How *UN*reliable are adult-reported suicide attempts?

An examination of correlates and underlying causal mechanisms of discordant reporting over time.

Niki P. Palmetto, MPhil, MPH

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences
COLUMBIA UNIVERSITY
2011

ABSTRACT

How *UN*reliable are adult-reported suicide attempts?

An examination of correlates and underlying causal mechanisms of discordant reporting over time.

Niki P. Palmetto, MPhil, MPH

The challenge of capturing suicide attempts in the population, plague its examination. The reliability of adult-reported lifetime suicide attempts had not been rigorously explored prior to this work, and therefore estimates have remained largely unchallenged. This dissertation explicitly sought to fill this research gap by utilizing a longitudinal study, comprising two waves of data collection, in which information on suicide attempts was obtained at both time points.

Chapter 1 presents a systematic review of the literature depicting the state of the literature with regards to the reliability of suicidality measures (e.g. ideation, plans, and attempts). Few studies assessed correlates of discordant reporting, and no studies examined the reliability of adult-reported suicide attempts. Drawing upon the correlates reported within studies, as well sources of heterogeneity across studies, I posited four plausible causal mechanisms underlying discordant suicidality reporting; recall failure, reinterpretation, conscious denial, and lack of construct comprehension. Extending these findings, I proposed that the likelihood of each mechanism is influenced by factors such as the severity of the suicidality, amount of time passed since the suicidal event, social desirability effects, mood context, and suicide construct validity.

In Chapter 2, I assessed the reliability of adult-reported lifetime attempts as reported in a large, population-based longitudinal study, and found reports to be moderately reliable, with a Kappa

coefficient of 0.51. I hypothesized that discordant reporters would be more similar to individuals who reported a past attempt at both waves (Concordant yes responders), compared with individuals who reported no attempt at both waves (Concordant no responders). I found that indeed, discordant reporters were more similar to the former, potentially signifying that discordant reporters are true attempters who underreported their attempt at one time point. Further, I hypothesized that discordant reporters would be less likely to have a history of depressive disorders compared with Concordant yes responders; positing that this history would serve as a marker for attempt severity, and that discordants would have less severe attempts, which would therefore be more easily forgotten or reinterpreted. Contrary to this hypothesis however, discordants were as likely as Concordant yes individuals to have a history of depressive disorders, and unexpectedly, discordants were much less likely to have a history of suicidal ideation. It is therefore plausible that a history of suicidal ideation serves as a marker for attempt severity, and/or that discordant reporters are characterized by more impulsive attempts.

In Chapter 3, I examined how a respondent's current depressed mood may influence the recall, and hence reporting of attempts. Based on established mood-recall theories, I tested three competing hypotheses to determine if a current depressed mood would enhance (mood-congruent recall), inhibit (mood memory deficit effect), or have no effect on the recall (mood-independent recall) and reporting of attempts. I hypothesized that discordant reporters would demonstrate a mood-congruent pattern of reporting, such that a depressed mood at the time of the interview would increase the likelihood that a respondent would report an attempt at that wave. There were in fact, distinct mood-congruent reporting effects among Recanters, yet mood-independent effects detected among New endorsers. This may indicate that New endorsers are a

unique group of discordant responders, which warrant further examination. Still, because respondents in our sample were over 20 times more likely to recant than newly endorse, and comparatively, there was limited power within our New endorser group, I believe these results may be generalized to assert that overall, discordant responders report in a mood-congruent fashion.

Table of contents

List of Tables and Charts

Chapter 1

Chapter 2

Chapter 3

Appendices

Acknowledgments

I would like to acknowledge my many mentors, colleagues, friends, and family, without which, my doctorate and dissertation would have undoubtedly not been possible.

