A Primer on Meta-Analytic Structural Equation Modeling: the Case of Depression

Jeffrey C. Valentine[1] · Mike W.-L. Cheung[2] · Eric J. Smith[1] · Olivia Alexander[1] · Jessica M. Hatton[1] · Ryan Y. Hong[2] · Lucas T. Huckaby[1] · Samantha C. Patton[3] · Patrick Pössel[1] · Hayley D. Seely[1]

Abstract

In this paper, we show how the methods of systematic reviewing and meta-analysis can be used in conjunction with structural equation modeling to summarize the results of studies in a way that will facilitate the theory development and testing needed to advance prevention science. We begin with a high-level overview of the considerations that researchers need to address when using meta-analytic structural equation modeling (MASEM) and then discuss a research project that brings together theoretically important cognitive constructs related to depression to (a) show how these constructs are related, (b) test the direct and indirect effects of dysfunctional attitudes on depression, and (c) test the effects of study-level moderating variables. Our results suggest that the indirect effect of dysfunctional attitudes (via negative automatic thinking) on depression is two and a half times larger than the direct effect of dysfunctional attitudes on depression. Of the three study-level moderators tested, only sample recruitment method (clinical vs general vs mixed) yielded different patterns of results. The primary difference observed was that the dysfunctional attitudes→automatic thoughts path was less strong for clinical samples than it was for general and mixed samples. These results illustrate how MASEM can be used to compare theoretically derived models and predictions resulting in a richer understanding of both the empirical results and the theories underlying them.

Keywords Systematic review · Meta-analysis · Meta-analytic structural equation modeling · MASEM

Structural equation modeling is a theory-driven statistical technique that involves analyzing and making sense of the correlations between variables. It can be used to assess the plausibility of different theoretical models, including those that impose a putative causal structure on non-manipulated constructs assessed at a single point in time and non-manipulated constructs assessed at multiple points in time, as well as the effects of manipulating a variable (such as in intervention research), and both mediating and moderating variables. The focus of this paper is on meta-analytic structural equation modeling (MASEM), which, as its name suggests, is a statistical technique that brings meta-analysis

and structural equation modeling together (Becker, 1992, 1995; Cheung & Chan, 2005; Cheung, 2015a, 2019; Jak & Cheung, 2020). It allows researchers to investigate how empirically distinct theoretically related constructs might be and whether the structural model serves as a reasonable way of organizing the underlying constructs and, hence, is well-suited to addressing topics of interest to both theoretically oriented and applied prevention scientists. While both meta-analysis and structural equation modeling are regularly used in prevention science, they are rarely applied in combination in the prevention science literature (the first such application in *Prevention Science* is Shen et al., 2021). We hope that, by providing a primer on MASEM, prevention scientists will make greater use of these methods in the future. In addition to this paper, good conceptual introductions to MASEM can be found in Cheung and Hafdahl (2016) and Cheung (2020).

MASEM usually involves two basic steps. First, the relationships observed in the studies in the meta-analytic database are used to create a meta-analytic (pooled) correlation matrix. This is the primary distinction between MASEM

✉ Jeffrey C. Valentine
  jeff.valentine@louisville.edu

1 College of Education and Human Development, University of Louisville, Louisville, USA

2 Department of Psychology, National University Singapore, Singapore, Singapore

3 School of Medicine, Emory University, Atlanta, USA

and a structural equation model that is based on the data observed in a single study. The average correlation matrix with its sampling covariance matrix can be analyzed with a weighted least square estimation method. After arriving at the meta-analytic correlation matrix, structural equation models are fitted to the pooled correlation matrix. Like all structural equation models, test statistics and goodness-of-fit indices can be used to help judge the exact and approximate fit of the proposed models.

To create and interpret MASEM models, then, researchers need familiarity with the methods of systematic reviewing, meta-analysis, and structural equation modeling. Fortunately, researchers with a good familiarity with structural equation modeling will transition easily to applying that knowledge to MASEM. Both the random-effects two-stage structural equation modeling (TSSEM) approach proposed by Cheung (2014, 2015a) and one-stage structural equation modeling (OSMASEM; Jak & Cheung, 2020), which are introduced in this paper, use structural equation models to meta-analyze correlation matrices and fit structural equation models. Therefore, model fit indices and parameter estimates in MASEM can be interpreted similarly as those in SEM. That said, there are some practical differences between SEM and MASEM. For example, applications of SEM usually involve models with latent variables with large degrees of freedom (*df*s), whereas regression or path models with small *df*s are fit in MASEM. This point is important because some fit indices, for example, the root mean square error of approximation (RMSEA), may not behave well in models with small *df*s (Kenny et al., 2015). Researchers in both SEM and MASEM should be cautious in these cases.

An Overview of the Order of Operations for MASEM

Researchers interested in using MASEM should begin by formulating a precise research question, or set of research questions, that can be addressed using the technique. Then, researchers should set the Participant, Intervention, Comparison, Outcome, Study designs (PICOS) criteria, that is, arrive at theoretical and operational definitions of the relevant participants, interventions (if applicable), comparisons (if applicable), outcome measures used, and study designs. The PICOS criteria will inform the systematic review's literature search. A structured process that yields reproducible results should be used to identify and extract information from eligible studies.

The five primary considerations for arriving at the meta-analytic (pooled) correlation matrix are (a) whether to use a fixed effects or a random effects approach, the considerations for which are the same in MASEM as they are in any meta-analysis, (b) how to handle study dependence (e.g., if a study presents multiple outcome estimates for the same construct), (c) how to address possible publication bias, (d) whether to correct the estimates for attenuation, and (e) how to incorporate judgments of study quality. We address each of these briefly in turn.

Fixed Effects vs Random Effects　The fixed effects model is defensible if it is reasonable to believe that all study effects are estimating the same (or common) population parameter or if all study-level influences are known and can be accounted for (e.g., via regression adjustment). In either case, this assumption implies that sampling error is the only reason that the observed effect varies across studies. If, on the other hand, researchers believe that study effects will vary across samples, setting, and study characteristics, and that some of these characteristics are unknown and cannot be accounted for, then the random effects model will generally be more appropriate. We should note that the fixed effects model is also used when the researchers' interest is limited to the studies at hand (i.e., they want to generalize only to the specific studies in the meta-analytic database and studies highly like them), something we suspect will be rare in MASEM applications.

There are two random-effects models in MASEM. Conventionally, we may fit a structural equation model on the average correlation matrix, whereas the random effects in a study represent the differences between the population correlation matrix in that study and the average correlation matrix. This is the model we use in the present study. An alternative model is to treat the random effects in a study as the differences between the population structural parameters, e.g., regression coefficients and factor loadings, in that study and the average structural parameters. Readers interested in these issues may refer to Cheung and Cheung (2016) and Ke et al. (2019) for the discussion.

Dependent Estimates　The estimates contributing to a meta-analytic correlation matrix are presumed to be independent. Non-independence can happen for many different reasons. For example, several studies included in the meta-analytic database we present below used multiple measures of depression and therefore had more than one estimate of the correlation between depression and another construct of interest. If the estimates contributing the meta-analytic correlation are not independent, then the meta-analytic standard errors will be too small, and studies contributing multiple effect sizes will be given too much weight in the analysis. There are two ways of addressing dependence. One approach is structural, in the sense that researchers can make intentional choices to limit dependence. In the case of having multiple measures of the same construct, for example, the research team might have a preference for one measure over another or might drop measures randomly until only

one remains (see Lipsey & Wilson, 2001). Of course, these kinds of decisions should not be contingent on the effect size observed and ideally will be articulated in a publicly available protocol that was created prior to data collection. The other approach is statistical. Common statistical methods for addressing dependence are multilevel modeling (see, for example, Konstantopoulos, 2011), robust variance estimation (see, for example, Tanner-Smith & Tipton, 2014), and averaging dependent effects (Cooper, 2017). For MASEM applications, we recommend averaging dependent effect sizes within studies because it provides a framework for testing moderators in MASEM. The average correlation matrix with its asymptotic sampling covariance matrix from multilevel modeling or robust variance estimation can be used to fit structural equation models with weighted least squares as the estimation method.

Publication Bias Sometimes, the decision about whether to publish a study depends on the nature of its findings. When a study goes unpublished because it does not have statistically significant findings on its main outcomes, this is known as publication bias, and it can represent an important threat to the validity of the conclusions arising from a systematic review and meta-analysis. In general, publication bias will tend to result in estimated relationships between constructs that are too large and heterogeneity estimates that are too small. In the context of MASEM, this means that the path coefficients will be too large in the presence of publication bias. Vevea et al. (2019) provide a good discussion of the various options for detecting the possible presence of publication bias and diagnosing the extent of its possible impact.

Prevention is always the best defense against publication bias (Jak & Cheung, 2020; Vevea et al., 2019), meaning that researchers should carry out an exhaustive search for studies relevant to their research question. This feature is always an important part of a systematic review, but it assumes even greater importance in a MASEM given that, to date, no consensus has emerged on how to address possible publication bias in MASEM. If researchers carrying out a MASEM project conclude that publication bias has an important impact on one or more of their paths of interest, then it might make sense to redouble efforts aimed at locating relevant unpublished studies. If the researchers choose to proceed with MASEM despite the possible influence of publication bias, they should clearly communicate the rationale for this decision to their readers, and should be very cautious when interpreting results.

Attenuation Corrections If researchers have item-level correlation matrices, MASEM can handle measurement error by fitting confirmatory factor analytic models or structural equation models. When path models on the composite scores are fitted in MASEM, problems like measurement error and artificial dichotomization of continuous variables reduce the correlation that will be observed between two variables. For example, even if two constructs are perfectly correlated, their observed correlation will not be 1.0 in the presence of measurement error. This phenomenon is known as attenuation, and it is not uncommon for researchers to correct correlations for attenuation caused by measurement error and other problems in meta-analysis (in fact, doing so is the norm in some fields, organizational psychology being one example). If attenuation corrections for measurement error are desirable, one way to implement these would be to rely on good systematic reviews and meta-analyses of coefficient alpha for measures of relevant constructs (these are often referred to as reliability generalization studies; Rodriguez & Maeda, 2006). Another would be to meta-analyze the reliability coefficients present in the studies in the meta-analysis. Schmidt et al. (2019) have a very good discussion of how this could be done. At the same time, users should be aware that attenuation corrections can sometimes lead to a meta-analytic correlation matrix that is non-positive definite (e.g., Charles, 2005). Furthermore, Michel et al. (2011) found that results of uncorrected and corrected for unreliability were similar in several published datasets. However, it remains unclear whether this finding is supported by simulation studies.

