

Alternative methods for examining response patterns—most notably person-fit statistics, based on item response theory—are less often used due to multiple drawbacks: they are more restrictive, often require response options be collapsed resulting in loss of data precision, and simply are not able to assess *intra*-individual response patterns as they are relative analyses by nature (Johnson, 2005). As such, no *intra*-individual consistency checks were performed for the present study.

Regardless of other restrictions based upon the data, careless responding could still be detected in the present study via analysis of response time. Response time has been shown to have the highest sensitivity of all careless responding detection methods when an appropriate cutoff is determined (Niessen et al., 2016) and can arguably be used to detect multiple forms of careless responding, including the focal form of the long string method, overly consistent responding.

As such—and because response time is easily captured in an online survey—the present study examined the response times of participants using a previously recommended minimum cutoff of 2 seconds per item (Huang et al., 2012), or 160 seconds for the entire survey. Participants with survey response times shorter than the 160 second cutoff were excluded from consequent analyses. This cutoff was set conservatively in that the survey also included elements not counted relative to the 2 seconds per item guideline, including a multi-page informed consent form and a few other brief passages of text.

Two additional considerations were made to avoid eliminating data based on false positive identification given the arbitrary nature of the cutoff. First, the descriptive statistics of response time in the sample were subjectively examined to help ensure the

recommended cutoff seemed appropriate for the data. Second, the resulting number of excluded cases were compared to the previously reported rates in the existing literature with a similar intent.

The response time investigation revealed 43 participants with a total time to take the survey that was less than the 160 second cutoff. The modal response time of the full sample was 340 seconds, and the median was 501 seconds (data were positively skewed by some extreme outliers, resulting in a less interpretable mean; $M = 2325$ seconds). This was judged to indicate the subsample below the 160 second cutoff was likely careless responders. The false positive rate was deemed low based on the fraction falling into this category (2.4%) compared to the reported rates in the literature of 3.5% to 12% (e.g., Johnson, 2005; Kurtz & Parish, 2001; Meade & Craig, 2012). As such, the 43 cases were removed prior to running any other analyses, leaving a remaining sample of $n = 1739$.

Due to their nature as outlier calculations, Mahalanobis distances require that the data being considered are normally distributed, which is often a concern with psychological survey data (Niessen et al., 2016). To confirm this assumption, tests of normality were performed at the item level for each measure using the Shapiro-Wilk's test (Shapiro & Wilk, 1965) as it perhaps the most common approach—in part due to its increased power relative to alternative tests (Stephens, 1974).

All 72 measure items showed significant results, indicating non-normality. However, real data is, in fact *never* normal in the sense that perfect normality is an asymptotic proposition (Micceri, 1989). That is to say, the real question is not binary in nature; instead, it is whether the data deviate *too* much from the normal distribution to substantially influence the veracity of any conclusions drawn—a process that includes

multiple arbitrarily-determined cutoffs (e.g., the generally accepted 5% chance of Type I errors on statistical significance).

Statistical philosophy aside, the Shapiro-Wilk's test is known to often exhibit significance in large sample sizes due to only minor deviations from normality in the data that are very unlikely to affect conclusions drawn from consequent analyses that assume normality of data (Field, 2005). As such, in instances of significant Shapiro-Wilk's test results with large samples, it is recommended that additional metrics be considered, including the skewness and kurtosis values and the character of the Quantile-Quantile (Q-Q) plots. Skewness or kurtosis absolute values larger than generally accepted guidelines of 1.96 should be taken as confirmatory evidence of non-normality. Q-Q plot examination is subjective by nature but can also be used to corroborate normality conclusions based on visually identified deviation from the normal baseline.

At the item level, the highest absolute skewness value was 1.58, and the highest absolute kurtosis value was 1.56. Additionally, no Q-Q plots showed apparent, significant deviation from the normal trend. As such, all item-level data were considered to exhibit *sufficient* normality in order to be considered robust for the calculation of valid Mahalanobis distances to detect outliers in response patterns, without resorting to data transformation procedures.

Mahalanobis distances were calculated separately for all measures, for each participant. Probability values for all calculated distances were then computed via comparison to the chi-square distribution, with critical p -values ($p < .001$) indicating likely careless responding. The purpose of this check process was to identify participants with consistently errant response patterns. Because of this, and in order to avoid high

rates of false positive errors, only participants with data exhibiting critical p values on three or more measures (i.e., a majority) were excluded from consequent analyses.

Although a number of participants evidenced critical p -values on at least one measure ($n = 278$), only 22 showed response patterns of concern on three or more measures compared to the sample. Data for these 22 participants were removed, leaving a sample of $n = 1717$ for planned analyses.

Confirmation of assumptions of primary analytic approaches. A few assumptions were particular to specific planned analyses and are therefore reported below, immediately preceding the results of that particular analysis. Other assumptions are shared across multiple analyses and are therefore reported here based on more omnibus confirmatory checks across measures.

Normality is not required (though is preferable) for the chosen method of EFA, and it *is* an assumption of all other planned analyses. Multivariate normality was assessed across total scores for all scales and subscales in the same manner described above for the item level confirmations related to Mahalanobis distances. All 14 scales/subscales showed significant results on the Shapiro-Wilk's test; again, this is likely an artifact of the very large sample size. However, the highest absolute skewness value was 0.93 and the highest absolute kurtosis value was 0.95. Additionally, no Q-Q plots showed significant deviation from the expected normal trend. As such, all scale/subscale data were considered to exhibit sufficient normality in order to be considered robust for all relevant planned analyses, without resorting to data transformation procedures.

Linearity is an assumption for the correlation and regression analyses. As there is no robust/objective statistical check for linearity, it was confirmed by subjectively

examining plots of residuals for scales/subscale total scores. Plots of the CFQ against all 14 scales/subscales comparing standardized residuals against predicted values showed little evidence of non-linearity and were all deemed appropriate for all relevant planned analyses.

Homoscedasticity (also referred to as homogeneity of variance) is required for all planned analyses except EFA and was confirmed using Levene's test for all scales/subscales, as it is less sensitive to departures from normality than alternative tests of homogeneity of variance (Levene, 1960). All scales/subscales had significant Levene's test results with the CFQ, with the exceptions of the BADS-SF Avoidance subscale score ($p = .19$) and the SOS-10 total score ($p = .29$), indicating possible heterogeneity of variance. However, similar to statistical tests of normality mentioned above, Levene's test is sensitive to small differences in variance as samples analyzed become large. As such, plots of data residuals were examined for subjective corroboration of heteroscedasticity character. Plots of the CFQ with all 14 scales/subscales comparing standardized residuals against predicted values showed little evidence of heteroscedasticity. Additionally, the chance for Type I error rates in such cases are significantly reduced by equally-sized and sufficiently large samples (Coombs, Algina, & Oltman, 1996), as was the case in the present study. As a result, data were deemed appropriate for all relevant planned analyses.

With the exception of regression, outliers are problematic for all planned analyses and were investigated for all scale/subscale data using the outlier labeling rule (Hoaglin, Iglewicz, & Tukey, 1986). This approach is arguably more robust than the more commonly used, simple guideline based on standard deviation (Hoaglin & Iglewicz,

1987). All scale/subscale data were investigated based on a multiplier coefficient of $g = 2.2$ based on the latest recommendations. Outlier investigations revealed no outliers for any of the 14 scales/subscales.

Perfect multicollinearity must not exist for the data used in the multiple linear regression analyses and is also a concern for EFA. This characteristic was investigated using the tolerance parameter according to modern guidelines (O'Brien, 2007).

Correlations between scales of interest were not high enough to imply concerns about perfect multicollinearity, as evidenced by tolerance values between the CFQ and the AAQ-II (.28), BADS-SF (.75), and ASI-3 (.42) all being above even the most conservative minimum guideline (tolerance $> .2$).

Primary Analyses

After calculations confirming the general robustness and appropriateness of the data, statistical analyses were performed to test each of the present study's hypotheses. As above for the supporting analyses, all applicable analyses used one-tailed $\alpha = .05$ and determined statistical significance using the guideline of $p < .05$.

Internal consistency reliability. To confirm the CFQ's internal consistency reliability (Hypothesis #1), an alpha coefficient was calculated using correlations between responses to the seven items on the CFQ. As predicted, the CFQ showed high internal consistency reliability, with a value of $\alpha = .96$ for the analyzed sample. The internal consistency reliability for all scales/subscales, as well as their mean scores and standard deviations, can be found in Table 3.

Table 3

Scale/Subscale Statistics

Scale/Subscale	<i>M</i>	<i>SD</i>	α
CFQ	23.9	11.8	.96
AAQ-II	21.0	10.2	.93
BADS-SF	29.0	7.9	.65
BADS-SF Activation subscale	20.4	7.1	.80
BADS-SF Avoidance subscale	8.6	2.6	.79
SOS-10	39.8	13.1	.94
DASS21	17.0	13.3	.95
DASS21 Depression subscale	5.2	4.9	.90
DASS21 Anxiety subscale	5.0	4.7	.87
DASS21 Stress subscale	6.8	4.9	.87
ASI-3	22.9	16.7	.95
ASI-3 Physical subscale	6.4	6.0	.90
ASI-3 Social subscale	10.0	6.2	.85
ASI-3 Cognitive subscale	6.5	6.3	.92

Factor structure of the CFQ. Prior to EFA, shared variance was investigated via the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy (Kaiser & Rice, 1974) to confirm the proportion of common variance among items. The resulting value of .95 is greater than the generally accepted minimum of .60 (Tabachnick & Fidell, 2001), so the sampling adequacy of the data was deemed sufficient. Bartlett's test of sphericity (Bartlett, 1954) was also performed to confirm the inter-item correlation matrix of the CFQ has non-zero covariance in the analyzed sample. The test was significant at the $p < .001$ level, indicating the data are acceptable for factor analysis.

Once the data were deemed appropriate, EFA was performed with the seven items from the CFQ to confirm its factor structure (Hypothesis #2). The EFA was run using the same methodology as the original development (Gillanders et al., 2014). This included a

principal axis factoring (also known as common or principal factor analysis) approach with parallel analysis for factor extraction and an oblique promax rotation to identify the latent construct(s). This approach allowed for a direct confirmation of the original development results. Additionally, the EFA parameter choices made were arguably among the stronger recommended approaches within the behavioral sciences statistics literature (for a discussion, see Fabrigar, Wegener, MacCallum, & Strahan, 1999).

The parallel analysis revealed a single factor with an eigenvalue of 5.62, with the next largest factor exhibiting an eigenvalue of only .33. This result did not vary between two parallel analyses: one run using normally distributed randomly generated data (the more common method), and another using permutations of the data set from the present study (a superior method according to some recommendations; e.g., O'Connor, 2000).

The initial factor analysis was run without restriction to corroborate the parallel analysis and revealed the same values of 5.62 for the first factor's eigenvalue (and .33 for the next largest factor), indicating a robust single-factor solution that explained 77.1% of the total variance. The most commonly implemented guidelines for EFA recommend consideration of any factor with an eigenvalue above 1, and omission of values less than .4 (Ferguson & Cox, 1993). Values between those cutoffs can be considered more carefully via subjective examination of a scree plot based on its point(s) of inflection. Despite the lack of eigenvalues in that range for the solution of the CFQ EFA, the scree plot was examined and corroborated the statistical result with a clear, single point of inflection. Because only one factor was extracted, the solution could not, by definition, be rotated. This solution confirmed the predicted factor structure of the CFQ. Factor loadings for individual measure items are shown in Table 4.