I would like to first thank the members of my dissertation committee; Bruce Link, Sharon Schwartz, Debbie Hasin, Matt Miller, and Maddy Gould. All of my committee members provided me with theoretical and methodological guidance and support, coming from their own unique disciplines and expertise to substantially strengthen my dissertation. I'd especially like to thank Bruce for encouraging me to think about my work in new and innovative ways, not only regarding my research questions, but also the application of my work within a broader context. Bruce regards his students with the same respect as he would a colleague; he was generous with his knowledge and expertise, but allowed me to develop my own identity as a researcher. I'd like to Sharon for not only her commitment to and belief in my dissertation work, but also her for her guiding role in the 6 years that led up to this point. Sharon has comfortingly witnessed many of my tears, and equally, shared in my progress. Without Sharon's continuous encouragement, I sincerely do not believe I would have finished the program how, and when I did. One could say that Sharon was a "necessary, while not sufficient cause", of my success in this program. Lastly, I'd like to thank Matt for his willingness to take in an orphan, but budding suicidologist. Matt selflessly offered his time and expertise to this project, without which, my dissertation would not have been what it was.

I would have not been able to navigate this program nearly as well, without the administrative support of Brenda Scariff, Liliane Zaretsky, and Elizabeth Ferrari. They are tireless and between their skills and connections, could probably rule the world (or at least the epidemiology world).

I'd like to acknowledge and thank my colleagues that I met along the way, and who's epidemiologic skills, as well as support and friendship made the process that much less painful. Among which, are the members of my writing group; Sarah Braunstein, Eduard Eduardo, Vera Frajzyngier, Heidi Jones, Kerry Keyes, Ann Madsen and Katherine Ornstein. I believe we are bonded by a very unique experience, and are who I would consider my "cohort".

I'd also like to thank my closest friends, all of whom have been right there with me from the very start; Laura, Bo, Syl, Maggie, Eva, and Rara. Laura has been my best friend throughout my entire academic career, kindergarten through now. Laura's unwavering friendship throughout the years provided me with a sense of stability and home, even during the most unstable periods in my life, and within this process. Bo, Syl, Maggie, Eva and Rara, my best friends from college, enthusiastically celebrated every step of the long and drawn out process along with me. I cannot thank them enough for during the most hectic periods, understanding that I was able to provide less of myself as a friend, and yet still providing their full support and love in return. Lastly, I'd like to thank my newest "friend", Pete. Pete has seen me through the last year of this process, and in my most neurotic of states, and that deserves thanks in and of itself. Pete always insisted that my work came first, and so I thank him for that strength and at times, sacrifice.

Lastly, I'd like to thank my family for their love and support. They are truly my biggest champions in life. I'd like to my Mom and Dad, for putting up with a child in prolonged adolescence, and for always reminding me that if I quit at any time, they'd be equally as proud of

me—that knowledge provided me the strength to persevere. And to my brothers Darin and Erik, and my sister-in-law Rachel who made me feel supported by bragging about me to their friends—although we've all chosen different paths, I appreciate knowing that you respect mine. Lastly, I'd like to thank my niece, Stella, and nephew, Ryder, for ensuring that my school breaks were filled with love and laughter.

Introduction

Epidemiologists rely on respondent-reported suicide attempts elicited through population-based surveys to obtain estimates of incidence, prevalence, and correlates. These estimates are necessary and critical for public health since they ultimately inform prevention and intervention efforts, and allow for trend surveillance and appropriate resource allocation. However, how reliable are these suicide attempt reports? Further, if respondents do not reliably report past attempts, what factors influence their reporting behaviors? It is necessary to quantify the reliability of lifetime attempts reported by adults, and further, to assess factors that influence the reporting of attempts, in order to understand potential causal mechanisms underlying discordant reporting. This dissertation aims to address this need.

Adult suicide attempts: reported prevalence and predictors

A suicide attempt is a self-injurious, but non-fatal act to take one's life; an act undertaken by approximately 5% of the US adult population in their lifetime.[1-3] Despite increased intervention efforts, the prevalence of adult lifetime attempts has remained relatively steady for the past two decades.[2,4] Suicide attempts are most common among women and young adults[4,5], and vary greatly in regards to degree of pre-meditation, intention to die, and medical consequences. The majority of attempts take place after at least some degree of contemplation (suicidal ideation) and planning, such as time, place, and method; still many take place after relatively little pre-meditation or planning[1,5], and accordingly, are considered "impulsive attempts".[6,7] Some attempts are carried out with a high degree of intention to die, and followed by hospitalizations, while some attempts are undertaken with more ambivalence and require little to no medical intervention. A consistent and pervasive commonality across attempters however, is their

likelihood of having had a mental health disorder at the time of the attempt; 90% of attempters can be diagnosed with a mental health problem, 60% of which are mood disorders.[8] There is some evidence that individuals who attempt outside the context of a mental health disorder are likely to make more impulsive,[6,7] and less medically and psychologically severe attempts.[9-14]