There are a couple of reasons for a non-positive definite matrix in MASEM, especially when using attenuation corrections. One is that the individual correlation matrices are not positive definite because pairwise deletion was used to calculate the primary correlation matrices (Wothke, 1993) or the corrected correlation was larger than one. Researchers should therefore check whether the correlation matrices are positive definite. The provided R code illustrates how to do this. Any non-positive definite matrices should be excluded from the analysis. Also, the average correlation matrix may be non-positive definite. Based on our experience, this is rare. If it happens, however, researchers should identify and exclude the variable causing it to happen. An alternative approach is to regularize the average correlation matrix using ridge SEM (e.g., Yuan et al., 2011). Finally, if the number of elements in the variance component is large comparing to the number of studies, it is difficult to estimate the full variance component. A common practice is to estimate the variances of the random effects, wherein the covariances are fixed at zero.

Incorporating Judgments of Study Quality There is widespread recognition that study quality, which in the context of a systematic review and meta-analysis can be defined as the extent to which the methods deployed to answer a research question fit the goals of the review (Valentine, 2019), can have an influence on study effect sizes. Therefore, it is

common for experts to recommend that study quality indicators be assessed and that judgments about study quality be incorporated into the review. For example, the AMSTAR-2 checklist holds that systematic reviews ought to assess the impact of varying study quality on meta-analytic results (Shea et al., 2017).

Assessing study quality is perhaps the most challenging aspect of conducting a systematic review, in part because the important study quality indicators vary as a function of study design — the issues relevant to a randomized experiment are not exactly the same as the issues involved in a cross-sectional investigation of a pattern of relationships. Important study quality indicators almost certainly also vary as a function of study context. For example, participant attrition is probably a more severe threat to the inferences arising from targeted prevention efforts than to those arising from universal prevention efforts.

In general, we recommend that individuals carrying out a MASEM project identify the relevant study quality indicators for their research question and then adopt a two-stage approach: Use the most important of these as inclusion criteria and then treat others empirically. Ideally, with a sufficient number of studies, the observed effect sizes could be adjusted (via meta-regression) in an attempt to take into account the impact of varying study quality on the meta-analytic results. A general approach for identifying relevant study quality indicators for a particular research question can be found in Valentine (2019).

Our Motivating Example: Cognitive Theories of Depression

One of the major cognitive theories explaining depression is Beck's cognitive theory of depression (1976). One reason for its status as a major theory is that it is the basis of one of the most effective approaches to treat (Butler et al., 2006) and prevent (e.g., Pössel et al., 2018) depression. Beck's theory includes multiple cognitive elements including that the relationship between dysfunctional attitudes and depression is mediated by negative automatic thinking. However, while the elements themselves are well established, their empirical relationships with each other and with depression

are less clear (e.g., Pössel & Black, 2014). Drawing on data collected for a larger project that examines the relationships between constructs from multiple cognitive models of depression (i.e., Beck's cognitive theory, 1976; the Hopelessness Model, Abramson et al., 1989; Response Style Theory, Nolen-Hoeksema, 1991) and personality traits known to be associated with depression, in this tutorial, we test two effects to illustrate MASEM: the direct effect of dysfunctional attitudes on depression and the extent to which negative automatic thoughts serve as a mediating effect of the relationship between dysfunctional attitudes and depression. We also test three theoretically relevant study-level moderators: sample age (children vs adults), sample recruitment (general vs mixed vs clinical), and the cultural context in which the study took place (North America vs other locations).

Method

The Literature Search

Two experts in theories of depression, supported by several students in psychology programs, developed a list of terms designed to capture constructs related to cognitive models of depression. We focus on three for this tutorial: depression, depressive schema/dysfunctional attitudes, and automatic thinking. We searched Medline, PsycINFO, and the Psychological and Behavioral Sciences Collection simultaneously in EBSCO, as well as ProQuest Dissertations and Theses. As shown in Table 1, we searched document abstracts for at least two constructs of interest (e.g., depression and automatic thoughts) and the full text of documents for terms suggesting that a correlation matrix might be available (i.e., matrix or correlation* or intercorrelation*).

This latter point merits further discussion. We focused on studies reporting a full correlation matrix for their measured variables not only in part for efficiency's sake (it is easier to search for a correlation matrix than for individually reported correlations) but also in part to reduce our exposure to outcome reporting bias (Pigott et al., 2013). This refers to within-study selective reporting of findings that are statistically significant. Outcome reporting bias has the potential to inflate effect size estimates in a meta-analysis. Fortunately,

Table 1 Search strategy for EBSCOhost (tutorial studies only)

Construct	Search strategy
Depression	AB (depress* not depressing) or dysthymi* or dysphori*
Dysfunctional attitude/dysfunctional schema	AB "dysfunctional attitude*" or ((maladaptive or depressive or depressogenic or cognitive) n1 schema*)
Automatic thoughts	AB "automatic thought*"
Correlation	Matrix or correlation* or intercorrelation*

most researchers create correlation matrices that report the intercorrelations between all of their measured variables and not just the statistically significant ones. Therefore, relative to using individually reported correlations, which are potentially selectively reported, relying on a full correlation matrix likely reduces the impact of outcome reporting bias on our results.

The search generated more than 3200 documents that were potentially eligible for review. Our goal for this stage was to identify documents clearly suggesting that at least two constructs of interest were *not* measured (i.e., we sought to eliminate all clearly irrelevant documents). To this end, we used three pilot rounds of testing (during which all screeners coded either 99 or 100 articles). During these pilot rounds, we tracked agreement and identified sources of disagreement. After the pilot rounds, two researchers working independently screened each study. All disagreements were resolved by one of the depression experts. We used abstrackr to manage the screening process (Wallace et al., 2012). Over 2700 documents passed this stage of screening.

We then attempted to obtain the full text of the 2770 studies that remained after eliminating the clearly not relevant documents. We were successful in obtaining 2424 of these. We proceeded to full-text screening, which involved a single researcher examining the full text of each document to determine (a) whether at least two constructs of interest were assessed in the study and (b) whether the study reported a full correlation matrix. We also screened out studies that included an experimental manipulation relevant to depression and reported only post-manipulation measures (pre-manipulation correlations were eligible for review). This process led to 161 documents that were provisionally eligible to be reviewed for this tutorial. Ultimately, after two researchers independently examined the full text of these 161 documents, 104 studies were included in the data on which this tutorial is based. By far, the most common reason studies dropped out at this stage was that at least one of the measures used did not pass expert screening (this was true for 41 of the 57 studies that did not pass full text screening). The next most common reason was that the study did not report a full correlation matrix ($k = 8$). Two studies assessed constructs of interest after a relevant experimental manipulation, one study was a duplicate, and another is awaiting translation. The remaining four studies had either unusable data (e.g., correlations based on change scores) or ambiguities that prevented us from using the data (e.g., inconsistent values reported for the data of interest). The references for the 104 included studies can be found in the Appendix.

Coding Studies

Studies were coded by two researchers working independently. One of us served as the first coder for all of the studies, supported by one of five additional coders. One of our experts in depression served to arbitrate disagreements. In addition to coding the relevant correlations of interest, we coded several study-level characteristics, including the year the study was published or released, sample characteristics (e.g., age, clinical sample vs not, reported race/ethnicity), and measure characteristics (e.g., measure name, reported reliability estimates). Overall, we coded 288 individual correlations, from 125 independent samples nested in 104 studies, that contributed to the data used in this tutorial. The number of coded findings in the studies ranged from 1 to 24.

With respect to assessing study quality, the ideal study for our research questions would involve giving a large probability sample perfectly reliable and valid measures of our constructs of interest. Because we (accurately) anticipated that no studies would be based on a probability sample, we did not code for this characteristic. We did consider coding for any validity evidence reported in a study that was based on the study's sample, but pilot testing indicated that most studies did not provide any information on validity apart from (a) mentioning validity work done in another study, (b) providing evidence of convergent validity by using multiple measures of the same construct, and (c) providing evidence of discriminant validity by providing correlations across different constructs. Therefore, the primary study quality characteristic that varied across the studies in this review was score reliability. When available, the sample coefficient alpha for each measure is reported in Table 2. Most of the measures used in the studies included in the meta-analysis will be recognized by scholars familiar with cognitive models of depression, adding a degree of face validity to our results, though it should also be noted that researchers occasionally revised existing measures by selecting particular items for use.

Participant Characteristics

As can be seen in Table 2, most studies included in the data supporting this tutorial were based on adults (88 of 104), which we defined as having an average sample age between 18 and 65. Only two were based on elderly adults (65 and older). The remaining studies were based on children. Most studies were conducted in the USA (78) or Canada (14). Twenty-seven studies recruited participants from a clinical sample (e.g., an inpatient hospital), while 68 studies recruited from a more general location, and 13 recruited from both clinical and non-clinical sources. Three studies recruited from both clinical and non-clinical sources and reported separate correlations by recruitment source — these studies are included twice in the counts above. For studies taking place in the USA, the median of the reported percentages of white participants was 75% (the range was 0% to 90 + %). Within studies, sample sizes ranged from 10 to