Table 4

Factor Loadings for the EFA of the CFQ

Measure Item	<i>Factor 1 Loading</i>
CFQ 4	.920
CFQ 6	.895
CFQ 5	.885
CFQ 1	.878
CFQ 7	.860
CFQ 2	.858
CFQ 3	.849

Relationship with the AAQ-II. To first confirm the construct validity of the CFQ relative to the AAQ-II (Hypothesis #3), the bivariate correlation between the two scales was examined. As predicted, it was found to be high: $r = .85, p < .001$.

Next, the EFA procedure described above was repeated after adding the AAQ-II's seven items to the CFQ's seven items to consider the factor structure relationship of the combined pool of measure items (Hypothesis #4). First, the KMO (.97) and Bartlett's test of sphericity (significant at the $p < .001$ level) confirmed the appropriateness of the combined item set for factor analysis. Next, the parallel analysis revealed a single factor with an eigenvalue of 9.85, with the next largest factor exhibiting an eigenvalue of only .95. This result again did not vary between the two methods of parallel analysis.

The initial factor analysis was again run without restriction to corroborate the parallel analysis and revealed the same values of 9.85 for the first factor's eigenvalue (and .95 for the next largest factor), indicating a robust single-factor solution that

explained 68.1% of the total variance. There were three additional factors above the .4 eigenvalue omission threshold (.95, .52, .46); however, no individual items from the CFQ or AAQ-II had factor loadings that were higher on any factor other than the first. Additionally, the scree plot showed a single, distinct point of inflection, corroborating the single factor statistical solution. This solution does not support the hypothesized prediction that the items from the CFQ and AAQ-II would load onto two separate factors. Instead, they exhibited a significant statistical overlap in this sample. Factor loadings for individual measure items are shown in Table 5.

Table 5

Factor Loadings for the EFA of the CFQ and AAQ-II Combined

Measure Item	<i>Factor 1 Loading</i>
CFQ 4	.898
CFQ 1	.879
CFQ 5	.863
CFQ 2	.853
CFQ 6	.851
AAQ-II 5	.848
AAQ-II 3	.838
CFQ 7	.837
AAQ-II 7	.818
CFQ 3	.814
AAQ-II 6	.801
AAQ-II 2	.782
AAQ-II 4	.727
AAQ-II 1	.727

Construct validity. To further confirm the CFQ's construct validity, bivariate scale/subscale correlations were calculated between the CFQ and the BADS-SF (Hypothesis #5), the DASS21 total scale, as well as the individual subscales of depression, anxiety, and stress separately (Hypothesis #6), and the SOS-10 (Hypothesis #7). All three predictions were supported with significant correlations in the direction predicted for the BADS-SF ($r = -.50, p < .001$), DASS21 total score ($r = .75, p < .001$), DASS21 depression subscale score ($r = .68, p < .001$), DASS21 anxiety subscale score ($r = .63, p < .001$), DASS21 stress subscale score ($r = .74, p < .001$), and SOS ($r = -.57, p < .001$). All correlations related to hypotheses, as well as all other correlational relationships between scales/subscales, can be seen in Table 6.

Table 6

Bivariate Correlation Matrix of All Scales/Subscales

	AAQ-II	ASI-3 Cognitive	ASI-3 Physical	ASI-3 Social	ASI-3	CFQ	BADS-SF Activation	BADS-SF Avoidance	BADS-SF	SOS-10	DASS21	DASS21 Stress	DASS21 Anxiety	DASS21 Depression
AAQ-II	1	.70	.59	.66	.73	.85	-.40	-.47	-.51	-.59	.75	.72	.63	.71
ASI-3 C		1	.76	.69	.91	.72	-.37	-.36	-.45	-.52	.73	.67	.68	.66
ASI-3 P			1	.66	.90	.61	-.30	-.30	-.37	-.42	.65	.58	.66	.55
ASI-3 S				1	.88	.70	-.32	-.39	-.41	-.47	.65	.63	.60	.56
ASI-3					1	.76	-.37	-.39	-.46	-.53	.76	.70	.72	.66
CFQ						1	-.37	-.53	-.50	-.57	.75	.74	.63	.68
BADS-SF Ac							1	.17	-.95	.63	.43	-.38	-.34	-.47
BADS-SF Av								1	-.47	.29	.46	-.48	-.40	-.41
BADS-SF									1	.65	.54	-.49	-.43	-.55
SOS-10										1	.60	-.53	-.50	-.62
DASS21											1	.93	.91	.91
DASS21 S												1	.80	.78
DASS21 A													1	.73
DASS21 D														1

Note: All correlations were significant at the $p < .001$ level (1-tailed)

Concurrent validity. To confirm the CFQ's concurrent validity, participants were first dichotomized into normative and severe groups on the DASS21 (Hypothesis #8) and ASI-3 (Hypothesis #9). Since the relevant measures lack empirically supported clinical cutoffs, a relative cutoff of one standard deviation above the mean of the sample was used for all group comparison analyses. Cutoffs were determined based on the statistical data shown above in Table 3 for the DASS21 total scale (31), DASS21 depression subscale (11), DASS21 anxiety subscale (10), DASS21 stress subscale (12), ASI-3 total scale (40), ASI-3 physical subscale (13), ASI-3 social subscale (17), and ASI-3 cognitive subscale (13) for the creation of eight separate dichotomies.

The normative group was then compared to the respective severe group using a separate independent *t*-test for each pair to detect if significant differences existed between the dichotomized groups on CFQ scores. Because equal variances cannot be assumed based on the previous results of heteroscedasticity, the standard Student's *t*-test approach to pooling the error terms would have been inappropriate. As such, all analyses employed the Welch-Satterwaite method (Welch, 1947) for adjusting the degrees of freedom in the analyses to account for the heteroscedasticity (otherwise known as the Welch's *t*-test or unequal variances *t*-test).

Analyses for all eight scales/subscales showed significant differences on two-tailed independent *t*-tests, with the severe groups scoring significantly higher on the CFQ than the normative groups. These results support both hypotheses' predictions and evidence concurrent validity between the CFQ and both the DASS21 and ASI-3. Table 7 shows the results and adjusted degrees of freedom for all eight analyses as well as the

CFQ means and standard deviations of each group within the pair for each of the eight analyses.

Table 7

Summary of Group Statistics and Comparison Analyses

Scale/Subscale	Normative		Severe		Comparison		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>g</i>
DASS21	21.02	10.32	37.04	8.48	-28.96	536.62	1.60
DASS21-D	21.38	10.64	36.26	8.72	-25.61	493.83	1.44
DASS21-A	21.32	10.60	35.31	9.60	-23.08	515.06	1.34
DASS21-S	21.01	10.37	36.74	8.50	-28.55	555.87	1.57
ASI-3	20.93	10.27	37.40	7.97	-31.22	568.38	1.67
ASI-3 P	21.58	10.86	34.96	9.31	-21.99	492.17	1.26
ASI-3 S	21.37	10.60	36.03	9.16	-24.47	484.27	1.41
ASI-3 C	20.75	10.15	37.42	7.92	-32.37	608.74	1.71

Note: Values are based on CFQ scores; All *t*-tests were significant at the $p < .001$ level (2-tailed)

Incremental validity. To confirm the CFQ's incremental validity, sets of measures corresponding to each hypothesis were entered into multiple linear regression analyses, with the CFQ being added on the second (and final) step using the enter method. Effect size was calculated using a local effect size variant of Cohen's f^2 (Selya, Rose, Dierker, Hedeker, & Mermelstein, 2012). Interpretation conventions for f^2 are small (.02), medium (.15), and large (.35; Cohen, 1988). The first regression investigated the CFQ's incremental validity over the AAQ-II in predicting the SOS-10 (Hypothesis #10). The analysis found that fusion predicted quality of life beyond psychological flexibility, with the full model explaining a significant amount of variance in quality of life: $F(2,1714) = 499.17, p < .001, R^2_{Adjusted} = .37$ (see Table 8 for additional statistics).

Table 8

Summary of Regression Analysis Predicting SOS-10 Scores

Variable	Model 1			Model 2		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
AAQ-II	-.764	.025	-.594	-.508	.047	-.395
CFQ				-.261	.041	-.234
R^2		.353			.368	
Adjusted R^2		.352			.367	
<i>F</i> for change in R^2		*934.9			*41.4	
f^2					.024	

Note: * $p < .001$; f^2 is the local effect size variant (not the effect size of the full predictor model)

The second regression investigated the CFQ's incremental validity over the AAQ-II in predicting the DASS21 (Hypothesis #11). The analysis found that fusion predicted symptoms of depression, anxiety, and stress beyond psychological flexibility, with the full model explaining a significant amount of variance in symptoms: $F(2,1714) = 1298.13, p < .001, R^2_{Adjusted} = .60$ (see Table 9 for additional statistics).

Table 9

Summary of Regression Analysis Predicting DASS21 Scores

Variable	Model 1			Model 2		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
AAQ-II	.977	.021	.746	.530	.038	.405
CFQ				.457	.033	.402
R^2		.557			.602	
Adjusted R^2		.557			.602	
<i>F</i> for change in R^2		*2157.9			*194.7	
f^2					.113	

Note: * $p < .001$; f^2 is the local effect size variant (not the effect size of the full predictor model)

The third regression investigated the CFQ's incremental validity over the BADS-SF in predicting the DASS21-Depression (Hypothesis #12). The analysis found that fusion predicted symptoms of depression beyond behavioral activation, with the full model explaining a significant amount of variance in depressive symptoms: $F(2,1714) = 943.20$, $p < .001$, $R^2_{Adjusted} = .52$ (see Table 10 for additional statistics).

Table 10

Summary of Regression Analysis Predicting DASS21 Depression Scores

Variable	Model 1			Model 2		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
BADS-SF	-.343	.013	-.550	-.172	.012	-.276
CFQ				.229	.008	.545
R^2		.303			.524	
Adjusted R^2		.302			.523	
F for change in R^2		*744.1			*796.9	
f^2					.464	

Note: * $p < .001$; f^2 is the local effect size variant (not the effect size of the full predictor model)

The fourth regression investigated the CFQ's incremental validity over the ASI-3 in predicting the DASS21-Anxiety (Hypothesis #13). The analysis found that fusion predicted symptoms of anxiety beyond anxiety sensitivity, with the full model explaining a significant amount of variance in anxious symptoms: $F(2,1714) = 985.00$, $p < .001$, $R^2_{Adjusted} = .53$ (see Table 11 for additional statistics).

Table 11

Summary of Regression Analysis Predicting DASS21 Anxiety Scores

Variable	Model 1			Model 2		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
ASI-3	.204	.005	.719	.160	.007	.563
CFQ				.083	.010	.206
R^2		.517			.535	
Adjusted R^2		.516			.534	
F for change in R^2		*1832.8			*66.8	
f^2					.039	

Note: * $p < .001$; f^2 is the local effect size variant (not the effect size of the full predictor model)

Discussion

Summary of Results

The present study supported all hypotheses with one exception. Analyses supported confirmation of strong internal consistency reliability and the unidimensional nature of the CFQ. Expected relationships between fusion (CFQ) and depression, anxiety, and stress symptoms (DASS21) as well as the constructs of anxiety sensitivity, including its physical, social, and cognitive dimensions (ASI-3); behavioral activation, including the activation and avoidance components (BADs-SF); and quality of life (SOS-10) were also supported by analyses in the present study. Fusion was also incrementally predictive of quality of life and symptoms of depression/anxiety/stress beyond psychological flexibility (AAQ-II), of symptoms of depression beyond behavioral activation, and of anxiety beyond anxiety sensitivity. However, the incremental validity of the CFQ beyond the AAQ-II must be interpreted with caution, as interpretation of those two investigations depends upon the interpretation of the factor analysis result (discussed below). Finally, analyses showed concurrent validity between a measure of

fusion, and measures of symptoms of depression/anxiety/stress and anxiety sensitivity respectively.