Capturing suicide attempts via self reports

The majority of descriptive and analytic attempt estimates are obtained through the use of large, epidemiological surveys, such as the National Comorbidity Study and the National Survey on Alcohol and Related Conditions, the dataset used in this dissertation. A significant benefit of these surveys is that their enormity provide for a viable sample size of this rare event, and capture a relatively representative sample of both attempts (e.g. medically serious and non-medically serious; impulsive and non-impulsive) and attempters. However, these large surveys also have constraints and limitations. First, because of the size of the surveys and respondent population, the number of questions regarding attempts is necessarily limited; most large population-based surveys contain only one or two questions regarding attempts[1,3,15-18], limiting the amount of details obtained on them. In addition, as with all self-report measures, they require the respondent to accurately report their experiences. Respondents must understand what constitutes an attempt, recall making an attempt, and be willing to reveal it to the interviewer; aspects that may hinder the validity and reliability of attempt reports.

Reliability of suicide attempt reports

Reliability is the consistency, or repeatability of a measure. In regards to suicide attempts, the test-retest reliability of a measure refers to the degree to which individuals respond to a question about past attempts the same way, at two separate time points. A discordant reporter is a respondent who reports a past attempt inconsistently over time (e.g. reports a past attempt at Time 1, but not at Time 2). The test-retest reliability of a measure is typically represented by a Kappa coefficient. The Kappa coefficient (K) is the amount of agreement across time, above and beyond that which would have occurred due to chance alone.[19] The reliability of a measure is also informative of a measure's validity, since a measure can be only as valid as it is reliable. If a respondent reports discordantly over time, one of the reports must be inaccurate. While there is no way to directly assess the validity of these self reports, with longitudinal reports of past attempts, reliability can be estimated.

There is no reliability estimate of adult-reported attempts in the literature; however there is ancillary evidence indicating likely compromised reliability. One indication of inconsistent reporting over time is the higher 12-month prevalence estimates among youth (7.3-10.6%) compared with the lifetime prevalence estimates among adults (1.9-4.6%).[5] Five studies documenting that attempt reporting is unreliable to varying degrees among youth (K= 0.58-0.80) provide further support for the likely hindered reliability of attempts reported by adults.[20-24] Therefore, while it is likely that adult-reported attempt estimates are unreliable to some degree, it has yet to be quantified. Further, it is unknown whether inconsistent reporting is the result of individuals underreporting true attempts, or falsely reporting attempts at one time point. Still,

regardless of the direction of reporting, discordant reporting over time has important implications for prevalence estimates, as well as for purposes of screening, resource allocation, and etiologic examination.

Why might individuals unreliably report their attempts? Predictors of reporting and underlying causal mechanisms

Very little is known about factors that influence a respondent's likelihood of reporting a past attempt. However, based on what is known to influence self-reports in general, as well as speculation from suicide researchers, there are many plausible underlying causal mechanisms; including issues surrounding lack of construct comprehension, social desirability, reinterpretation, and recall failure.

Respondent understanding of what constitutes a suicide attempt may vary. For example, respondents may conflate constructs such as suicidal gestures (without true intention to die) and non-suicidal self-harm (e.g. cutting, burning), with attempts. Without a concrete appreciation of what comprises an attempt, individuals may be more likely to respond differently over time. Factors that could plausibly influence attempt construct comprehension include age, education, as well as the specificity, and wording of the question. Respondents must also be willing to reveal making an attempt to the interviewer. Some respondents feel shame, regret, or embarrassment regarding their past attempt,[21,25-27] and therefore may choose not to report it at one of the time points. Factors that may influence social desirability include age, sex, and characteristics of the interviewer. Respondents may also reinterpret their attempts over