Table 2 Study characteristics

First author year	Location	n	Age	Recruitment	Depression measure (alpha)	Automatic thoughts measure (alpha)	Depressive schema measure (alpha)
Ahnberg 1997	Canada	81	20	Clinical	BDI-II	NA	DAS SF (author might have shortened, pp. 15–16)
Anthony 2001	Canada	150	19	General	MASQ (general distress: depression subscale) (.92)	ATQ (positive items) (.95) and ATQ (negative items) (.96)	NA
Arpin-Cribbie 2002	Canada	300	20	General	CES-D (.85)	ATQ (.97)	DAS
Asdigian 1993	USA	226	Adults	General	BDI (.82)	NA	DAS (.79)
Baggett 1994	USA	65	44	Clinical	HRSD	ATQ-R (ratio of positive to negative thoughts)	NA
Batmaz 2016	Turkey	885	35	Clinical	BDI	ATQ	DAS (revised and abbreviated form) — Turkish version (.84)
Batmaz 2015	Turkey	300	38	Mixed	BDI (Turkish version)	ATQ (.90)	DAS (.83)
Beevers 2004	USA	121	38	Clinical	BDI	NA	DAS (perfectionism) and DAS (need for approval)
Berle 2012	Australia	134	19	General	BDI-II and Nepean Dysphoria Scale (total score) (.91)	NA	DAS
Bernholtz 2013	Canada	54	42	Clinical	BDI and SCL-90 R (depression subscale)	NA	DAS
Bianchi 2016	USA	1063	43	General	PHQ-9 (.87)	NA	DAS SF 1 (.85)
Bianchi 2016	USA	323	43	General	PHQ-9 (.88)	NA	DAS SF 1 (.84)
Bianchi 2016	NZ	184	43	General	PHQ-9	NA	DAS SF 1
Block 1988	USA	222	19	Mixed	BDI	NA	DAS
Bond 1984	USA	101	19	General	BDI	NA	DAS
Braithwaite 2003	USA	200	20	General	BDI-II (.88)	ATQ-SF (.91)	NA
Brinker 2007	Canada	118	19	General	BDI-II (.88)	ATQ (.95)	NA
Brinker 2007	Canada	309	19	General	BDI-II (.88)	ATQ (.95)	NA
Buschmann 2016	USA	539	Adults	General	DASS-21 — Depression subscale	NA	DAS
Choon 2015	Malaysia	294	15	Clinical	CES-D (.79)	ATQ (.83)	NA
Church 2005	Australia	406	29	General	EPDS (.87)	NA	DAS (.90)
Ciesla 2007	USA	133	Adults	General	BDI-II (.90) and MAACL-R (depression items) (.94)	NA	DAS (.92)
Ciesla 2007	USA	126	Adults	General	BDI-II (.92) and MAACL-R (depression items) (.94)	NA	DAS (.90)
Clark 2002	USA	291	34	Mixed	HRSD (.79)	NA	DAS Form A (.94)
Cowden 1992	USA	256	19	General	BDI	NA	DAS
Cox 2011	USA	164	19	General	CES-D (.89)	NA	DAS (.90)
Cui 2012	China	613	14	General	CES-D (.91)	ATQ (.93)	DAS (.83)
D'Agostino 2016	Italy	132	21	General	BDI-II (Italian) and Nepean Dysphoria Scale (Italian) (.95)	NA	DAS-A (Italian)

Table 2 (continued)

First author year	Location	n	Age	Recruitment	Depression measure (alpha)	Automatic thoughts measure (alpha)	Depressive schema measure (alpha)
d'Alessandro 2004	Canada	453	12	General	CDI (.84)	NA	CADAS (.87)
d'Alessandro 2004	Canada	241	11	General	CDI (.88)	NA	CADAS (.85)
Du 2015	China	298	20	General	Zung SRDS (.81)	ATQ (.95)	NA
Dunkley 2008	Canada	223	40	General	BDI (.88)	NA	DAS — self-critical perfectionism items (.88) and DAS — adapted (referred to as self-critical/defensive separation items)
Dyer 1991	USA	103	45	General	NA	ATQ	DAS Form A
Ekstrom 2003	USA	143	Adults	General	MAACL (depression subscale) (.94)	NA	DAS – Expanded (.95)
Finer-Greenberg 1987	USA	108	Adults	General	SCL-90-R (depression subscale) (.90)	CCL (depression-related thoughts subscale) (.92)	NA
Fischer 2009	USA	280	11	General	CDI (.91)	CATS (personal failure subscale) (.83)	NA
Frantom 1994	USA	98	22	General	BDI	NA	DAS Form A
Gerstein 2010	USA	217	20	Mixed	SADS (lifetime) — days spent in a depressive episode/total days in study	NA	DAS Form A (performance evaluation) and DAS Form A (need for approval)
Grant 2013	USA	345	19	General	BDI	NA	DAS
Gül et al. 2015	Turkey	53	33	Clinical	BDI — Turkish version	ATQ — Turkish version	NA
Gunter 2008	USA	10	22	Mixed	BDI-II	ATQ	DAS Form A
Haeffel 2005	USA	237	19	General	BDI and MASQ (anhedonic depressive symptoms subscale)	NA	DAS (.91)
Haeffel 2005	USA	251	19	General	BDI and MASQ (anhedonic depressive symptoms subscale)	NA	DAS (.90)
Hamill Skoch 2011	USA	127	13	General	CDI (.91) and mini MASQ (anhedonic items) (.78)	NA	DAS-C (.91)
Hamill Skoch 2011	USA	197	15	General	CDI (.91) and mini MASQ (anhedonic items) (.78)	NA	DAS-C (.91)
Hamilton 1982	USA	20	40	General	BDI	NA	DAS
Hamilton 1982	USA	20	37	Clinical	BDI	NA	DAS
Hamilton 1982	USA	20	39	Clinical	BDI	NA	DAS
Hankin 2011	USA	223	Adults	General	BDI (.88), HDSQ (.97), MASQ (general distress/depression subscale) (.92), and MASQ (anhedonic depression subscale) (.92)	CCL-D (.91)	DAS (.89)
Healy 1991	USA	105	20	General	BDI (.88)	NA	DAS (.89)
Heinz et al. 2009	USA	182	19	General	IDD (.91)	ATQ (.92)	DAS (.92)
Hiramura 2008	Japan	546	19	General	Zung SRDS	ATQ-R	NA
Holden 2012	USA	63	45	General	BDI-II (.91)	ATQ	NA

Table 2 (continued)

First author year	Location	n	Age	Recruitment	Depression measure (alpha)	Automatic thoughts measure (alpha)	Depressive schema measure (alpha)
Hughes 2008	USA	364	19	General	BDI (.90), MASQ (anhedonic depression subscale) (.92), and MASQ (general distress/depression subscale) (.93)	CCL (depression) (.94)	NA
Hundt 2011	USA	299	20	General	BDI (.90) and CES-D (.89)	ATQ (positive items) (.96) and ATQ (negative items) (.96)	NA
Hyers 1995	USA	359	73	General	CES-D (.85)	NA	DAS Form A (.88)
Ito 2005	Japan	191	19	General	Zung SRDS — Japanese version	NA	DAS Form A — Japanese version
Jarrett 1983	USA	37	37	Clinical	BDI, DACL, and MMPI (depression subscale)	ATQ (frequency items) and ATQ (beliefs items)	NA
Karatepe 2013	Turkey	262	Adults	General	BDI	ATQ	NA
Keller 2010	Germany	266	15	General	Allgemeine Depressions-Skala	NA	DAS
Kelly 2005	USA	60	19	Clinical	BDI-II	NA	DAS
Knowles 2005	UK	528	19	General	BDI	NA	DAS
Kohlbeck 1995	USA	68	38	Mixed	BDI	ATQ	DAS
Konecky 1996	USA	23	Adults	Clinical	BDI	NA	DAS
Konecky 1996	USA	17	Adults	General	BDI	NA	DAS
LaGrange 2008	USA	161	8	General	CDI (.73–.88) and CESD-C (.81–87)	ATQ (.89–.96)	NA
LaGrange 2008	USA	173	9	General	CDI (.73–.88) and CESD-C (.81–87)	ATQ (.89–.96)	NA
LaGrange 2008	USA	164	10	General	CDI (.73–.88) and CESD-C (.81–87)	ATQ (.89–.96)	NA
LaGrange 2008	USA	149	11	General	CDI (.73–.88) and CESD-C (.81–87)	ATQ (.89–.96)	NA
Lam 2015	Singapore	194	49	General	HADS (depression) (.78)	NA	DAS Form A (.87)
Lee 2007	USA	152	21	General	BDI-II (.91)	NA	DAS (dysfunctional achievement attitude items) (.87)
Lee 2007	USA	135	21	General	BDI-II (.91)	NA	DAS (dysfunctional achievement attitudes items) (.87)
Leucht 1997	Germany	42	52	Clinical	BDI	Automatische Gedanken	Skala dysfunktionaler Einstellungen
Lewis 2011	USA	222	28	Clinical	BDI-I	NA	DAS Form A (total score); DAS Form A (achievement), and DAS Form A (interpersonal)
Linaman 1996	USA	166	40	Mixed	BDI and CES-D	ATQ (positive) and ATQ (negative)	DAS Form A
Lipinski 2015	USA	332	Children	Clinical	CDRS — Revised (.70)	NA	DAS
Lipinski 2015	USA	96	Children	Clinical	CDRS — Revised (.71)	NA	DAS
Lutz 1984	USA	37	Adults	General	BDI and DACL	NA	DAS
Lutz 1984	USA	36	Adults	Clinical	BDI and DACL	NA	DAS
Marčinko 2013	Croatia	234	44	Clinical	DASS — Depression subscale (.91)	NA	DAS (perfectionism subscale) (.82) and DAS (dependency subscale) (.74)

Table 2 (continued)