These findings are an independent replication of the important basic psychometrics of the CFQ as well as a confirmation and extension of similar results from the development studies by Gillanders and colleagues (2014) of the relationship between CFQ/fusion and other important measures/constructs. Further, the CFQ has now been psychometrically evaluated with 28 measures of over 30 distinct constructs, many of which are important and frequently utilized within the clinical literature. Replication allows for confident application of what appears to be a very robust measure of the unidimensional construct of fusion, based on behaviorally operationalized questions in a practical form.

Relationship between fusion and psychological flexibility. The exception to this affirmational trend in results was the factor analysis of the CFQ and AAQ-II combined. The resulting solution was a robust single factor, indicating both instruments were in fact measuring the same construct in the present study's sample. This result is the same as two of the five samples Gillanders and colleagues (2014) analyzed informally, only with a much larger sample; it stands in contrast with the other three samples they analyzed that showed a two factor solution based on the scales of origin—including the sample from the development studies' most demographically similar (Sample 1, a young adult sample).

Certainly, the high correlation between the two measures ($r = .85$; in the range of the development values of $.72 < r < .87$) is expected based on the ACT model, where fusion is conceptualized as a component of the overarching construct purportedly

measured by the AAQ-II: psychological flexibility. Similarly, the AAQ-II includes at least two items out of the seven that could easily be considered questions about fusion: “I worry about not being able to control my worries and feelings” and “Worries get in the way of my success.”

The factor analytic result, however, does not indicate simple relatedness, but rather that the AAQ-II and the CFQ are *so* interdependent that they are in fact measuring the same underlying construct in this sample. Gillanders and colleagues (2014) noted that the presumption, then, would be that they are measuring the same thing differently. That is, the CFQ may be measuring content of this shared construct that is specific to cognition, whereas the AAQ-II is measuring that same construct in a broader context that also includes emotions/feelings and memories/experiences. They also hinted that this explanation has additional support in their findings related to thought control (Thought Control Questionnaire) and rumination (Positive Beliefs About Rumination Scale [PBRs] and Ruminative Response Style Questionnaire [RSQ]), but did not fully explicate their understanding. This interpretation follows logically from the factor analytic result—that it, it is consistent with the statistical implications of the result. However, it is incongruent with the ACT model from a conceptual standpoint, where psychological flexibility is purported to include defusion *as well as* several other constructs/facets that, while related to defusion, are also distinct. Regardless, this interpretation can be readily evaluated empirically in a future study, as the CFQ should predict outcomes related to that construct and primarily involving cognition better than the AAQ-II, due to its greater specificity with that particular psychological context.

Related to the above interpretation, the derivation of the CFQ may have contributed to the issue. Gillanders and colleagues (2014) noted that its item development began broadly, defined by a larger pool of items incorporating several aspects of fusion (described earlier in this paper). However, during the empirical process of refining the measure, it may have become too narrowly defined. The result could be that the one or few aspects of fusion represented by the CFQ are highly interdependent with the construct of psychological flexibility (e.g., dominance of thoughts over behavior), whereas the absent aspects of fusion (e.g., evaluative extent of thoughts) are significantly less correlated with psychological flexibility. It could be that these aspects could comprise a measure more unique from psychological flexibility, whereas the included aspects are more interdependent aspects that need to be further investigated as noted above. Again, this interpretation has logical support from the item content and is consistent with the factor analytic result but also implies a result conceptually that is inconsistent with the ACT model—specifically that only part(s) of the defusion construct is/are subsumed under psychological flexibility, whereas other aspects of defusion are distinct.

The interpretation possibility perhaps most discussed in the literature is that the AAQ-II does not robustly measure what it is purported to measure. The most applicable instance of this argument may be the studies done by Gámez and colleagues (2011, 2014) developing full and brief measures of experiential avoidance, the foil of acceptance and another process included within psychological flexibility according to the ACT model. They highlighted the issue of the AAQ-II lacking sufficient discriminant validity with other constructs—in particular, neuroticism. It is worth noting that discriminant validity

was not robustly addressed by the development studies nor the present study, particularly in relation to other ACT constructs (e.g., experiential avoidance). Gámez and colleagues (2011, 2014) also made efforts in item selection to avoid conflation between experiential avoidance and other relevant constructs for the design of their measures. Overall, the data are mixed on the distinction between psychological flexibility and other important constructs (for a discussion, see Gloster, Klotsche, Chaker, Hummel, & Hoyer, 2011).

A more direct investigation of a related concern was performed by Wolgast (2014), who considered the possibility that the AAQ-II does not sufficiently distinguish between process and outcome. This concern has been raised by other researchers as well (e.g., Chawla & Ostafin, 2007). Wolgast (2014) found that the items from the AAQ-II loaded more significantly onto a factor that included items designed to measure general distress (i.e., outcomes) than onto another factor with items designed to measure acceptance as an attitude or response (i.e., process). This result risks circularity in measurement via the AAQ-II and could undermine its use in measuring, for example, process and outcomes of treatment studies. It may also be an explanation for the present study's results, in that fusion (as measured by the CFQ) may be a primary process that can/often results in the outcome of psychological inflexibility (as measured by the AAQ-II). It must be noted, however, that such an interpretation is incongruent with ACT's definition of psychological flexibility as a way of relating to our own experiences (i.e., a process, not an outcome).

There is a reasonable argument to be made that the ACT model does not appear to describe psychological flexibility as a unidimensional construct, but rather an "umbrella" construct inclusive of the six separate but related processes of defusion, acceptance,

mindfulness, self-as-context, values, and committed action. This conceptualization would prescribe replacing the AAQ-II (a unidimensional scale) with a measure that, for instance, could include one subscale for each process. Such a hierarchical relationship could be examined with established measures of each construct in the ACT model via a factor analytic approach to identifying a higher order construct (i.e., psychological flexibility) that encompasses each of the lower order constructs of the ACT model (for an example of the approach, see Bardeen, Fergus, & Orcutt, 2013). This approach is based on a paradigm that psychological flexibility behaves in a $1 + 1 = 2$ manner. The alternative is that the sum is greater than its parts, meaning that psychological flexibility is related to the six processes but is also different in some way at the structural level. It would be helpful if clearer distinctions were made between the processes that comprise psychological flexibility in the ACT model and the construct of psychological flexibility itself, such that these relationships could be tested empirically. Potentially because of current ambiguity, the literature presently has very few studies directly examining the integration of elements of the ACT model. Further, some of the few that exist are confounded by the inappropriate use of measures (e.g., the use of the AAQ-II as a measure of acceptance; Scott, McCracken, & Norton, 2016).

Ultimately, it is unclear whether the one factor finding implies conceptual problems with one or both measures, with the ACT model, or with some combination of these; it only points to a mismatch between the measures and the model as described (or interpreted). However, it would appear as though the content of the CFQ is consistent with the ACT definition of fusion and the psychometrics of the measure are robust. Simultaneously, the content of the AAQ-II items does not comprehensively match well

the definition of psychological flexibility as “...contacting the present moment as a conscious human being, fully and without needless defense—as it is and not as what it says it is—and persisting with or changing behavior in the service of chosen values” (Hayes, Strosahl, & Wilson, 2012, p. 96–7), and several concerns and confusions have been raised within the literature about the measure. Perhaps a useful path forward, in addition to a conceptual clarification of the ACT model, is to develop robust measures of the remaining ACT processes. Such measures could then either be used as a composite measure of psychological flexibility directly or be used to develop a measure that better represents the unique combination of these processes that represents psychological flexibility—whichever experts believe is a better conceptual representation of the ACT model.

Limitations of the present study. Although the very large sample size is, in general, a boon of the study, a couple of specific points are worth noting. The strengths of a large sample are that it allows for confident conclusions to be drawn from even complex analyses—in the sense that it provides substantial statistical power—and can also reduce the chances of imperfect data impacting the veracity of said conclusions (e.g., in the case of heteroscedasticity).

On the other hand, increased power widens the gap between statistical significance and practical or clinical significance. Virtually every relationship in the present study that was analyzed (not only those hypothesized) resulted in a statistically significant result, often extremely so (i.e., at a $p < .001$ level), presumably due to the increased power of a very large sample. This sensitivity may statistically identify a relationship that is less meaningful practically; for example, the CFQ’s incremental

validity in predicting the DASS21 Anxiety subscale score beyond the ASI-3 may be such a case (based on a change value of $\Delta R^2 = .018$).

Further, although many statisticians downplay situational violations in the data of assumptions of chosen methods, many issues are far from having robust, rigorously based consensuses within the statistical community. It is therefore possible to proceed with analysis in a large sample in spite of, for instance, a certain level of non-normality, toward false conclusions. That said, even in a perfect data situation, the profession still allows a 5% Type I error rate across all conclusions in most cases in clinical psychology. Ultimately, all conclusions should be viewed with care and aware of these considerations.

Limitations of the CFQ and Future Directions

Although the present study confirms that the CFQ is psychometrically robust, it is not without limitations. Near future studies need to address remaining questions about the psychometric characteristics of the CFQ, related to both conceptual issues and to clinical applications of the measure.

Construct validity. First, the overall validity of the construct as measured by the CFQ must be confirmed. As discussed earlier, the CFQ arguably does not comprehensively assess all aspects of fusion included in the conceptual literature. The empirical approach taken by the developers was important to ensure statistical robustness of a unidimensional construct. Further, it clearly has items assessing some important aspects of fusion, such as one's relationship with thoughts (e.g., "I tend to get very entangled in my thoughts"), emotion-elicitation (e.g., "My thoughts cause me distress or emotional pain"), and dominance of thoughts over behavior (e.g., "I get so caught up in my thoughts that I am unable to do the things that I most want to do"). However, the

CFQ may also omit important aspects of the construct, such as literality or even a more direct assessment of the evaluative aspects of thought that defusion interventions attempt to address. In fact, some items eliminated during development may represent precisely these aspects (e.g., “I make judgments about whether my thoughts are good or bad” and “my thoughts are facts” for evaluation and literality, respectively), calling into question the assumption of a unidimensional nature of the fusion construct.

Generalizability. Additionally, although the present study confirmed basic psychometrics in a way that expands its utility, broad generalization of the measure across cultural facets remains unclear. For example, there have been no studies of the CFQ with adolescents or children; virtually all samples investigated so far were predominantly female; and other than the translational studies in Spain and Korea, few cultures other than the U.K. and U.S. have been studied—particularly more eastern cultures that differ the most from the cultures of the U.K. and the U.S. and that may also show increased prevalence of many of the ideas that ACT is based on (e.g., those from Buddhism).