First author year	Location	n	Age	Recruitment	Depression measure (alpha)	Automatic thoughts measure (alpha)	Depressive schema measure (alpha)
Margolis 1982	USA	54	45	Clinical	BDI and MAACL (depression subscale)	NA	DAS
Mayer 2004	USA	80	24	General	BDI-II (.83) and HDSQ (.90)	NA	DAS (autonomy items) (.92) and DAS (sociotropy items) (.80)
McCarron 1980	USA	19	68	General	BDI and HRSD	NA	DAS
Morris 2014	USA	68	24	General	BDI-II (.85)	NA	DAS (.78)
Noh 2016	Korea	159	22	General	CES-D (Korean translation) (.92)	NA	DAS (Korean translation) (.91)
Palmer 2015	UK	177	40	General	PHQ-8 (.92)	NA	DAS SF 2 (.88)
Pedrelli 2006	USA	59	52	Clinical	CDSS and HRDS	NA	DAS
Philpot 1995	USA	145	26	General	BDI	ATQ-R (positive), ATQ-R (negative), and ATQ (ratio of positive to negative thoughts)	NA
Pirbaglou 2013	Canada	992	21	General	CES-D (.79)	ATQ (.97)	NA
Poppleton 2008	USA	22	33	Clinical	SCL-90 (depression subscale)	ATQ	DAS
Quiggle 2001	USA	240	39	General	BDI (.91) and BSI (depression subscale) (.91)	NA	DAS (.90)
Randolph 1998	USA	246	20	General	BDI (.85) and Depression Proneness Rating Scale (.92)	NA	DAS Form A (.92)
Raynor 1988	USA	19	15	Clinical	BDI and Behavior Inventory for Depressed Adolescents	NA	DAS Form A
Reeder 2014	USA	145	19	General	BDI-II	ATQ-30	NA
Robinson 1987	Canada	94	40	Mixed	BDI	ATQ	DAS
Ross 1986	USA	58	Adults	Clinical	BDI and MMPI (depression subscale)	ATQ-30	NA
Ross 1986	USA	20	Adults	Clinical	BDI and MMPI (depression subscale)	ATQ-30	NA
Ross 1986	USA	21	Adults	Clinical	BDI and MMPI (depression subscale)	ATQ-30	NA
Ross 1986	USA	23	Adults	Clinical	BDI and MMPI (depression subscale)	ATQ-30	NA
Rough 2016	USA	64	46	Mixed	BDI-II (.91–.97) and Structured Interview Guide (HDRS, Seasonal Affective Disorder version) (.63–.91)	NA	DAS (.92–.93)
Russell 1983	USA	108	37	Clinical	BDI (autonomy) (probably a selection of items, p. 68) and BDI (sociotropy) (probably a selection of items, p. 68)	NA	DAS (achievement), DAS (perfectionism), DAS (approval), DAS (love) (all were possibly a selection of items, p. 70)
Sadowski 1994	USA	118	15	Mixed	Reynolds Adolescent Depression Scale	ATQ	NA
Schatz 2009	USA	24	40	Clinical	BDI	ATQ	NA
Scherrer 2009	Canada	33	39	Mixed	BDI-II	ATQ-R (positive), ATQ-R (negative), and CCL (depressive thoughts items)	DAS
Scherrer 2009	Canada	22	32	Mixed	BDI-II	ATQ-R (positive), ATQ-R (negative), and CCL (depressive thoughts items)	DAS
Schroeder 1995	USA	258	21	General	BDI	ATQ	DAS

Table 2 (continued)

First author year	Location	n	Age	Recruitment	Depression measure (alpha)	Automatic thoughts measure (alpha)	Depressive schema measure (alpha)
Schwartz-Stav 2006	Israel	32	18	Clinical	BDI and CDSS (.79)	CCL (.92)	NA
Shapira 2012	Multiple	578	33	General	CES-D (.93)	NA	DAS (dependent and self-critical attitudes items only) (.85)
Sommer 1999	USA	49	15	Clinical	K-SADS-P (depression items) (.91)	ATQ-C (.96)	NA
Stanley 1991	USA	40	Adults	General	NA	ATQ-R	DAS Form A
Stevens 2006	USA	95	15	Mixed	BDI	NA	DAS
Stolow 2011	China	662	20	General	CES-D (translated for this study) (.90–.94)	NA	DAS Form A (translated for this study) (.88–.89)
Sutton 2011	USA	550	17	General	IDD (.88), MASQ (general depression subscale) (> .84), MASQ (anhedonic depression subscale) (> .84)	NA	DAS (.95)
Szymanski 1997	USA	249	19	General	BDI, CES-D, and Zung SRDS	ATQ	DAS
Tajima 2007	Japan	248	22	General	BDI-II	ATQ-R (total)	DAS-24 (Japanese translation) (.89)
Thimm 2013	Norway	149	29	Mixed	BDI	NA	DAS Form A
Turkoglu 2015	Turkey	70	34	Clinical	BDI	NA	DAS
Vatanasin 2012	Thailand	800	17	General	CES-D (Thai translation)	ATQ (Thai translation)	NA
Wierzbicki 1989	USA	143	Adults	General	BDI	ATQ	DAS
Wierzbicki 1989	USA	57	Adults	Clinical	BDI	ATQ	DAS
Wong 2008	Hong Kong	868	35	General	BDI (.91) and DASS-21 (Depression subscale) (.89)	ATQ (.96)	NA
Wong 2008	China	1001	34	General	BDI (.85) and DASS-21 (Depression subscale) (.84)	ATQ (.94)	NA
Yilmaz 2015	Turkey	236	23	General	BDI	NA	DAS (.88)
You 2007	USA	306	19	General	BDI-II (.91) and IDD (lifetime version) (.92)	NA	DAS (.90)
Zheng 2014	China	535	19	General	CES-D	NA	DAS

ATQ automatic thoughts questionnaire, *ATQ-R* automatic thoughts questionnaire-revised, *BDI* beck depression inventory, *CATS* children's automatic thoughts scale, *CCL* cognition checklist, *CDI* children's depression inventory, *CDSS* calgary depression scale for schizophrenia, *CES-D* center for epidemiologic studies — depression, *CDRS* children's depression rating scale, *DACL* depression adjective checklist, *DAS* dysfunctional attitudes scale, *DAS SF* dysfunctional attitudes scale short form, *DASS-21* depression and anxiety stress scale 21, *EPDS* edinburgh postnatal depression scale, *HADS* hospital anxiety and depression scale, *HDSQ* hopelessness depression symptom questionnaire, *HRSD* hamilton rating scale for depression, *IDD* inventory to diagnose depression, *K-SADS-P* schedule for affective disorders for schizophrenia for school-age children — present episodes, *MAACL-R* multiple affect adjective checklist — revised, *MASQ* mood and anxiety symptom questionnaire, *MMPI* Minnesota multiphasic personality inventory, *NA* construct not measured in the study, *PHQ* patient health questionnaire, *SCL-90 R* symptom checklist 90 revised, *SADS* seasonal affective disorder scale, *Zung SRDS* Zung self-rating depression scale

1063, with a median of 143 and the middle 50% of the sample size distribution ranging from 57 to 241.

Measure Characteristics

With respect to measurement, the BDI I and II were the most used in the studies in this review — their usage was more than five times greater than the Center for Epidemiologic Studies — Depression subscale, the next most used measure. In all, 24 unique measures of depression were used in this review. The Automatic Thoughts Questionnaire was, by far, the most commonly used measure of automatic thoughts. In some cases, authors assessed positive automatic thinking (which should be negatively correlated with depression) — for these, we reversed the sign of the reported correlation coefficient. The Dysfunctional Attitudes Scale was the only measure of dysfunctional attitudes used in the studies in this tutorial (though studies differed in the specific subscales and items used).

Statistical Models

The random-effects TSSEM and OSMASEM approaches were used to conduct the MASEM. Correlation coefficients are pooled in a multivariate way by taking their dependence (sampling covariances) into account. Specifically, given the diversity of the studies included in this review, including sample and measure characteristics, we could not support the assertion that the studies share a common effect size, and therefore, the correlation matrices were combined with a random-effects model in the first stage of analysis. To address non-independence, when a study provided multiple estimates for the same construct pair, we averaged the correlations within studies. We chose this method because (a) dependencies arose due to having multiple estimates of the same construct pair with equal or nearly equal sample sizes and (b) it minimized measurement errors by including all relevant effect sizes. Although a multilevel model or robust variance estimation might also be used, one major limitation was that these models do not allow including study characteristics to moderate the structural parameters.

As can be seen in Table 2, about 60% of the studies in our meta-analytic dataset provided a sample-based coefficient alpha for at least one measure. The coefficient alphas were quite high. For depression, the median coefficient alpha was .90 and the middle 50% of the distribution ranged from .85 to .91. For automatic thoughts, the median coefficient alpha was .94 and the middle 50% of the distribution ranged from .92 to .96. For depressive schema, the median coefficient alpha was 0.91 and the middle 50% of the distribution also ranged from .85 to .91. In part because attenuation corrections for measurement error would have little impact given the typically high reliability estimates observed in these studies, we opted not to perform this correction.

An average correlation matrix and the variance component of the heterogeneity variances were estimated with the full information maximum likelihood estimation method (FIML), which is unbiased and efficient in handling data missing completely at random and missing at random (Enders, 2010). FIML is preferable to other methods such as listwise and pairwise deletion in handling missing data in MASEM (Cheung, 2015a). If a particular correlation of interest is not present in any of the included studies, then the associated variable has to be dropped from analysis.

Both the TSSEM and OSMASEM fit the structural equation models using the precision of the correlation matrices. Therefore, the study sample sizes do not affect the parameter estimates, their standard errors, and the fit statistics of the proposed model once the asymptotic sampling covariance matrices of the correlations are calculated. The study sample sizes do slightly affect some of the approximate fit indices, e.g., RMSEA, which depend on the sample size used to fit the structural model. In TSSEM and OSMASEM, the sum of the sample sizes is used as the sample size in calculating fit indices. This follows the practice of how sample size is calculated in multiple-group SEM.

The average correlation matrix and its asymptotic sampling covariance matrix were used to fit structural equation models in the second stage of analysis in TSSEM. Figure 1
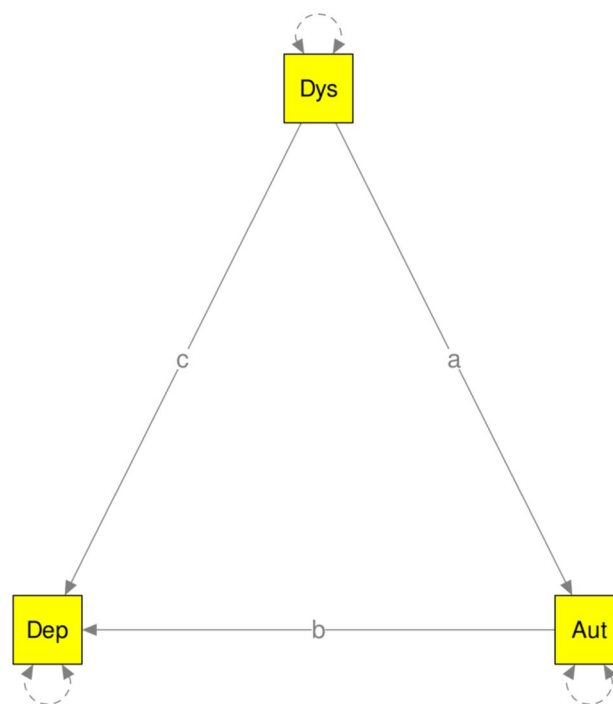
Fig. 1 Proposed mediation model. Dys, dysfunctional attitudes; Aut, automatic thoughts; Dep, depression. Path c represents the hypothesized direct effect of dysfunctional attitudes on depression. The indirect effect of dysfunctional attitudes on depression via automatic thoughts is represented by the $a \times b$ pathway

displays the proposed mediation model. The corresponding path coefficients are labelled as *a*, *b*, and *c* for ease of reference. As is usual in SEM, the indirect effect is estimated by the product term *a*b*. Conventional confidence intervals, also known as Wald confidence intervals, are usually based on parameter estimate ± 1.96*standard error. The sampling distribution of an indirect effect is complicated and tends not to be normal (MacKinnon et al., 2004), making the Wald confidence intervals less appropriate. Therefore, following the advice provided in Cheung (2009), in the second stage of the analysis, we report the 95% likelihood-based confidence interval (LBCI), which captures the non-normal distribution of the parameters. The main drawback to this approach is that it is computationally intensive. It is therefore advisable to use the LBCI only when researchers are interested in the functions of the parameters, e.g., the indirect effect in this paper.

Conventional MASEM, e.g., the TSSEM approach proposed by (Cheung & Chan, 2005; Cheung, 2014), cannot handle continuous moderators. To address the moderating effects, we used the OSMASEM recently proposed by Jak and Cheung (2020) to test whether the selected study characteristics moderate the regression coefficients. The key idea is that the structural parameters, say *a*, *b*, and *c* in Fig. 1, are formulated as functions of the moderators. Then, we may test whether the moderators can be used to explain these parameters. All the analyses were conducted using the metaSEM package (Cheung, 2015b) implemented in the R statistical platform (R Development Core Team, 2020). The data, R code, and output are available as a registered Open Science Foundation project at https://osf.io/h976y.

Results

Descriptive Statistics

In Table 3, we report descriptive statistics for each correlation pair, including the number of studies and the total number of participants on which the meta-analytic data are based. As can be seen, the random effects meta-analytic correlation between automatic thinking and depression is quite large (+.67). The other two correlations are smaller but still

substantial: +.47 for the correlation between dysfunctional attitudes and automatic thinking and +.40 for correlation between dysfunctional attitudes and depression. The test statistic for the null hypothesis that all the correlation matrices are homogeneous was rejected, $\chi^2(df = 123) = 360.38$, $p < .001$, suggesting that the correlation matrices are not homogeneous. A prediction interval for each meta-analytic correlation can be found by multiplying the average correlation by $\pm 1.96*\tau$ in Table 3.

Assessing Publication Bias

Because almost two-thirds of the studies in this review are dissertations or theses, we did not expect that publication bias would have an important influence on the meta-analytic correlations we observed. Still, to be safe, we used the trim and fill approach to assess the plausibility of publication bias on these results. We ran the trim and fill analysis three times, once for each correlation of interest, using the metafor package in R (Viechtbauer, 2010). In two of the three cases, the trim and fill analysis did not impute studies to the left of the mean effect size (this is the direction that we would expect would be censored if publication bias were operating). In the other case, which involved correlations between measures of automatic thoughts and measures of depression, the trim and fill procedure identified two possibly censored studies. Based on the large percentage of dissertations that comprise our evidence base and on the results of the trim and fill analysis, we conclude that publication bias is unlikely to be an important influence on our results.

Testing a Structural Model

Because the proposed model is saturated, the model fit is perfect. Referring to Fig. 1, the estimated coefficients and their 95% LBCIs are 0.47 (0.42, 0.52) for path *a* (dysfunctional attitudes → automatic thoughts), 0.62 (0.56, 0.69) for path *b* (automatic thoughts → depression), and 0.11 (0.05, 0.17) for path *c* (dysfunctional attitudes → depression). As can be inferred from the confidence intervals, all individual paths are statistically significant at $\alpha = .05$. The estimated direct effect (path c) is only 0.11, whereas the estimated indirect effect of paths *a* and *b* is $0.47 \times 0.62 = 0.29$ (0.24, 0.35), more than two and a half times larger than the direct effect. A likelihood ratio test rejects

Table 3 Weighted descriptive statistics for the correlations of interest

Construct pair	Mean correlation	SE	τ	I^2	k	n
Dysfunctional attitudes — automatic thoughts	0.4701	.0269	.0721	53%	19	3,718
Depression — dysfunctional attitudes	0.4026	.0135	.0978	71%	90	18,550
Depression — automatic thoughts	0.6719	.0240	.0605	49%	52	11,980

SE standard error, τ estimated standard deviation of the true effect sizes, I^2 proportion of variability in true effect size that appears to be attributable to between-study differences, *k* number of correlations in the meta-analytic database, *n* total number of participants represented in the meta-analysis

Table 4 Moderating effects on the path coefficients and their standard errors

	Path *a* (dysfunctional attitudes → automatic thoughts)	Path *b* (automatic thoughts → depression)	Path *c* (dysfunctional attitudes → depression)
General samples	0.4831 (0.0166)	0.5911 (0.0428)	0.1312 (0.0303)
Clinical samples	0.2847 (0.0257)	0.5604 (0.0640)	0.1513 (0.0391)
Mixed samples	0.5163 (0.0328)	0.6805 (0.0627)	0.0923 (0.0572)

Values in parentheses are standard errors

the null hypothesis that the direct effect and the indirect effect are the same magnitude in the population, $\chi^2(df=1) = 13.88$, $p < .001$. This suggests that, consistent with the theory, the primary effect of dysfunctional attitudes on depression mediated by automatic thoughts is stronger than that of the direct effect.

Testing Study-Level Moderators in a Structural Model

We next tested three study-level characteristics that might moderate the relationships we observed. The first potential study-level moderator we examined was whether studies sampled children vs adults and older adults. As there were 3 path coefficients, we compared the models with and without the moderating effect. The test statistic was not statistically significant with $\chi^2(df=3) = 3.77$, $p = .29$, suggesting that there is not enough evidence to support that the path coefficients are different in samples based on children relative to path coefficients obtained from other samples. We conducted the same analysis on examining effects based on North American samples vs samples obtained from other locations. The test statistic was not statistically significant with $\chi^2(df=3) = 5.14$, $p = .16$, suggesting that there is not enough evidence to support that the path coefficients are different in North America samples vs other samples.

Finally, we tested whether the samples were general ($k=78$), clinical ($k=32$), or mixed ($k=15$). The test statistic was statistically significant with $\chi^2(df=6) = 25.41$, $p < .001$, suggesting that there are differences in the path coefficients among these groups. Further analyses on the individual paths show that they were all statistically significant at $\alpha = .05$. Table 4 shows the path coefficients in different samples. As can be seen, the primary difference is that for clinical samples, the path from dysfunctional attitudes to automatic thoughts is somewhat smaller (.28) than it is for other samples (.48 for general samples and .52 for mixed samples).

Discussion

We had several goals in writing this primer. One was to highlight how the strengths of systematic reviewing, meta-analysis, and structural equation modeling can be used to help inform theories and practices relevant to prevention science. To this end, we brought together 104 studies, all of which had one thing in common: They provided a correlation between at least two of three constructs relevant to the Beck's model of depression. This model posits, in part, that dysfunctional attitudes lead to negative automatic thinking, which leads to depression.

Although the paper is primarily focused on articulating how MASEM can be used, based on our data, we concluded that the indirect effect is substantially larger than the direct effect of dysfunctional attitudes on depression. Therefore, our results suggest that researchers developing prevention and treatment strategies for depression, and clinicians implementing these, should be aware of the importance of negative automatic thinking as a mediator of the relationship between dysfunctional attitudes and depression. We also tested three possible study-level moderators of this relationship. For two of these, we were unable to reject the null hypothesis that the models are similar. For the other, which examined whether the study sample was recruited from a general population, a clinical population, or a mix of these two sources, results suggested that while the overall patterns of the relationships between the variables were similar, the path from dysfunctional attitudes to automatic thoughts was somewhat smaller for clinical samples than it was for other samples. In generating these conclusions, we outlined the major considerations that potential users of MASEM will need to consider (model choice, dealing with dependent effect sizes, assessing publication bias, and whether to carry out attenuation corrections).

In highlighting how the strengths of systematic reviewing, meta-analysis, and structural equation modeling can be used to help inform theories and practices relevant to prevention science, we aimed not only to fulfill a tutorial function and show readers how this can be done, but also to convince readers about the value of doing so. To this end, it is worth emphasizing that of the three relationships that we address in this paper, the one with the least amount of information (the relationship between dysfunctional attitudes and automatic thoughts) is based on more than 25 times the number of participants that the median study had (3718 participants across 19 studies vs a median sample size of 143). And because these data were drawn from studies that varied along a number of theoretically important dimensions, there

is much more opportunity to test mediation and moderation using the MASEM framework than there is in an individual study. Finally, while the results of an individual study are conditional on the specific operations used, a systematic review involves multiple samples using diverse populations collected at different times, and these features almost certainly translate into better generalizability. All of these are reasons to considering adding MASEM to the toolbox of methods that are used to advance prevention science.

Appendix

References for studies included in the meta-analytic database

Ahnberg, J. L. (1997). *The prediction of recovery from dysphoria in a college sample* (Order No. MQ24640) [Doctoral dissertation, University of Calgary]. ProQuest Dissertations & Theses.

Anthony, A. J. (2003). Relationship of a multidimensional well-being measure to broad dimensions of personality, affect, thinking, optimism, and social hopelessness (Order No. NQ67916) [Doctoral dissertation, York University]. ProQuest Dissertations & Theses.

Arpin-Cribbie, C. A. (2002). *Understanding fatigue: An exploratory study of the role of psychosocial variables* (Order No. MQ76887) [Doctoral dissertation, University of Manitoba]. ProQuest Dissertations & Theses.

Asdigian, N. L. (1993). Toward an understanding of the cognitive etiology of depressive reactions to life stressors: An evaluation of the hopelessness theory of depression (Order No. 9420569) [Doctoral dissertation, University of New Hampshire]. ProQuest Dissertations & Theses.

Baggett, M. R. (1994). *States of mind as predictors of psychotherapeutic outcome in the treatment of depression* (Order No. 95147290 [Doctoral dissertation, Pacific Graduate School of Psychology]. ProQuest Dissertations & Theses.