Mono-method assessment. Ultimately, even if the above recommended confirmations are performed, the CFQ is inherently limited by the fact that it is a single self-report measure. Although single measure investigations are very common in psychology, construct validity is threatened by single instrument measurement as well as mono-method bias. Currently, the CFQ is arguably the best self-report measure of the construct of fusion, as represented in the ACT model. Threats to construct validity could be reduced with the additional use of one of the other fusion measures discussed earlier—the best likely being the DDS. While the incremental improvement in validity is unlikely

to be substantial due to the flaws in the DDS noted earlier, given the narrowed result of the CFQ's items (discussed above), it may be advisable to include a second self-report measure in order to reduce the chances that an under-representation of the construct contributes to erroneous results (i.e., a mono-operation bias).

A superior approach would be to augment the measurement of fusion by the CFQ with a second measure that is not self-report. This would follow the recommended gold standard of multi-trait, multi-method assessment, and could significantly reduce the chance of bias in measurement (in particular mono-method bias). Unfortunately, there does not appear to be a broad and conceptually sound non-self-report measure of fusion in the literature, as of this writing. Accordingly, this approach would first require the development of such a measure.

Perhaps the best potential candidate for development would be to harness a method designed and executed by a group of RFT researchers called the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). In this approach, participants interact with a computer and have to match words according to the instructions given. They first engage in a practice phase where they are provided with feedback that is explicit in the case of incorrect responding, in order to confirm understanding of and compliance with instructions. They then engage in a test phase where they respond according to the same basic protocol. The key outcome is the response latency of participants. Additional steps are taken in the delivery of the IRAP method and analysis of the resulting data to reduce problems such as contamination of results by individual difference variables like motor skills and cognitive ability.

This method could be applied to behaviorally measure the implicit construct of fusion. It would first require the development of appropriate task items related to fusion and defusion. These could then be administered to participants with response latency used as a dimensional measure of fusion. The expectation would be that participants with high levels of fusion would respond with shorter latencies on the fusion items than participants with low levels of fusion (and longer latencies on the defusion items than those with low fusion).

Once developed, the IRAP method of fusion measurement could be combined with the self-report CFQ in a study to confirm the validity of the pairing. The ideal (though seldom performed) multitrait-multimethod matrix approach (Campbell & Fiske, 1959) should be performed to confirm the convergent validity of the two measures of fusion. The study would necessitate at least one other construct be studied (with similar methods of measurement)—perhaps of one of the other cognitive constructs discussed next—which would also allow for simultaneous examination of divergent validity, “killing two birds with one stone.” Once validated the CFQ and IRAP fusion measure could be used in tandem, minimizing threat to construct validity for future studies.

Discriminant/divergent validity. Discriminant or divergent validity was not well addressed by the present study, and was only considered somewhat in the original development studies by Gillanders and colleagues (2014; in relation to socially desirable responding). Absent from the literature is a thorough investigation of fusion and other important cognitive constructs. It is unclear why Gillanders and colleagues (2014) did not investigate some of these, as they had the Thought Control Questionnaire, Automatic Thoughts Questionnaire, and two measures of rumination: the PBRS and RSQ. The

present study's results reinforced the validity of the CFQ as a measure of fusion, and therefore, its value in being able to examine such relationships in future studies. Of particular importance may be the relationships between fusion and the frequently emphasized clinical constructs of worry, rumination, and decentering.

Despite the nature of the relationship between fusion and decentering outlined earlier, competing relationships between the constructs have been conceptualized (e.g., Bernstein, Hadash, Lichtash, Tanay, Shepherd, & Fresco, 2015) and the empirical literature on the subject is currently unclear. Some studies have found weak to moderate correlations between the two constructs (e.g., Naragon-Gainey & DeMarree, 2016), while others have hypothesized fusion as a factor within decentering, implying high levels of interdependence (e.g., McCracken, Barker, & Chilcot, 2014; Scott et al., 2016). Of course, some of these investigations are possibly muddled by the fusion measure(s) they chose, with some even using the AAQ-II due to a paucity of validated fusion measures.

Also unclear is the relationship between fusion and rumination. Researchers have variously proposed fusion as a similar process to rumination (e.g., McCracken et al, 2014), fusion as a mediating process for rumination (Nitzan-Assayag, Aderka, & Bernstein, 2015), rumination as a mediating process for fusion (e.g., Fresnics & Borders, 2016), both constructs as separate processes (e.g., Romero-Moreno, Márquez-González, Losada, Fernández-Fernández, & Nogales-González, 2015), or defusion as an intervention for rumination (Yovel, Mor, & Shakarov, 2014). Furthermore, many previous study conclusions are again confounded by issues with instruments used to measure fusion. The first study to examining (a Spanish translation of) the CFQ and a measure of rumination found a large correlation ($r = .63$) between the two constructs but

did not investigate further (Romero-Moreno et al., 2014). Multiple future studies will be needed to confirm the relationship between fusion and rumination more precisely.

The relationship between fusion and worry is perhaps the least studied of the three pairs. Worry has most frequently been studied in relation to the construct of thought-action fusion (e.g., Coles, Mennin, & Heimberg, 2001); however, this construct represents only a small subset of the construct of fusion as represented in ACT (see earlier discussion for a distinction from the ACT construct of fusion). Investigations into ACT concepts and worry have more commonly focused on the relationship between worry and mindfulness or the overarching construct of psychological flexibility (e.g., Ruiz, 2014). In the only identifiable extant, peer-reviewed, empirical investigation, defusion was shown to mediate worry and predict reductions in worry based on ACT and CBT interventions in a sample diagnosed with anxiety disorders (Arch, Wolitzky-Taylor, Eifert, & Craske, 2012). Future research should utilize the CFQ to more closely study promising models of worry that involve conceptualized relations with fusion (e.g., Roemer & Orsillo, 2002; Salters, Raffa, & Orsillo, 2005).

Requirements for clinical application. Beyond theoretical issues related to the construct, presumably a primary purpose of developing the CFQ is for use in applied clinical studies including intervention trials. For use in these contexts, a number of psychometrics should be more robustly confirmed. Test-retest reliability was not examined with a clinical sample, and in fact, only two samples so far considered were comprised of clinical mental health participants. The CFQ has also not yet been tested for longitudinal predictive capability nor for its ability to detect change based on clinical intervention. Further, the lack of time-based contextual cues in the measure may result in

lowered sensitivity to the short-term change hoped for in treatment trials; interestingly, this design decision was made in spite of the fact that nowhere in the literature is fusion conceptualized as a trait characteristic. Each of these characteristics should be considered in relation to the clinically applied use of the CFQ.

Conclusion

All but one of the present study's hypotheses were supported in a very large, adult, U.S. sample, resulting in the independent confirmation of the CFQ's basic psychometric properties. Results also extended the understanding of the CFQ's relation to other clinically important constructs beyond the original development studies. The single exception within the hypotheses was the relationship between the CFQ and the AAQ-II, with the present study finding a single factor solution suggesting that they are both measuring the same underlying construct in the present sample. This issue will have to be resolved by future investigations of the ACT model and component processes. Future studies should also perform further psychometric confirmation of the CFQ, particularly related to clinical application. While the CFQ is a substantive step in the direction of the empirical study of the construct of fusion, these confirmations are needed before it can be used confidently in the pursuit of an enhanced and causal understanding of how and why ACT works.

References

- A-Tjak, J. G. L., Davis, M. L., Morina, N., Powers, M. B., Smits, J. A. J., & Emmelkamp, P. M. G. (2015). A meta-analysis of the efficacy of acceptance and commitment therapy for clinically relevant mental and physical health problems. *Psychotherapy and Psychosomatics, 84*, 30-36. doi:10.1159/000365764
- American Psychological Association. (2012). *Resolution on the recognition of psychotherapy effectiveness*. Retrieved from <http://www.apa.org/about/policy/resolution-psychotherapy.aspx>
- American Psychological Association, Division of Clinical Psychology, Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48*(1), 3-23.
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176-181. doi:10.1037/1040-3590.10.2.176
- Arch, J. J., Eifert, G. H., Davies, C., Plumb, J. C., Rose, R. D., & Craske, M. G. (2012). Randomized clinical trial of cognitive behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for mixed anxiety disorders. *Journal of Consulting and Clinical Psychology, 80*(5), 750-765. doi:10.1037/a0028310

- Arch, J. J., Wolitzky-Taylor, K. B., Eifert, G. H., & Craske, M. G. (2012). Longitudinal treatment mediation of traditional cognitive behavioral therapy and acceptance and commitment therapy for anxiety disorders. *Behaviour Research and Therapy*, *50* (7-8), 469-478. doi:10.1016/j.brat.2012.04.007
- Bach, P. & Hayes, S. C. (2002). The use of acceptance and commitment therapy to prevent the rehospitalization of psychotic patients: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *70*(5), 1129-1139. doi:10.1037/0022-006X.70.5.1129
- Bach, P., Hayes, S. C., & Gallop, R. (2012). Long-term effects of brief acceptance and commitment therapy for psychosis. *Behavior Modification*, *36*(2), 165-181. doi:10.1177/0145445511427193
- Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2013). Testing a hierarchical model of distress tolerance. *Journal of Psychopathology and Behavioral Assessment*, *35*, 495-505. doi:10.1007/s10862-013-9359-0
- Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & Ehrenreich-May, J. (2010). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. Oxford, U.K.: Oxford University Press.
- Bartlett, M. S. (1954). A further note on the multiplying factors for various χ^2 approximations in factor analysis. *Journal of the Royal Statistical Society*, *16*, 296-298.

- Bernstein, A., Hadash, Y., Lichtash, Y., Tanay, G., Shepherd, K., & Fresco, D. M. (2015). Decentering and related constructs: A critical review and metacognitive processes model. *Perspectives on Psychological Science, 10*(5), 599-617. doi:10.1177/1745691615594577
- Blackledge, J. T. & Hayes, S. C. (2006). Using acceptance and commitment training in the support of parents of children diagnosed with autism. *Child & Family Behavior Therapy, 28*(1), 1-18. doi:10.1300/J019v28n01_01
- Blais, M. A., Lenderking, W. R., Baer, L., deLorell, A., Peets, K., Leahy, L., & Burns, C. (1999). Development and initial validation of a brief mental health outcome measure. *Journal of Personality Assessment, 73*(3), 359-373. doi:10.1207/S15327752JPA7303_5
- Bloom, D. E., Cafiero, E. T., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R.,...Weinstein, C. (2011). *The global economic burden of non-communicable diseases*. Geneva: World Economic Forum. Retrieved from: <http://www.weforum.org/reports/global-economic-burden-non-communicable-diseases>
- Bond, F. W. & Bunce, D. (2000). Mediators of change in emotion-focused and problem-focused worksite stress management interventions. *Journal of Occupational Health Psychology, 5*(1), 156-163. doi:10.1037/1076-8998.5.1.156

- Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K.,...,Zettle, R. D. (2011). Preliminary psychometric properties of the acceptance and action questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*, 676-688.
doi:10.1016/j.beth.2011.03.007
- Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*, 17-31. doi:10.1016/j.cpr.2005.07.003
- Cahill, S. P., Carrigan, M. H., Frueh, B. C. (1999). Does EMDR work? And if so, why?: A critical review of controlled outcome and dismantling research. *Journal of Anxiety Disorders, 13*(1-2), 5-33.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.
- Centers for Disease Control and Prevention, Division of Violence Prevention, National Center for Injury Prevention and Control (2012). Suicide: Facts at a glance 2012. Retrieved from: http://www.cdc.gov/ViolencePrevention/pdf/Suicide_DataSheet-a.pdf.
- Chawla, N., & Ostafin, B. (2007). Experiential avoidance as a functional dimensional approach to psychopathology: An empirical review. *Journal of Clinical Psychology, 63*, 871–890.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale, N.J.: Lawrence Erlbaum.

- Coles, M. E., Mennin, D. S., & Heimberg, R. G. (2001). Distinguishing obsessive features and worries: The role of thought-action fusion. *Behaviour Research and Therapy, 39*(8), 947-959. doi:10.1016/S0005-7967(00)00072-3
- Coombs, W. T., Algina, J., & Oltman, D. O. (1996). Univariate and multivariate omnibus hypothesis tests selected to control type I error rates when population variances are not necessarily equal. *Review of Educational Research, 66*(2), 137-179.
- Creswell, J. D., Welch, W. T., Taylor, S. E., Sherman, D. K., Gruenewald, T. L., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science, 16*, 846-851.
doi:10.1111/j.1467-9280.2005.01624.x
- Dahl, J. A., Wilson, K. G., & Nilsson, A. (2004). Acceptance and commitment therapy and the treatment of persons at risk for long-term disability resulting from stress and pain symptoms: A preliminary randomized trial. *Behavior Therapy, 35*(4), 785-801. doi:10.1016/S0005-7894(04)80020-0
- Dalrymple, K. L. & Herbert, J. D. (2007). Acceptance and commitment therapy for generalized social anxiety disorder: A pilot study. *Behavior Modification, 31*(5), 543-568. doi:10.1177/0145445507302037
- Drossel, C., Waltz, T. J., & Hayes, S. C. (2007). An introduction to principles of behavior. In D. Woods & J. Kantor (Eds.) *Understanding behavior disorders: A contemporary behavior analytic perspective* (pp. 21-46). Reno, NV: Context Press.

- 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*(3), 272-299.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/BF03193146
- Fergus, T. A. (2015). I really believe I suffer from a health problem: Examining an association between cognitive fusion and healthy anxiety. *Journal of Clinical Psychology, 71*(9), 920-934. doi:10.1002/jclp.22194.
- Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment, 1*, 84-94. doi:10.1111/j.1468-2389.1993.tb00092.x
- Fernández, R., Secades, R., Terrados, N., García, E., & García, J. M. (2004). Efecto de la hipnosis y de la terapia de aceptación y compromiso (ACT) en la mejora de la fuerza física en piragüistas. *International Journal of Clinical and Health Psychology, 4*(3), 481-493. Retrieved from:
http://www.aepc.es/ijchp/articulos_pdf/ijchp-119.pdf
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London, England: SAGE Publications.
- Folkman, S., Lazarus, R. S., Gruen, R. J., & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of Personality and Social Psychology, 50*(3), 571-579. doi:10.1037/0022-3514.50.3.571

- Fonagy, P., & Target, M. (2002). Early intervention and the development of self-regulation. *Psychoanalytic Inquiry, 22*(3), 307-335.
doi:10.1080/07351692209348990.
- Forman, E. M., Herbert, J. D., Juarascio, A. S., Yeomans, P. D., Zebell, J. A., Goetter, E. M., & Moitra, E. (2012). The Drexel defusion scale: A new measure of experiential distancing. *Journal of Contextual Behavioral Science, 1*, 55-65.
doi:10.1016/j.jcbs.2012.09.001.
- Forman, E. M., Herbert, J. D., Moitra, E., Yeomans, P. D., & Geller, P. A. (2007). A randomized controlled effectiveness trial of acceptance and commitment therapy and cognitive therapy for anxiety and depression. *Behavior Modification, 31*(6), 772-799. doi:10.1177/0145445507302202
- Fresco, D. M., Moore, M. T., van Dulmen, M. H. M., Segal, Z. V., Ma, S. H., Teasdale, J. D., & Williams, J. M. G. (2007). Initial psychometric properties of the experiences questionnaire: Validation of a self-report measure of decentering. *Behavior Therapy, 38*, 234-246.
- Fresco, D. M., Segal, Z. V., Buis, T., & Kennedy, S. (2007). Relationship of posttreatment decentering and cognitive reactivity to relapse in major depression. *Journal of Consulting and Clinical Psychology, 75*(3), 447-455.
doi:10.1037/0022-006X.75.3.447.
- Fresnics, A., & Borders, A. (2016). Angry rumination mediates the unique associations between self-compassion and anger and aggression. *Mindfulness, 1*-11.
doi:10.1007/s12671-016-0629-2

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, C. (2011).

Development of a measure of experiential avoidance: The multidimensional experiential avoidance questionnaire. *Psychological Assessment, 23*, 692-713. doi:10.1037/a0023242.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, C. (2014).

The brief experiential avoidance questionnaire: Development and initial validation. *Psychological Assessment, 26*, 35-45. doi:10.1037/a0034473.

Gaudiano, B. A. & Herbert, J. D. (2006). Acute treatment of inpatients with psychotic symptoms using acceptance and commitment therapy: Pilot results. *Behaviour Research and Therapy, 44*, 415-437. doi:10.1016/j.brat.2005.02.007

Gifford, E. V., Kohlenberg, B. S., Hayes, S. C., Antonuccio, D. O., Piasecki, M. M., Rasmussen-Hall, M. L., & Palm, K. M. (2004). Acceptance-based treatment for smoking cessation. *Behavior Therapy, 35*, 689-705. doi:10.1016/S0005-7894(04)80015-7

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., ..., & Remington, B. (2014). The development and initial validation of the cognitive fusion questionnaire. *Behavior Therapy, 45*(1), 83-101.

doi:10.1016/j.beth.2013.09.001

Gloster, A. T., Klotsche, J., Chaker, S., Hummel, K. V., & Hoyer, J. (2011). Assessing psychological flexibility: What does it add above and beyond existing constructs? *Psychological Assessment, 23*(4), 970-982. doi:10.1037/a0024135

- Godinho, A., Kushnir, V., & Cunningham, J. A. (2016). Unfaithful findings: Identifying careless responding in addictions research. *Addiction, 111*, 955-956.
doi:10.1111/add.13221
- Gratz, K. L., Tull, M. T., & Gunderson, J. G. (2008). Preliminary data on the relationship between anxiety sensitivity and borderline personality disorder: The role of experiential avoidance. *Journal of Psychiatric Research, 42*, 550-559.
doi:10.1016/j.jpsychires.2007.05.011
- Greco, L. A., Lambert, W., & Baer, R. A. (2008). Psychological inflexibility in childhood and adolescence: Development and evaluation of the avoidance and fusion questionnaire for youth. *Psychological Assessment, 20*(2), 93-102.
doi:10.1037/1040.3590.20.2.93.
- Gregg, J. A., Callaghan, G. M., Hayes, S. C., & Glenn-Lawson, J. L. (2007). Improving diabetes self-management through acceptance, mindfulness, and values: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 75*(2), 336-343. doi:10.1037/0022-006X.75.2.336
- Gross, J. J. (Ed.). (2009). *Handbook of emotion regulation*. New York, NY: Guilford Press.
- Gumley, A. (2010). The developmental roots of compromised mentalization in complex mental health disturbances of adulthood: An attachment-based conceptualization. In G. DiMaggio & P. H. Lysaker (Eds.), *Metacognition and severe adult mental disorders: From research to treatment*. Hove, U.K.: Routledge.

- Gutiérrez, O., Luciano, C., Rodríguez, M., & Fink, B. C. (2004). Comparison between an acceptance-based and a cognitive-control-based protocol for coping with pain. *Behavior Therapy, 35*, 767–784. doi:10.1016/S0005-7894(04)80019-4
- Hayes, S. C. (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. New York, NY: Plenum Press.
- Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*, 639-665. doi:10.1016/S0005-7894(04)80013-3
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic/Plenum.
- Hayes, S. C., Bissett, R., Roget, N., Padilla, M., Kohlenberg, B. S., Fisher, G., . . . , & Niccolls, R. (2004). The impact of acceptance and commitment training and multicultural training on the stigmatizing attitudes and professional burnout of substance abuse counselors. *Behavior Therapy, 35*(4), 821-835. doi:10.1016/S0005-7894(04)80022-4
- Hayes, S. C., Hayes, L. J., & Reese, H. W. (1988). Finding the philosophical core: A review of Stephen C. Pepper's World Hypotheses: A study in evidence. *Journal of the Experimental Analysis of Behavior, 50*(1), 97-111. doi:10.1901/jeab.1988.50-97

- Hayes, S. C., Levin, M., Plumb-Villardaga, J., Villatte, J., & Pistorello, J. (2013). Acceptance and Commitment Therapy and contextual behavioral science: Examining the progress of a distinctive model of behavioral and cognitive therapy. *Behavior Therapy* (44), 180-198. doi:10.1016/j.beth.2009.08.002
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., Lillis, J. (2006). Acceptance and commitment therapy: Model process and outcomes. *Behaviour Research and Therapy*, 44, 1-25. doi:10.1016/j.brat.2005.06.006
- Hayes, S. C., Pankey, J., Gifford, E. V., Batten, S. V., & Quiñones, R. (2002). Acceptance and commitment therapy in experiential avoidance disorders. In F. W. Kaslow & T. Patterson (Eds.), *Comprehensive handbook of psychotherapy: Cognitive-behavioral approaches, volume 2* (pp. 319-351). New York, NY: Wiley.
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). New York, NY: The Guilford Press.
- Hayes, S. C., Wilson, K. G., Gifford, E. V., Bissett, R. Piasecki, M., Batten, S. V.,...Gregg, J. (2004). A preliminary trial of twelve-step facilitation and acceptance and commitment therapy with polysubstance-abusing methadone-maintained opiate addicts. *Behavior Therapy*, 35, 667-688. doi:10.1016/S0005-7894(04)80014-5

- Henry, J. D. & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227-239.
doi:10.1348/014466505X29657
- Herzberg, K. N., Sheppard, S. C., Forsyth, J. P., Credé, M., Earleywine, M., & Eifert, G. H. (2012). The believability of anxious feelings and thoughts questionnaire: A psychometric evaluation of cognitive fusion in a nonclinical and highly anxious community sample. *Psychological Assessment, 24*(4), 877-891.
doi:10.1037/a0027782
- Hesser, H., Gustafsson, T., Lundén, C., Henrikson, O., Fattahi, K., Johnsson, E.,..., & Andersson, G. (2012). A randomized controlled trial of internet-delivered cognitive behavior therapy and acceptance and commitment therapy in the treatment of tinnitus (2012). *Journal of Consulting and Clinical Psychology, 80*(4), 649-661. doi:10.1037/a0027021
- Hilsenroth, M. J., Ackerman, S. J., & Blagys, M. D. (2001). Evaluating the phase model of change during short-term psychodynamic psychotherapy. *Psychotherapy Research, 11*(1), 29-47. doi:10.1080/713663851
- Hoaglin, D. C., and Iglewicz, B. (1987). Fine tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association, 82*, 1147-1149.
- Hoaglin, D.C., Iglewicz, B., and Tukey, J.W. (1986). Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association, 81*, 991-999.