Batmaz, S., Yuncu, O. A., & Kocbiyik, S. (2015). Assessing negative automatic thoughts: Psychometric properties of the Turkish version of the cognition checklist. *Iranian Journal of Psychiatry and Behavioral Sciences*, *9*(4). https://doi.org/10.17795/ijpbs-3444

Batmaz, S., Kocbiyik, S., Yalçınkaya-Alkar, Ö., & Turkcapar, M. (2016). Cognitive distortions mediate the relationship between defense styles and depression in female outpatients. *The European Journal of Psychiatry*, 30(4), 237–247.

Beevers, C. G., & Miller, I. W. (2004). Depression-related negative cognition: Mood-state and trait dependent properties. *Cognitive Therapy and Research*, *28*(3), 293-307.

Berle, D., & Starcevic, V. (2012). Preliminary validation of the Nepean dysphoria scale. *Australasian Psychiatry*, *20*(4), 322-326. https://doi.org/10.1177/1039856212447966

Bernholtz, B. (2013). *The client's and therapist's vocal qualities in CBT and PE-EFT for depression* (Order No. 1764221250) [Doctoral dissertation, University of Toronto]. ProQuest Dissertations & Theses.

Bianchi, R., & Schonfeld, I. S. (2016). Burnout is associated with a depressive cognitive style. *Personality and Individual Differences*, *100*, 1-5. https://doi.org/10.1016/j.paid.2016.01.008

Block, P. (1988). *A psychosocial expansion of the cognitive theory of depression* (Order No. 303696924) [Doctoral dissertation, New York University]. ProQuest Dissertations & Theses.

Bond, R. N. (1984). *Inducing depressed mood: A test of basic assumptions* (Order No. 303301746) [Doctoral dissertation, New York University]. ProQuest Dissertations & Theses.

Braithwaite, K. (2004). *Predictors of depressive symptomatology among black college women* (Order No. 305328687) [Doctoral dissertation, Howard University]. ProQuest Dissertations & Theses.

Brinker, J. K. (2007). *Dispositional rumination and depressed mood* (Order No. 304721708) [Doctoral dissertation, University of Western Ontario]. ProQuest Dissertations & Theses.

Buschmann, T. (2016). Towards an integrative cognitive therapy model: The primary role of cognitive rigidity in cognitive-behavioral therapies (Order No. 10188585) [Doctoral dissertation, Northern Arizona University]. ProQuest Dissertations & Theses.

Choon, M. W., Abu Talib, M., Yaacob, S. N., Awang, H., Tan, J. P., Hassan, S., & Ismail, Z. (2015). Negative automatic thoughts as a mediator of the relationship between depression and suicidal behaviour in an at-risk sample of Malaysian adolescents. *Child and Adolescent Mental Health*, *20*(2), 89-93. https://doi.org/10.1111/camh.10275

Church, N. F., Brechman-Toussaint, M. L., & Hine, D. W. (2005). Do dysfunctional cognitions mediate the relationship between risk factors and postnatal depression symptomatology? *Journal of Affective Disorders*, *87*(1), 65-72. https://doi.org/10.1016/j.jad.2005.03.009

Ciesla, J. A. (2004). *Rumination, negative cognition, and their interactive effects on depressed mood* (Order No. 3141263) [Doctoral dissertation, State University of New York at Buffalo]. ProQuest Dissertations & Theses.

Clark, C. G. (2002). *Depression, exercise schemas, and physical self-efficacy* (Order No. 3037965) [Doctoral dissertation, Fielding Graduate Institute]. ProQuest Dissertations & Theses.

Cowden, C. R. (1992). *Dysfunctional attitudes and self-concept discrepancy in depression and social anxiety* (Order No. 9226673) [Doctoral dissertation, The

Pennsylvania State University]. ProQuest Dissertations & Theses.

Cox, S. (2011). *Reflection, brooding, and negative cognitive content: Untangling associations with depressive symptoms and problem solving* (Order No. 3458358) [Doctoral dissertation, Seattle Pacific University]. ProQuest Dissertations & Theses.

Cui, L. X., Shi, G. Y., Zhang, Y. J., & Yu, Y. (2012). A study of the integrated cognitive model of depression for adolescents and its gender difference. Acta Psychological Sinica, 44(11), 1501-1514. https://doi.org/10.3724/SP.J.1041.2012.01501

D'Agostino, A., Manganelli, E., Aportone, A., Rossi Monti, M., & Starcevic, V. (2016). Development, cross-cultural adaptation process and preliminary validation of the Italian version of the Nepean Dysphoria Scale. *Journal of Psychopathology*, *22*, 149-156.

D'Alessandro, D. U. (2004). Development and validation of the child and adolescent dysfunctional attitudes scale: Tests of beck's cognitive diathesis -stress theory of depression, of its causal mediation component, and of developmental effects (Order No. NQ98233) [Doctoral dissertation, McGill University]. ProQuest Dissertations & Theses.

Du, X., Luo, W., Shen, Y., Wei, D., Xie, P., Zhang, J., Zhang, Q., & Qiu, J. (2015). Brain structure associated with automatic thoughts predicted depression symptoms in healthy individuals. *Psychiatry Research: Neuroimaging*, *232*(3), 257-263. https://doi.org/10.1016/j.pscychresns.2015.03.002

Dunkley, D. M., & Kyparissis, A. (2008). What is DAS self-critical perfectionism really measuring? Relations with the five-factor model of personality and depressive symptoms. *Personality and Individual Differences*, *44*(6), 1295-1305. https://doi.org/10.1016/j.paid.2007.11.022

Dyer, J. G. (1991). *A cross-sectional study of the cognitions of nuclear family members of unipolar major depressed patients* (Order No. 9205382) [Doctoral dissertation, The Catholic University of America]. ProQuest Dissertations & Theses.

Ekstrom, B. J. (2003). *Cognitive vulnerability to depression: Personality traits and differential response to depressed mood induction* (Order No. 3100751) [Doctoral dissertation, Southern Illinois University at Carbondale]. ProQuest Dissertations & Theses.

Finer-Greenberg, R. (1987). *Factors contributing to the degree of psychopathology in first- and second-generation holocaust survivors* (Order No. 8801241) [Doctoral dissertation, California School of Professional Psychology – Los Angeles]. ProQuest Dissertations & Theses.

Fischer, R. (2009). *Etiology of anxiety, depression, and aggression in children and youth* (Order No. 3425225) [Doctoral dissertation, The University of Toledo]. ProQuest Dissertations & Theses.

Frantom, L. V. (1994). *The role of attributional style, dysfunctional attitudes, and negative life events on depression in college students* (Order No. 1358562) [Doctoral dissertation, University of Nevada, Las Vegas]. ProQuest Dissertations & Theses.

Gerstein, R. K. (2010). *The long-term course of bipolar spectrum disorder: Applications of the behavioral approach system (BAS) model* (Order No. 3412334) [Doctoral dissertation, Temple University]. ProQuest Dissertations & Theses.

Grant, D. A. (2012). *Cognitive vulnerability and the actuarial prediction of depressive course* (Order No. 3552322) [Doctoral dissertation, Temple University]. ProQuest Dissertations & Theses.

Gül, A. I., Simsek, G., Karaaslan, Ö., & Inanir, S. (2015). Comparison of automatical thoughts among generalized anxiety disorder, major depressive disorder and generalized social phobia patients. *European Review for Medical and Pharmacological Sciences*, *19*(15), 2916-2921.

Gunter, J. D. (2008). *Examining the cognitive theory of depression using descriptive experience sampling* (Order No. 1456338) [Doctoral dissertation, University of Nevada, Las Vegas]. ProQuest Dissertations & Theses.

Haeffel, G. J. (2005). *Cognitive vulnerability to depression: Distinguishing between implicit and explicit cognitive processes* (Order No. 3186167) [Doctoral dissertation, University of Wisconsin - Madison]. ProQuest Dissertations & Theses.

Hamill Skoch, S. K. (2011). The relationship between temperament, cognitive vulnerabilities and depressive symptoms in American Indian youth from northern plains tribes (Order No. 3490555) [Doctoral dissertation, University of Wyoming]. ProQuest Dissertations & Theses.

Hamilton, E. W. (1982). *A longitudinal study of attributional style and dysfunctional attitudes in a clinical population* (Order No. 8307397) [Doctoral dissertation, State University of New York at Stony Brook]. ProQuest Dissertations & Theses.

Hankin, B. L. (2001). *Cognitive vulnerability-stress theories of depression: A 2-year prospective study to predict depressive symptoms and disorder* (Order No. 3020729) [Doctoral dissertation, University of Wisconsin - Madison]. ProQuest Dissertations & Theses.

Healy, K. J. (1991). Depression in college students: A reciprocal process of stressful life events, dysfunctional attitudes, and social support (Order No. EP11974) [Doctoral dissertation, Kean University]. ProQuest Dissertations & Theses.

Heinz, A. J., Veilleux, J. C., & Kassel, J. D. (2009). The role of cognitive structure in college student problem drinking. *Addictive Behaviors*, *34*(2), 212-218. https://doi.org/10.1016/j.addbeh.2008.10.011

Hiramura, H., Shono, M., Tanaka, N., Nagata, T., & Kitamura, T. (2008). Prospective study on suicidal ideation

among Japanese undergraduate students: Correlation with stressful life events, depression, and depressogenic cognitive patterns. *Archives of suicide research*, *12*(3), 238-250. https://doi.org/10.1080/13811110802100924

Holden, K. B., Hall, S. P., Robinson, M., Triplett, S., Babalola, D., Plummer, V., Treadwell, H., & Bradford, L. D. (2012). Psychosocial and sociocultural correlates of depressive symptoms among diverse African American women. *Journal of the National Medical Association*, *104*(11-12), 493-504. https://doi.org/10.1016/S0027-9684(15)30215-7

Hughes, M. E., Alloy, L. B., & Cogswell, A. (2008). Repetitive thought in psychopathology: The relation of rumination and worry to depression and anxiety symptoms. *Journal of Cognitive Psychotherapy*, *22*(3), 271-288. https://doi.org/10.1891/0889-8391.22.3.271

Hundt, N. E. (2011). *Reinforcement sensitivity, cognitive biases, stressful life events, and depression symptoms* (Order No. 3473465)) [Doctoral dissertation, University of North Carolina at Greensboro]. ProQuest Dissertations & Theses.