- Hofmann, S. G. & Asmundson, G. J. G. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review, 28*, 1-16.
doi:10.1016/j.cpr.2007.09.003
- Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427-440. doi:10.1007/s10608-012-9476-1
- Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 169-183. doi:10.1037/a0018555
- Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*, 383-395.
- Hooper, N., Saunders, J., McHugh, L. (2010). The derived generalization of thought suppression. *Learning & Behavior, 38*(2), 160-168. doi:10.3758/LB.38.2.160
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27*(1), 99-114. doi:10.1007/s10869-011-9231-8
- Insel, T. R. (2008). Assessing the economic costs of serious mental illness. *The American Journal of Psychiatry, 165*(6), 663-665. doi:10.1176/appi.ajp.2008.08030366
- Institute for Health Metrics and Evaluation (2014). *United States Global Burden of Disease Study 2010 (GBD 2010) results 1990-2010*. Retrieved from:
<http://ghdx.healthdata.org/record/united-states-global-burden-disease-study-2010-gbd-2010-results-1990-2010>.

- Johns, L. C., Oliver, J. E., Khondoker, M., Byrne, M., Jolley, S., Wykes, T., . . . , Morris, E. M. J. (2016). The feasibility and acceptability of a brief acceptance and commitment therapy (ACT) group intervention for people with psychosis: The ‘ACT for life’ study. *Journal of Behavior Therapy and Experimental Psychiatry*, *50*, 257-263. doi:10.1016/j.jbtep.2015.10.001.
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, *39*, 103-129. doi:10.1016/j.jrp.2004.09.009
- Johnston, M., Foster, M., Shennan, J., Starkey, N. J., & Johnson, A. (2010). The effectiveness of an acceptance and commitment therapy self-help intervention for chronic pain. *Clinical Journal of Pain*, *26*(5), 393-402. doi:10.1097/AJP.0b013e3181cf59ce
- Kahl, K. G., Winter, L., & Schweiger, U. (2012). The third wave of cognitive behavioural therapies: What is new and what is effective? *Current Opinion in Psychiatry*, *25*(6), 522-528. doi:10.1097/YCO.0b013e328358e531
- Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, *34*, 111-117.
- Kessler R. C., Chiu, W. T., Demier, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617-627. doi:10.1001/archpsyc.62.6.617

- Kessler, R. C., Berglund, P., Demier, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 593-602. doi:10.1001/archpsyc.62.6.593
- Kim, B., & Cho, S. (2015). Psychometric properties of a Korean version of the cognitive fusion questionnaire. *Social Behavior and Personality*, 43(10), 1715-1724. doi:10.2224/sbp.2015.43.10.1715.
- Kurtz, J. E., & Parrish, C. L. (2001). Semantic response consistency and protocol validity in structured personality assessment: The case of the NEO-PI-R. *Journal of Personality Assessment*, 76, 315-322. doi:10.1207/S15327752JPA7602_12
- Lappalainen, R., Lehtonen, T., Skarp, E., Taubert, E., Ojanen, M., & Hayes, S. C. (2007). The impact of CBT and ACT models using psychology trainee therapists: A preliminary controlled effectiveness trial. *Behavior Modification*, 31(4), 488-511. doi:10.1177/0145445506298436
- Lazarus, A. A. (1967). In support of technical eclecticism. *Psychological Reports*, 21(2), 415-416.
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin & H. Hotelling (Eds.), *Contributions to probability and statistics: Essays in honor of Harold Hotelling* (pp. 278-292). Stanford, CA: Stanford University Press.
- Levin, M. & Hayes, S. C. (2009). Is acceptance and commitment therapy superior to established treatment comparisons? *Psychotherapy and Psychosomatics*, 78, 380. doi:10.1159/000235978

- Lillis, J. & Hayes, S. C. (2007). Applying acceptance, mindfulness, and values to the reduction of prejudice: A pilot study. *Behavior Modification*, 31(4), 389-411. doi:10.1177/0145445506298413
- Lillis, J., Hayes, S. C., Bunting, K., & Masuda, A. (2009). Teaching acceptance and mindfulness to improve the lives of the obese: A preliminary test of a theoretical model. *Annals of Behavioral Medicine*, 37, 58-69. doi:10.1007/s12160-009-9083.x
- Lovibond, S. H. & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney, Australia: Psychology Foundation.
- Lundgren, T., Dahl, J., & Hayes, S. C. (2008). Evaluation of mediators of change in the treatment of epilepsy with acceptance and commitment therapy. *Journal of Behavioral Medicine*, 31(3), 225–235. doi:10.1007/s10865-008-9151-x
- Lundgren, T., Dahl, J., Melin, L., & Kies, B. (2006). Evaluation of acceptance and commitment therapy for drug refractory epilepsy: A randomized controlled trial in South Africa – A pilot study. *Epilepsia*, 47(12), 2173-2179. doi:10.1111/j.1528-1167.2006.00892.x
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2, 49-55.
- Manos, R. C., Kanter, J. W., & Luo, W. (2011). The behavioral activation for depression scale – short form: Development and validation. *Behavior Therapy*, 42, 726-739. doi:10.1016/j.beth.2011.04.004

- Mantar, A., Yemez, B., Alkın, T., & Dergisi, P. (2010). The validity and reliability of the Turkish version of the Anxiety Sensitivity Index-3. *Turkish Journal of Psychiatry*, 21(3), 225–234.
- Marshall, E-J., & Brockman, R. N. (2016). The relationships between psychological flexibility, self-compassion, and emotional well-being. *Journal of Cognitive Psychotherapy: An International Quarterly*, 30(1), 60-72. doi:10.1891/0889-8391.30.1.60.
- Masuda, A., Hayes, S. C., Fletcher, L. B., Seignourel, P. J., Bunting, K., Herbst, S. A.,..., & Lillis, J. (2007). Impact of acceptance and commitment therapy versus education on stigma toward people with psychological disorders. *Behaviour Research and Therapy*, 45, 2764-2772. doi:10.1016/j.brat.2007.05.008
- Masuda, A., Hayes, S. C., Sackett, C. F., & Twohig, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: Examining the impact of a ninety year old technique. *Behaviour Research and Therapy*, 42, 477-485. doi:10.1016/j.brat.2003.10.008
- McDermott, M. J., Tull, M. T., Gratz, K. L., Daughters, S. B., & Lejuez, C. W. (2009). The role of anxiety sensitivity and difficulties in emotion regulation in posttraumatic stress disorder among crack/cocaine dependent patients in residential substance abuse treatment. *Journal of Anxiety Disorders*, 23(5), 591–599. doi:10.1016/j.janxdis.2009.01.006
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, 136(3), 450-470. doi:10.1037/a0019216

- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437-455. doi:10.1037/a0028085
- Merikangas, K. R., He, J., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2009). Prevalence and treatment of mental disorders among US children in the 2001-2004 NHANES. *Pediatrics, 125*(1), 75-81. doi:10.1542/peds.2008-2598
- Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S.,...Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980-989. doi:10.1016/j.jaac.2010.05.017
- Micceri, T. (1989). The unicorn, the normal curve, and other improbably creatures. *Psychological Bulletin, 105* (1), 156-166.
- Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D.,...Lopez, A. D. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*, 2197-2223. doi:10.1016/S0140-6736(12)61689-4
- Murray, C. J. L., Abraham, J., Ali, M. K., Alvarado, M., Atkinson, C.,...Lopez, A. D. (2013). The state of U.S. health, 1990-2010: Burden of diseases, injuries, and risk factors. *The Journal of the American Medical Association, 310*(6), 591-608. doi:10.1001/jama.2013.13805

- Muto, T., Hayes, S. C., & Jeffcoat, T. (2011). The effectiveness of acceptance and commitment therapy bibliotherapy for enhancing the psychological health of Japanese college students living abroad. *Behavior Therapy, 42*, 323-335. doi:10.1016/j.beth.2010.08.009
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2016). Detecting careless respondents in web-based questionnaires: Which method to use? *Journal of Research in Personality, 63*, 1-11. doi:10.1016/j.jrp.2016.04.010
- Nitzan-Assayag, Y., Aderka, I. M., & Bernstein, A. (2015). Dispositional mindfulness in trauma recovery: Prospective relations and mediating mechanisms. *Journal of Anxiety Disorders, 36*, 32-32. doi:10.1016/j.janxdis.2015.07.008
- Ng, F., Trauer, T., Dodd, S., Callaly, T., Campbell, S., & Berk, M. (2007). The validity of the 21-item version of the depression anxiety stress scales as a routine clinical outcome measure. *Acta Neuropsychiatrica, 19*, 304-310. doi:10.1111/j.1601-5215.2007.00217.x
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*, 673-690.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32*, 396-402.
- Omer, H., & Dar, R. (1992). Changing trends in three decades of psychotherapy research: The flight from theory into pragmatics. *Journal of Consulting and Clinical Psychology, 60*(1), 88-93.

- Ossman, W. A., Wilson, K. G., Storaasli, R. D., & McNeill, J. W. (2006). A preliminary investigation of the use of acceptance and commitment therapy in group treatment for social phobia. *International Journal of Psychology and Psychological Therapy*, 6(3), 397-416. Retrieved from:
<http://www.ijpsy.com/volumen6/num3/151/a-preliminary-investigation-of-the-use-of-EN.pdf>
- Öst, L. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy*, 46, 296-321.
doi:10.1016/j.brat.2007.12.005
- Powers, M. B., Vörding, M. B. Z. V. S., & Emmelkamp, P. M. G. (2009). Acceptance and commitment therapy: A meta-analytic review. *Psychotherapy and Psychosomatics*, 78, 73-80. doi:10.1159/000190790
- Rachman, S., Thordarson, D. S., Shafran, R., & Woody, S. R. (1995). Perceived responsibility: Structure and significance. *Behaviour Research and Therapy*, 33, 779-784.
- Rehm, J., Mathers, C. Popova, S., Thavorncharoensap, M., Teerawattananon, Y., & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *Lancet*, 373(9682), 2223-2233.
doi:10.1016/S0140-6736(09)60746-7
- Reiss, S., Peterson, R. A, Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*, 24(1), 1-8. doi:10.1016/0005-7967(86)90143-9

- Roemer, L., & Orsillo, S. M. (2002). Expanding our conceptualization of and treatment for generalized anxiety disorder: Integrating mindfulness/acceptance-based approaches with existing cognitive-behavioral models. *Clinical Psychology: Science and Practice, 9*, 54-68. doi:10.1093/clipsy.9.1.54
- Roemer, L., Salters, K., Raffa, S. D., & Orsillo, S. M. (2005). Fear and avoidance of internal experiences in GAD: Preliminary tests of a conceptual model. *Cognitive Therapy and Research, 71-88*. doi:10.1007/s10608-005-1650-2
- Romero-Moreno, R., Márquez-González, M., Losada, A., Fernández-Fernández, V., & Nogales-González, C. (2015). Rumiación y fusión cognitiva en el cuidado familiar de personas con demencia. *Revista Española de Geriatria y Gerontología, 50(5)*, 216-222. doi:10.1016/j.regg.2015.02.007
- Romero-Moreno, R., Márquez-González, M., Losada, A., Gillanders, D., & Fernández-Fernández, V. (2014). Cognitive fusion in dementia caregiving: Psychometric properties of the Spanish version of the “cognitive fusion questionnaire.” *Behavioral Psychology, 22(1)*, 117-132.
- Ruiz, F. J. (2010). A review of acceptance and commitment therapy (ACT) empirical evidence: Correlational, experimental psychopathology, component and outcome studies. *International Journal of Psychology and Psychological Therapy, 10(1)*, 125-162. Retrieved from: <http://www.ijpsy.com/volumen10/num1/256/a-review-of-acceptance-and-commitment-therapy-EN.pdf>