Hyers, D. A. (1995). *A structural analysis of relationships among stress, social support, dysfunctional attitudes, and depression in older adults* (Order No. 9531843) [Doctoral dissertation, University of North Carolina at Greensboro]. ProQuest Dissertations & Theses.

Ito, T., Takenaka, K., & Agari, I. (2005). Psychological vulnerability to depression: Negative rumination, perfectionism, immodithymic personality, dysfunctional attitudes, and depressive states. *Japanese Journal of Educational Psychology*, *53*(2), 162–171. https://doi.org/10.5926/jjep1953.53.2_162

Jarrett, R. B. (1983). *Mechanisms of change in cognitive-behavioral therapy in relation to depressives' dysfunctional thoughts* (Order No. 8408996) [Doctoral dissertation, University of North Carolina at Greensboro]. ProQuest Dissertations & Theses.

Karatepe, H. T., Yavuz, F. K., & Turkcan, A. (2013). Validity and reliability of the Turkish version of the ruminative thought style questionnaire. *Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology*, *23*(3), 231-241. https://doi.org/10.5455/bcp.20121130122311

Keller, F., Kirchner, I., & Pössel, P. (2010). Die Skala dysfunktionaler Einstellungen für Jugendliche (DAS-J). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 39, 234-243. https://doi.org/10.1026/1616-3443/a000054

Kelly, M. A. R. (2005). *Recovery from depression: Factors impacting non-treatment related improvements in depressive symptomatology* (Order No. 3185317) [Doctoral dissertation, State University of New York at Buffalo]. ProQuest Dissertations & Theses.

Knowles, R., Tai, S., Christensen, I., & Bentall, R. (2005). Coping with depression and vulnerability to mania: A factor analytic study of the Nolen-Hoeksema (1991) Response Styles Questionnaire. *British Journal of Clinical Psychology*, *44*(1), 99-112. https://doi.org/10.1348/014466504X20062

Kohlbeck, P. A. (1995). *Specificity of cognitive processes in depression* (Order No. 1376773) [Doctoral dissertation, Finch University of Health Sciences – The Chicago Medical School]. ProQuest Dissertations & Theses.

Konecky, M. S. (1996). *Autobiographical memory and clinical depression* (Order No. 9633706) [Doctoral dissertation, Rutgers, The State University of New Jersey]. ProQuest Dissertations & Theses.

LaGrange, B. (2008). *Disentangling the prospective relations between cognitive style and depressive symptoms* (Order No. 3365656) [Doctoral dissertation, Vanderbilt University]. ProQuest Dissertations & Theses.

Lam, K. F., Lim, H. A., Tan, J. Y., & Mahendran, R. (2015). The relationships between dysfunctional attitudes, rumination, and non-somatic depressive symptomatology in newly diagnosed Asian cancer patients. *Comprehensive Psychiatry*, *61*, 49-56. https://doi.org/10.1016/j.comppsych.2015.06.001

Lee, J. H. (2007). *Psychosocial risk and protective factors for depression among young adults with and without a history of depression* (Order No. 3288117) [Doctoral dissertation, University of Hawai'i at Manoa]. ProQuest Dissertations & Theses.

Leucht, S., Wada, M., & Kurz, A. (1997). Sind negative Kognitionen Symptome einer Depression oder auch Ausdruck von Persönlichkeitszügen?. *Der Nervenarzt*, *68*(7), 563-568.

Lewis, C. C. (2011). *Understanding patterns of change: Predictors of response profiles for clients treated in a CBT training clinic* (Order No. 3472378) [Doctoral dissertation, University of Oregon]. ProQuest Dissertations & Theses.

Linaman, T. E. (1996). Recollected parental verbal abuse and its relationship to dysfunctional attitudes, now and possible-future selves and depressive symptoms in adults (Order No. 9717682) [Doctoral dissertation, The Fielding Institute]. ProQuest Dissertations & Theses.

Lipinski, C. (2015). *Cognitive mechanisms of change in adolescent depression* (Order No. 3663317) [Doctoral dissertation, St. John's University]. ProQuest Dissertations & Theses.

Lutz, R. W. (1984). *Mood and the selective recall and evaluation of naturally occurring pleasant and unpleasant events (depression)* (Order No. 8415476) [Doctoral dissertation, Arizona State University]. ProQuest Dissertations & Theses.

Marčinko, D., Jakšić, N., Ivezić, E., Skočić, M., Surányi, Z., Lončar, M., Franić, T., & Jakovljević, M. (2014). Pathological narcissism and depressive symptoms in psychiatric outpatients: Mediating role of dysfunctional

attitudes. *Journal of Clinical Psychology*, *70*(4), 341-352. https://doi.org/10.1002/jclp.22033

Margolis, M. F. (1982). *The use of exercise and group cognitive therapy in the treatment of depression* (Order No. 8212763) [Doctoral dissertation, Hofstra University]. ProQuest Dissertations & Theses.

Mayer, J. L. (2004). *Postpartum mood disturbance in first time mothers: Application of cognitive diathesis-stress models of depression* (Order No. 3131331) [Doctoral dissertation, Idaho State University]. ProQuest Dissertations & Theses.

McCarron, J. A. (1980). *The relative efficacy of cognitive therapy and chemotherapy for the treatment of depression among the retired elderly* (Order No. 8028693) [Doctoral dissertation, California School of Professional Psychology]. ProQuest Dissertations & Theses.

Morris, M. C., Kouros, C. D., Fox, K. R., Rao, U., & Garber, J. (2014). Interactive models of depression vulnerability: The role of childhood trauma, dysfunctional attitudes, and coping. *British Journal of Clinical Psychology*, *53*(2), 245-263. https://doi.org/10.1111/bjc.12038

Noh, D., & Kim, S. (2016). Dysfunctional attitude mediates the relationship between psychopathology and Internet addiction among Korean college students: A cross-sectional observational study. *International Journal of Mental Health Nursing*, *25*(6), 588-597. https://doi.org/10.1111/inm.12220

Palmer, C. (2015). *Does "hot cognition" mediate the relationship between dysfunctional attitudes (cold cognition) and depression?* (Order No. 10293887) [Doctoral dissertation, University of Surrey]. ProQuest Dissertations & Theses.

Pedrelli, P. (2006). *Generalizability of the cognitive diathesis-stress model of depression to depressive symptoms in schizophrenia* (Order No. 3208619) [Doctoral dissertation, University of California, San Diego and San Diego State University]. ProQuest Dissertations & Theses.

Philpot, V. D., Holliman, W. B., & Madonna Jr, S. (1995). Self-statements, locus of control, and depression in predicting self-esteem. *Psychological Reports*, *76*(3), 1007-1010. https://doi.org/10.2466%2Fpr0.1995.76.3.1007

Pirbaglou, M., Cribbie, R., Irvine, J., Radhu, N., Vora, K., & Ritvo, P. (2013). Perfectionism, anxiety, and depressive distress: Evidence for the mediating role of negative automatic thoughts and anxiety sensitivity. *Journal of American College Health*, *61*(8), 477-483. https://doi.org/10.1080/07448481.2013.833932

Poppleton, L. E. (2008). Evaluating mediators and moderators of cognitive behavioral telephone treatment for adult major depression in a rural population (Order No. 3293517) [Doctoral dissertation, Brigham Young University]. ProQuest Dissertations & Theses.

Quiggle, N. L. (2001). *The relation between childhood abuse and adult depression: Negative cognitions as a mediator* (Order No. 3038832) [Doctoral dissertation, Vanderbilt University]. ProQuest Dissertations & Theses.

Randolph, J. J., & Dykman, B. M. (1998). Perceptions of parenting and depression-proneness in the offspring: Dysfunctional attitudes as a mediating mechanism. *Cognitive Therapy and Research*, *22*(4), 377-400. https://doi.org/10.1023/A:1018761229824

Raynor, C. M., & Manderino, M. A. (1988). The prevalence of adolescent depression in a residential treatment home program. *Residential Treatment for Children & Youth*, *5*(3), 17-27. https://doi.org/10.1300/J007v05n03_03

Reeder, J. (2014). *Cognitive differences among individuals with clinically significant depression, alcohol use, and comorbidity* (Order No. 1525728) [Doctoral dissertation, California State University, Fullerton]. ProQuest Dissertations & Theses.

Robinson, P. J. (1987). *Depressotypic cognitive patterns in conjugal bereavement and major depression* (Order No. NL35793) [Doctoral dissertation, York University]. ProQuest Dissertations & Theses.

Ross, S. M., Gottfredson, D. K., Christensen, P., & Weaver, R. (1986). Cognitive self-statements in depression: Findings across clinical populations. *Cognitive Therapy and Research*, *10*(2), 159-165. https://doi.org/10.1007/BF01173722

Rough, J. N. (2016). *An integrative chronobiological-cognitive approach to seasonal affective disorder* (Order No. 3738922) [Doctoral dissertation, University of Vermont and State Agricultural College]. ProQuest Dissertations & Theses.

Russell, N. N. (1983). The relationship between autonomous and sociotropic types of depression and dimensions of depression: Personality clusters; symptomatology; and cognitive theme content (Order No. 8326188) [Doctoral dissertation, Fordham University]. ProQuest Dissertations & Theses.

Sadowski, C. (1994). *Cognitive distortions, impulsivity, and stressful life events in suicidal adolescents* (Order No. 9524482) [Doctoral dissertation, Louisiana State University and Agricultural & Mechanical College]. ProQuest Dissertations & Theses.

Schatz, N. K. (2009). *Maternal depression and parenting stress among families of children with AD/HD: Child and family correlates* (Order No. 1464713) [Master's thesis, University of North Carolina at Greensboro]. ProQuest Dissertations & Theses.

Scherrer, M. C. (2009). Predictors and cognitive consequences of a negative mood induction procedure: The development of depressed mood across diagnostic categories (Order No. NR49624) [Doctoral dissertation, University of Calgary]. ProQuest Dissertations & Theses.

Schroeder, R. M. (1995). *The relationship between depression, cognitive distortion, and life stress among African-American college students* (Order No. 9605168)

[Doctoral dissertation, University of North Carolina at Chapel Hill]. ProQuest Dissertations & Theses.

Schwartz-Stav, O., Apter, A., & Zalsman, G. (2006). Depression, suicidal behavior and insight in adolescents with schizophrenia. *European Child & Adolescent Psychiatry*, *15*(6), 352-359. https://doi.org/10.1007/s00787-006-0541-8

Shapira, L. B. (2012). *Exploring online loving-kindness and CBT interventions: Do individuals vulnerable to depression improve?* (Order No. NS00174) [Doctoral dissertation, York University]. ProQuest Dissertations & Theses.

Sommer, D. H. (1999). *Relationships between sexual abuse, cognitive style, and depression in adolescent psychiatric inpatients* (Order No. 9947395) [Doctoral dissertation, University of Texas at Austin]. ProQuest Dissertations & Theses.

Stanley, P. H. (1991). *Asymmetry in internal dialogue, core assumptions, valence of self-statements and counselor trainee effectiveness* (Order No. 9204464) [Doctoral dissertation, University of North Carolina at Greensboro]. ProQuest Dissertations & Theses.

Stevens, E. A. (2006). Adolescent psychopathy and cognitive style: Associations between psychopathic traits, depressive symptoms, attributional style, and dysfunctional attitudes among youth (Order No. 3214309) [Doctoral dissertation, Yale University]. ProQuest Dissertations & Theses.

Stolow, D. (2011). *A prospective examination of beck's cognitive theory of depression in university students in mainland china* (Order No. 1500353) [Master's thesis, Rutgers, The State University of New Jersey]. ProQuest Dissertations & Theses.

Sutton, J. M., Mineka, S., Zinbarg, R. E., Craske, M. G., Griffith, J. W., Rose, R. D., Waters, A. M., Nazarian, M., & Mor, N. (2011). The relationships of personality and cognitive styles with self-reported symptoms of depression and anxiety. *Cognitive Therapy and Research*, *35*(4), 381-393. https://doi.org/10.1007/s10608-010-9336-9

Szymanski, J. B. (1997). Investigating the underlying factor structure of measures of depressive symptomatology and correlates of depression using structural equation modeling (Order No. 9733545) [Doctoral dissertation, Northern Illinois University]. ProQuest Dissertations & Theses.

Tajima, M., Akiyama, T., Numa, H., Kawamura, Y., Okada, Y., Sakai, Y., Miyake, Y., Ono, Y., & Power, M. J. (2007). Reliability and validity of the Japanese version of the 24-item Dysfunctional Attitude Scale. *Acta Neuropsychiatrica*, *19*(6), 362-367. https://doi.org/10.1111/j.1601-5215.2007.00203.x

Thimm, J. C. (2010). Personality and early maladaptive schemas: A five-factor model perspective. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(4), 373-380. https://doi.org/10.1016/j.jbtep.2010.03.009

Turkoglu, S. A., Essizoglu, A., Kosger, F., & Aksaray, G. (2015). Relationship between dysfunctional attitudes and childhood traumas in women with depression. *International Journal of Social Psychiatry*, *61*(8), 796-801. https://doi.org/10.1177%2F0020764015585328

Vatanasin, D., Thapinta, D., Thompson, E. A., & Thungjaroenkul, P. (2012). Testing a model of depression among Thai adolescents. *Journal of Child and Adolescent Psychiatric Nursing*, *25*(4), 195-206. https://doi.org/10.1111/jcap.12012

Wierzbicki, M., & Rexford, L. (1989). Cognitive and behavioral correlates of depression in clinical and non-clinical populations. *Journal of Clinical Psychology*, *45*(6), 872-877. https://doi.org/10.1002/1097-4679(198911)45:6%3C872::AID-JCLP2270450607%3E3.0.CO;2-T

Wong, K. C. (2008). Psychometric investigation into the construct of neurasthenia and its related conditions: A comparative study on Chinese in Hong Kong and Mainland China (Order No. 3348888) [Doctoral dissertation, The Chinese University of Hong Kong]. ProQuest Dissertations & Theses.

Yilmaz, A. E., Gençöz, T., & Wells, A. (2015). Unique contributions of metacognition and cognition to depressive symptoms. *The Journal of General Psychology*, *142*(1), 23-33. https://doi.org/10.1080/00221309.2014.964658

You, S. (2007). A gender comparison of cognitive vulnerability as a function of moderation and mediation between negative life events and depressive mood (Order No. 3291219) [Doctoral dissertation, Purdue University]. ProQuest Dissertations & Theses.

Zheng, X. Y., Xuan, L., Yang, L. L., Yao, S. Q., & Xiao, J. (2014). Dysfunctional attitude and depressive symptoms: The moderator effects of daily hassles. *Chinese Journal of Clinical Psychology*.

Declarations

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International University Press.

Becker, B. J. (1992). Using results from replicated studies to estimate linear models. *Journal of Educational Statistics, 17*, 341–362. https://doi.org/10.3102/10769986017004341

Becker, B. J. (1995). Corrections to "using results from replicated studies to estimate linear models." *Journal of Educational and Behavioral Statistics, 20*, 100–102. https://doi.org/10.2307/1165390

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioural therapy: A review of meta-analyses. *Clinical Psychology Review, 26*, 17–31. https://doi.org/10.1016/j.cpr.2005.07.003

Charles, E. P. (2005). The correction for attenuation due to measurement error: Clarifying concepts and creating confidence sets. *Psychological Methods, 10*, 206–226. https://doi.org/10.1037/1082-989X.10.2.206

Cheung, M.W.-L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods, 41*, 425–438. https://doi.org/10.3758/BRM.41.2.425

Cheung, M.W.-L. (2014). Fixed- and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods, 46*, 29–40. https://doi.org/10.3758/s13428-013-0361-y

Cheung, M. W.-L. (2015a). *Meta-analysis: A structural equation modeling approach*. John Wiley & Sons, Inc.

Cheung, M. W.-L. (2015b). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*(1521). https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M.W.-L. (2019). Some reflections on combining meta-analysis and structural equation modeling. *Research Synthesis Methods, 10*, 15–22. https://doi.org/10.1002/jrsm.1321

Cheung, M. (2020). *Meta-analytic structural equation modeling*. PsyArXiv. https://doi.org/10.31234/osf.io/epsqt

Cheung, M.W.-L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods, 10*, 40–64. https://doi.org/10.1037/1082-989X.10.1.40

Cheung, M.W.-L., & Cheung, S. F. (2016). Random-effects models for meta-analytic structural equation modeling: Review, issues, and illustrations. *Research Synthesis Methods, 7*, 140–155. https://doi.org/10.1002/jrsm.1166

Cheung, M.W.-L., & Hafdahl, A. R. (2016). Special issue on meta-analytic structural equation modeling: Introduction from the guest editors. *Research Synthesis Methods, 7*, 112–120. https://doi.org/10.1002/jrsm.1212

Cooper, H. (2017). *Research synthesis and meta-analysis* (5th ed.). SAGE Publications.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Jak, S., & Cheung, M.W.-L. (2020). Meta-analytic structural equation modeling with moderating effects on SEM parameters. *Psychological Methods, 25*, 430–455. https://doi.org/10.1037/met0000245

Ke, Z., Zhang, Q., & Tong, X. (2019). Bayesian meta-analytic SEM: A one-stage approach to modeling between-studies heterogeneity in structural parameters. *Structural Equation Modeling: A Multidisciplinary Journal, 26*, 348–370. https://doi.org/10.1080/10705511.2018.1530059

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*, 486–507. https://doi.org/10.1177/0049124114543236

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2*(1), 61-76. https://doi.org/10.1002/jrsm.35

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99–128. https://doi.org/10.1207/s15327906mbr3901_4

Michel, J. S., Viswesvaran, C., & Thomas, J. (2011). Conclusions from meta-analytic structural equation models generally do not change due to corrections for study artifacts. *Research Synthesis Methods, 2*, 174–187. https://doi.org/10.1002/jrsm.47

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100*, 569–582. https://doi.org/10.1037/0021-843X.100.4.569

Pigott, T. D., Valentine, J. C., Polanin, J. R., Williams, R., & Canada, D. D. (2013). Outcome-reporting bias in education research. *Educational Researcher, 42*, 424–432. https://doi.org/10.3102/2F0013189X13507104

Pössel, P., & Black, S. W. (2014). Testing three different sequential mediational interpretations of Beck's cognitive model of the development of depression. *Journal of Clinical Psychology, 70*, 72–94. https://doi.org/10.1002/jclp.22001

Pössel, P., Smith, E., & Alexander, O. (2018). LARS&LISA: A universal school-based cognitive-behavioral program to prevent adolescent depression. *Psicologia: Reflexão e Crítica | Psychology: Research and Review, 31,* 23. https://doi.org/10.1186/s41155-018-0104-1.

R Development Core Team. (2020). *R: A language and environment for statistical computing*. http://www.R-project.org/

Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient alpha. *Psychological Methods, 11*, 306–322. https://doi.org/10.1037/1082-989X.11.3.306

Schmidt, F. L., Le, H., & Oh, I.-S. (2019). Correcting for the distorting effects of study artifacts in meta-analysis and second order meta-analyses. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 315–337). Russell Sage Foundation.

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*. https://doi.org/10.1136/bmj.j4008

Shen, J., Wang, Y., Kurpad, N., & Schena, D. A. (2021). A systematic review on the impact of hot and cool executive functions on pediatric injury risks: A meta-analytic structural equation modeling approach. *Prevention Science*. https://doi.org/10.1007/s11121-021-01271-2

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5*, 13–30. https://doi.org/10.1002/jrsm.1091

Valentine, J. C. (2019). Incorporating judgments about study quality into research syntheses. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 129–140). Russell Sage Foundation.

Vevea, J. L., Coburn, K., & Sutton, A. (2019). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 383–432). Russell Sage Foundation.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). *Deploying an interactive machine learning system in an evidence-based practice center: abstrackr*. In Proceedings of the ACM International Health Informatics Symposium (IHI), 819–824.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Sage.

Yuan, K.-H., Wu, R., & Bentler, P. M. (2011). Ridge structural equation modelling with correlation matrices for ordinal and continuous data. *British Journal of Mathematical and Statistical Psychology, 64*, 107–133. https://doi.org/10.1348/000711010X497442

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.