- Ruiz, F. J. (2012). Acceptance and commitment therapy versus traditional cognitive behavioral therapy: A systematic review and meta-analysis of current empirical evidence. *International Journal of Psychology and Psychological Therapy, 12*(2), 333-357. Retrieved from:
<http://www.ijpsy.com/volumen12/num3/334/acceptance-and-commitment-therapy-versus-EN.pdf>
- Ruiz, F. J. (2014). The relationship between low levels of mindfulness skills and pathological worry: The mediating role of psychological inflexibility. *Anales de Psicología, 30*(3), 887-897. doi:10.6018/analesps.30.3.150651
- Ruiz, F. J. & Luciano, C. (2009). Eficacia de la Terapia de Aceptación y Compromiso (ACT) en la mejora del rendimiento ajedrecístico de jóvenes promesas. *Psicothema, 21*(3), 347-352. Retrieved from:
<http://www.psicothema.es/pdf/3637.pdf>
- Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement, 10*(2), 135-154.
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29*(4), 304-321. doi:10.1177/0734282911406653
- Scott, W., McCracken, L. M., & Norton, S. (2016). A confirmatory factor analysis of facets of psychological flexibility in a sample of people seeking treatment for chronic pain. *Annals of Behavioral Medicine, 50*, 285-296. doi:10.1007/s12160-015-9752-x

- Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2 , a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, 3, 1-6. doi:10.3399/fpsyg.2012.00111
- Shafran, R., Thordarson, D., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders*, 10(5), 379-391.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. doi:10.1093/biomet/52.3-4.591.
- Smout, M. F., Longo, M., Harrison, S., Minniti, R., Wickes, W., & White J. M. (2010). Psychosocial treatment for methamphetamine use disorders: A preliminary randomized controlled trial of cognitive behavior therapy and acceptance and commitment therapy. *Substance Abuse*, 31, 98-107.
doi:10.1080/08897071003641578
- Solé, E., Racine, M., Castarlenas, E., de la Vega, R., Tomé-Pires, C., Jensen, M., & Miró, J. (2015). The psychometric properties of the cognitive fusion questionnaire in adolescents. *European Journal of Psychological Assessment*, 32(3), 181-186.
doi:10.1027/1015-5759/a000244.
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730-737.
doi:10.2307/2286009.

- Substance Abuse and Mental Health Services Administration (2013a). *Results from the 2012 National Survey on Drug Use and Health: Mental health findings*, NSDUH Series H-47, HHS Publication No. (SMA) 13-4805. Rockville, MD: Substance Abuse and Mental Health Services Administration. Retrieved from: http://www.samhsa.gov/data/NSDUH/2k12MH_FindingsandDetTables/2K12MHF/NSDUHmhfr2012.htm.
- Substance Abuse and Mental Health Services Administration (2013b). *Results from the 2012 National Survey on Drug Use and Health: Summary of national findings*, NSDUH Series H-46, HHS Publication No. (SMA) 13-4795. Rockville, MD: Substance Abuse and Mental Health Services Administration. Retrieved from: <http://www.samhsa.gov/data/NSDUH/2012SummNatFindDetTables/NationalFindings/NSDUHresults2012.htm>.
- Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and commitment therapy in the treatment of anxiety: A systematic review. *Clinical Psychology Review, 33*, 965-978. doi:10.1016/j.cpr.2013.07.002
- Tabachnick, B. G., & Fidell, L. S. (2001). Principal components and factor analysis. In *Using multivariate statistics* (4th ed., pp. 582–633). Needham Heights, MA: Allyn & Bacon.
- Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R.,..., Bouvard, M. (2007). Robust dimensions of anxiety sensitivity: development and initial validation of the Anxiety Sensitivity Index-3. *Psychological Assessment, 19*(2), 176–188. doi:10.1037/1040-3590.19.2.176

- Törneke, N. (2010). *Learning RFT: An introduction to relational frame theory and its clinical application*. Oakland, CA: Context Press.
- Twohig, M. P., Hayes, S. C., Plumb, J. C., Pruitt, L. D., Collins, A. B., Hazlett-Stevens, H., & Woidneck, M. R. (2010). A randomized clinical trial of acceptance and commitment therapy versus progressive relaxation training for obsessive compulsive disorder. *Journal of Consulting and Clinical Psychology, 78*(5), 705-716. doi:10.1037/a0020508
- Twohig, M. P. (2012). Introduction: The basics of acceptance and commitment therapy. *Cognitive and Behavioral Practice, 19*, 499-507. doi:10.1016/j.cbpra.2012.04.003
- United States Department of Justice, National Drug Intelligence Center (2011). *The economic impact of illicit drug use on American society*, Product No. 2011-Q0317-002. Washington D. C.: United States Department of Justice. Retrieved from: <http://www.justice.gov/archive/ndic/pubs44/44849/44849p.pdf>.
- Van Brunt, D. L. (2000). Modular cognitive-behavioral therapy: Dismantling validated treatment programs into self-standing treatment plan objectives. *Cognitive and Behavioral Practice, 7*, 156-165.
- Vowles, K. E., McCracken, L. M., & O'Brien, J. Z. (2011). Acceptance and values-based action in chronic pain: A three-year follow-up analysis of treatment effectiveness and process. *Behaviour Research and Therapy, 49*, 748-755. doi:10.1016/j.brat.2011.08.002
- Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record, 61*(3), 371-390. Retrieved from: <http://opensiuc.lib.siu.edu/tpr/vol61/iss3/5>

- Welch, B. L. (1947). The generalization of 'student's' problem when several different population variances are involved. *Biometrika*, 34(1/2), 28-35.
- Wetherell, J. L., Afari, N., Ayers, C. R., Stoddard, J. A., Ruberg, J., Sorrell, J. T.,..., & Patterson, T. L. (2011). Acceptance and commitment therapy for generalized anxiety disorder in older adults: A preliminary report. *Behavior Therapy*, 42, 127-134. doi:10.1016/j.beth.2010.07.002
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J.,... Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *The Lancet*, 382, 1575-1586. doi:10.1016/S0140-6736(13)61611-6
- Williams, L.M. (2006). *Acceptance and commitment therapy: An example of third-wave therapy as a treatment for Australian Vietnam War veterans with posttraumatic stress disorder*. Unpublished dissertation, Charles Stuart University, Bathurst, New South Wales.
- Wilson, K. G. & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, 66(3), 267-281. doi:10.1901/jeab.1993.66-267
- Wolgast, M. (2014). What does the acceptance and action questionnaire (AAQ-II) really measure? *Behavior Therapy*, 45, 831-839.
- World Health Organization (2014). Suicide Prevention (SUPRE). Retrieved from: http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/.

World Health Organization, Department of Mental Health and Substance Abuse. (2011).

Mental health atlas: United States of America profile. Retrieved from:

http://www.who.int/mental_health/evidence/atlas/profiles/usa_mh_profile.pdf?ua=1.

Young, J. L., Waehler, C. A., Laux, J. M., McDaniel, P. S., & Hilsenroth, M. J. (2003).

Four studies extending the utility of the Schwartz Outcome Scale (SOS-10).

Journal of Personality Assessment, *80*(2), 130-138.

doi:10.1207/S15327752JPA8002_02

Yovel, I., Mor, N., & Shakarov, H. (2014). Examination of the core cognitive

components of cognitive behavioral therapy and acceptance and commitment

therapy: An analogue investigation. *Behavior Therapy*, *45*(4), 482-494.

doi:10.1016/j.beth.2014.02.007

Zettle, R. D., & Hayes, S. C. (1986). Dysfunctional control by client verbal behavior: The

context of reason giving. *The Analysis of Verbal Behavior*, *4*, 30–38. Retrieved

from:

<http://www.ncbi.nlm.nih.gov.ezproxy.emich.edu/pmc/articles/PMC2748463/pdf/anverbbehav00041-0032.pdf>

Zettle, R. D., & Hayes, S. C. (1987). Component and process analysis of cognitive

therapy. *Psychological Reports*, *61*, 939-953.

Zettle, R. D. & Rains, J. C. (1989). Group cognitive and contextual therapies in treatment

of depression. *Journal of Clinical Psychology*, *45*, 436-445. doi:10.1002/1097-

4679(198905)45:3%3C436::AID-JCLP2270450314%3E3.0.CO;2-L

Zvolensky, M. J., Bernstein, A., & Vujanovik, A. A. (Eds.). (2010). *Distress tolerance: Theory, research, and clinical applications*. New York, NY: Guilford Press.

APPENDICES

Appendix A: IRB Approval Letter

RESEARCH @ EMU**UHSRC Determination: FULL BOARD INITIAL APPROVAL**

DATE: December 12, 2014

TO: Barry Eye, M.S.
Department of Psychology
Eastern Michigan University

Re: UHSRC: # 661999-1
Approval Date: December 12, 2014
Expiration Date: December 11, 2015

Title: A Brief Acceptance and Commitment Therapy Group Intervention for Anxiety Sensitivity

Your research project, entitled **A Brief Acceptance and Commitment Therapy Group Intervention for Anxiety Sensitivity**, has been approved in accordance with all applicable federal regulations.

This approval includes the following:

1. Enrollment of up to 500 subjects to participate in the approved protocol.
2. Use of the following study measures: AAQ-II, ASI-3, BADS-SF, CFQ, DASS-21, Demographic questionnaire, SOS-10, and Treatment Satisfaction Survey.
3. Use of the stamped screening and treatment consent forms.

Renewals: This approval is valid for one year and expires on December 11, 2015. If you plan to continue your study beyond December 11, 2015, you must submit a Continuing Review Form by October 25, 2015 to ensure the approval does not lapse.

Modifications: All changes must be approved prior to implementation. If you plan to make any minor changes, you must submit a **Minor Modification Form**. For any changes that alter study design or any study instruments, you must submit a **Human Subjects Approval Request Form**. These forms are available through IRBNet on the UHSRC website. Please note that major modifications will require Full Board review and should be submitted at least 30 days in advance to allow for the UHSRC monthly meeting schedule.

Problems: All major deviations from the reviewed protocol, unanticipated problems, adverse events, subject complaints, or other problems that may increase the risk to human subjects **or** change the category of review must be reported to the UHSRC via an **Event Report** form, available through IRBNet on the UHSRC website

Follow-up: If your Expedited research project is not completed and closed after **three years**, the UHSRC office requires a new **Human Subjects Approval Request Form** prior to approving a continuation beyond three years.

Please use the UHSRC number listed above on any forms submitted that relate to this project, or on any correspondence with the UHSRC office.

Good luck in your research. If we can be of further assistance, please contact us at 734-487-3090 or via e-mail at human.subjects@emich.edu. Thank you for your cooperation.

Sincerely,

Jennifer Kellman Fritz, PhD
Chair
University Human Subjects Review Committee

Appendix B: Informed Consent Form

**INFORMED CONSENT FOR RESEARCH SCREENING
A Brief Acceptance and Commitment Therapy Group Intervention for Anxiety
Sensitivity**

Barry Eye, M.S., Doctoral Fellow – Principal Investigator
Ellen Koch, PhD, Associate Professor of Psychology – Co-Investigator

Purpose of the Study and How Long it Will Last

The purpose of this screening is to determine whether you are eligible to participate in a research study examining an intervention to help more effectively manage stress, based on set inclusion and exclusion criteria. We cannot tell you in advance what the eligibility criteria are, but it is anticipated that many students will be eligible. The screening will be completed online and should only take approximately ten minutes to complete.

Participation Withdrawal or Refusal to Participate

Participation in this study is completely voluntary. If you decide to participate, you can change your mind at any time and withdraw from the study. Refusing to participate will involve no penalty or loss of benefits. In order to withdraw, simply close the browser window at any time during this phase of the study.

Description of Study Procedures

For this study you will be asked to fill out an online survey that includes questionnaires about a specific type of stress, symptoms of stress, coping strategies, quality of life, and demographic and background information such as gender, age, and race/ethnicity. Once you have completed this survey, you will be asked if you are interested in participating in the treatment phase of this study. If you indicate interest in the treatment phase of this study, you will be asked for your name and contact information. You are not obligated to participate in the treatment phase of this study, but can elect to do so if interested.

Confidentiality of Information Obtained

All responses and personally identifiable information will be kept confidential within the confines of Qualtrics's privacy policy (see <http://www.qualtrics.com/privacy-statement/> for further information). Your personal responses will only be released to the principal investigator, who will download all the responses off the internet at the end of the study and delete the information off of qualtrics.com. At this point, any identifying information will be separated from your survey responses and you will be given an identification number to use throughout the study to protect your confidentiality. However, to ensure that you are using the same number throughout the study, the principal investigator will keep a separate password-protected file linking personally identifiable information and identification numbers. Once all data has been collected, this file will be destroyed. Information from this study may be reported or published in aggregated form, but your anonymity will be maintained in any publications or presentations.

Expected Risks of the Study

There are no known or anticipated risks for participating in this phase of the study. Nevertheless, you may experience some mild emotional discomfort when completing the study, but it is not expected to last longer than it takes you to complete the study. If, however, you experience emotional reactions that are difficult for you to manage, you can contact the principal investigator for referral information.

Expected Benefits of the Study

Your participation in this study will help us identify potential participants for the treatment phase of the study. Personal benefits of participation include learning a bit about how psychologists conduct research as well as contributing to the psychological literature.

Compensation for Participation

If you are an EMU psychology student, it is possible that you may receive extra course credit in accordance with the guidelines established by your psychology course instructor. In such cases, we will provide your instructor with your name and verification of participation so that this extra credit can be awarded to you per your instructor's course policy. There is no monetary compensation for your participation in this phase of the study.

Use of Research Results

Findings from this study may be published in psychological journals and may also be presented at professional conferences. In addition, the data being collected will be used in the Principal Investigator's dissertation and may appear in that published document. As a participant you will also be sent a brief summary of the results of the study at its conclusion, and are entitled to meet with the principal investigator to discuss any other questions or concerns regarding the study.

Questions About the Study

If at any time you have questions about the study procedures or your participation in the study, please contact the principal investigator, Mr. Barry Eye (*Phone: 734-487-6715; Email: beye@emich.edu*) or the co-investigator, Dr. Ellen Koch (*Phone: 734-487-0189; Email: ellen.koch@emich.edu*).

Human Subjects Review Board

For questions about your rights as a participant in human subject research, you can contact the Eastern Michigan University Human Subjects Review Committee (UHSRC) at human.subjects@emich.edu or 734-487-3090. The UHSRC is responsible for the safety and protection of people who participate in research.

CONSENT TO PARTICIPATE: I have read this form. I have been given the opportunity to ask questions, and all my questions have been answered to my satisfaction. I give my consent to participate in this study. By consenting I am also indicating that I am eighteen or older and am able to freely consent to participation.

You may also print out a copy of this consent form for your future reference if you desire. If you have read all of the above and would like to take part in the screening phase of this study, please click the “next” button below. By doing so, you are giving informed consent for us to use your responses in this study.

If you do not wish to take part in this study, please close the browser window now.

Appendix C: Demographics Questionnaire

Gender: Male Female Other: _____

Age: _____

Race/Ethnicity (*Check all that apply*):

- Mexican, Mexican American, Chicano
- Puerto Rican
- Cuban
- Other Hispanic, Latino, or Spanish: _____
- Black or African American
- White
- American Indian or Alaska Native
- Native Hawaiian
- Guamanian or Chamorro
- Samoan
- Other Pacific Islander: _____
- Asian Indian
- Chinese
- Filipino
- Japanese
- Korean
- Vietnamese
- Other Asian: _____
- Some other race: _____

Are you currently engaged in therapy or any similar services? Yes No

Are you currently taking any psychotropic medications (e.g., anti-depressants)? Yes No

Which of your family members have past and/or present anxiety (*check all that apply*):

- One or more primary caregiver(s), such as a parent
- One or more other immediate family member(s), such as a brother or sister
- One or more other relatives, such as a grandparent

Appendix D: AAQ-II

Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.

1	2	3	4	5	6	7
never true	very seldom true	seldom true	sometimes true	frequently true	almost always true	always true

1. My painful experiences and memories make it difficult for me to live a life that I would value.	1	2	3	4	5	6	7
2. I'm afraid of my feelings.	1	2	3	4	5	6	7
3. I worry about not being able to control my worries and feelings.	1	2	3	4	5	6	7
4. My painful memories prevent me from having a fulfilling life.	1	2	3	4	5	6	7
5. Emotions cause problems in my life.	1	2	3	4	5	6	7
6. It seems like most people are handling their lives better than I am.	1	2	3	4	5	6	7
7. Worries get in the way of my success.	1	2	3	4	5	6	7

Appendix E: ASI-3

Please select the number that best corresponds to how much you agree with each item. If any items concern something that you have never experienced (e.g., fainting in public), then answer on the basis of how you think you might feel *if you had* such an experience. Otherwise, answer all items on the basis of your own experience. Be careful to circle of one number for each item and please answer all items.

	0	1	2	3	4	
	never true	very seldom true	seldom true	sometimes true	frequently true	
1. It is important for me not to appear nervous	0	1	2	3	4	S
2. When I cannot keep my mind on a task, I worry that I might be going crazy	0	1	2	3	4	C
3. It scares me when my heart beats rapidly	0	1	2	3	4	P
4. When my stomach is upset, I worry that I might be seriously ill	0	1	2	3	4	P
5. It scares me when I am unable to keep my mind on a task	0	1	2	3	4	C
6. When I tremble in the presence of others, I fear what people might think of me	0	1	2	3	4	S
7. When my chest feels tight, I get scared that I won't be able to breath properly	0	1	2	3	4	P
8. When I feel pain in my chest, I worry that I'm going to have a heart attack	0	1	2	3	4	P
9. I worry that other people will notice my anxiety	0	1	2	3	4	S
10. When I feel "spacey" or spaced out I worry that I may be mentally ill	0	1	2	3	4	C
11. It scares me when I blush in front of people	0	1	2	3	4	S
12. When I notice my heart skipping a beat, I worry that there is something seriously wrong with me	0	1	2	3	4	P
13. When I begin to sweat in a social situation I fear people will think negatively of me	0	1	2	3	4	S
14. When my thoughts seem to speed up, I worry that I might be going crazy	0	1	2	3	4	C
15. When my throat feels tight, I worry that I could choke to death	0	1	2	3	4	P
16. When I have trouble thinking clearly, I worry that there is something wrong with me	0	1	2	3	4	C
17. I think it would be horrible for me to faint in public	0	1	2	3	4	S
18. When my mind goes blank, I worry that there is something terribly wrong with me	0	1	2	3	4	C

*S = Social subscale; C = Cognitive subscale; P = Physical sub-scale

Appendix F: CFQ

Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.

1	2	3	4	5	6	7
never true	very seldom true	seldom true	sometimes true	frequently true	almost always true	always true

1. My thoughts cause me distress or emotional pain	1	2	3	4	5	6	7
2. I get so caught up in my thoughts that I am unable to do the things that I most want to do	1	2	3	4	5	6	7
3. I over-analyse situations to the point where it's unhelpful to me	1	2	3	4	5	6	7
4. I struggle with my thoughts	1	2	3	4	5	6	7
5. I get upset with myself for having certain thoughts	1	2	3	4	5	6	7
6. I tend to get very entangled in my thoughts	1	2	3	4	5	6	7
7. It's such a struggle to let go of upsetting thoughts even when I know that letting go would be helpful	1	2	3	4	5	6	7

Thank you for completing this questionnaire

Appendix G: BADS-SF

Please read each statement carefully and then circle the number which best describes how much the statement was true for you DURING THE PAST WEEK, INCLUDING TODAY.

	0	1	2	3	4	5	6		
	Not at all		A little		A lot			Completely	
1. There were certain things I needed to do that I didn't do.	0	1	2	3	4	5	6	AC	R
2. I am content with the amount and types of things I did.	0	1	2	3	4	5	6	AC	
3. I engaged in many different activities.	0	1	2	3	4	5	6	AC	
4. I made good decisions about what type of activities and/or situations I put myself in.	0	1	2	3	4	5	6	AC	
5. I was an active person and accomplished the goals I set out to do.	0	1	2	3	4	5	6	AC	
6. Most of what I did was to escape from or avoid something unpleasant.	0	1	2	3	4	5	6	AV	
7. I spent a long time thinking over and over about my problems.	0	1	2	3	4	5	6	AV	R
8. I engaged in activities that would distract me from feeling bad.	0	1	2	3	4	5	6	AV	R
9. I did things that were enjoyable.	0	1	2	3	4	5	6	AC	

*AC = Activation subscale; AV = Avoidance subscale; R = reverse-scored item

Appendix H: DASS21

Please read each statement and circle a number 0, 1, 2, or 3 that indicates how much the statement applied to you *over the past week*. There are no right or wrong answers. Do not spend too much time on any statement.

0	1	2	3	
Did not apply to me at all	Applied to me to some degree, or some of the time	Applied to me a considerable degree, or a good part of the time	Applied to me very much, or most of the time	
1. I found it hard to wind down			0 1 2 3	S
2. I was aware of dryness of my mouth			0 1 2 3	A
3. I couldn't seem to experience any positive feeling at all			0 1 2 3	D
4. I experienced breathing difficulty (e.g., excessively rapid breathing, breathlessness in the absence of physical exertion)			0 1 2 3	A
5. I found it difficult to work up the initiative to do things			0 1 2 3	D
6. I tended to over-react to situations			0 1 2 3	S
7. I experienced trembling (e.g., in the hands)			0 1 2 3	A
8. I felt that I was using a lot of nervous energy			0 1 2 3	S
9. I was worried about situations in which I might panic and make a fool of myself			0 1 2 3	A
10. I felt that I had nothing to look forward to			0 1 2 3	D
11. I found myself getting agitated			0 1 2 3	S
12. I found it difficult to relax			0 1 2 3	S
13. I felt down-hearted and blue			0 1 2 3	D
14. I was intolerant of anything that kept me from getting on with what I was doing			0 1 2 3	S
15. I felt close to panic			0 1 2 3	A
16. I was unable to become enthusiastic about anything			0 1 2 3	D
17. I felt I wasn't worth much as a person			0 1 2 3	D
18. I felt that I was rather touchy			0 1 2 3	S
19. I was aware of the action of my heart in the absence of physical exertion (e.g., sense of heart rate increase, missing a beat)			0 1 2 3	A
20. I felt scared without any good reason			0 1 2 3	A
21. I felt that life was meaningless			0 1 2 3	D

*D = Depression subscale; A = Anxiety subscale; S = Stress subscale