time;[21,25,26] that is, a respondent may report a history of suicidality at Time 1, but later, reinterpret that event as not of suicidal nature, and therefore not report it at follow-up. Lastly, respondents may fail to remember a past attempt. This recall failure may be the result of active memory suppression, as a way to heal and move on[25], or the result of simple recall failure and the natural decline of memory over time.[21,28] Both reinterpretation and recall failure would likely be influenced by factors such as the amount of time since the attempt, context in which the question is embedded, the severity of the attempt, and history of a mental health disorder. One factor in particular that has been highlighted as a likely influential factor is the respondents' current mood at the time of the interview.[25,26,28] Whether or not the respondent has a depressed mood at the time of the interview, may affect their likelihood of reporting a past attempt. Because respondent mood can vary over time, it is a strong candidate for explaining varying attempt reports over time.

Potential effects of respondent mood on the reporting of attempts

There are three possible ways in which a respondent's current mood may affect the reporting of attempts. A depressed mood at the time of the interview may enhance, inhibit, or have no effect on the likelihood that a respondent reports a past attempt during a particular interview. "Mood congruent recall" purports that mood influences memory by making an individual's memories that are congruent with their mood at the time of retrieval, more accessible to them.[29,30] Therefore, when a respondent is currently depressed, they may be more likely to recall a past suicide attempt, and hence report it. Alternatively, a depressed mood at the time of the interview may inhibit recall of past attempts. Memory deficit effects, particularly with regards to episodic

memory, are typical in individuals suffering from depression.[31-40] In particular, individuals with depression have been found to have an "overgeneral" memory, which affects their recall of specific events.[41-45] Therefore, when a respondent is depressed at the time of the interview, they may be less likely to recall a past attempt, and hence not report it. Finally, a depressed mood at the time of the interview may have no effect on a respondent's likelihood of reporting a past attempt. Rather, discordant reporting over time may be affected by other, non-mood related factors, such as the severity of the attempt.

Aims of the Dissertation

The objective of this dissertation is to advance the understanding of suicide attempt reporting by adults. Specifically, I examine the extent to which attempts are reported unreliably, or inconsistently over time, and factors which may influence the likelihood of reporting an attempt. In particular, by testing competing hypotheses, I aim to evaluate how a respondent's mood at the time of the interview may impact their likelihood of reporting a past attempt.

I examine the aims of my dissertation in a series of three papers (Chapters 1, 2 and 3). I begin with a systematic examination of the literature in Chapter 1. This literature review critically assesses and synthesizes the findings across studies that have examined the test-retest reliability of suicidality constructs (ideation, plans or attempts). This chapter serves as the motivating foundation for this dissertation, as it highlights the critical need for an estimate of reliability for adult-reported attempts. Further, by drawing upon studies among youth, and across other suicidal

constructs, I posit potentially significant correlates of discordant reporting, as well as causal mechanisms underlying attempt reporting.

Chapter 2 is the first empirical paper of this dissertation. This paper's objective is to add to the body of literature by providing the first reliability estimate of lifetime suicide attempts as reported by adults. Further, I will assess sociodemographic and psychiatric correlates of discordant reporting over time, particularly focusing on the respondent's history of depressive disorders as a potential influential factor. In addition, by comparing discordant and concordant responders on these various correlates, I will evaluate the likelihood of discordant reporters being true attempters who underreported their attempt, or true non-attempters who falsely reported an attempt at one time point.

Chapter 3 aims to extend the findings of Chapter 2, with an in-depth examination of how a respondent's current mood at the time of the interview may affect their likelihood of reporting a past attempt. Specifically, through testing three competing theories, this paper will examine if a current depressed mood enhances, inhibits, or has no effect on the reporting of attempts.

Together, these three chapters aim to further our understanding of the reliability of suicide attempt measures. I attempt to quantify the degree of discordant reporting over time, and posit plausible reasons for the likely underreporting of attempts in the population, particularly the effects of a respondent's mood state at the time of an interview.

References

1. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry.* Jul 1999;56(7):617-626.

2. Kessler RC, Berglund P, Borges G, Nock M, Wang PS. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990-1992 to 2001-2003. *Jama.* May 25 2005;293(20):2487-2495.

3. Nock MK, Kessler RC. Prevalence of and risk factors for suicide attempts versus suicide gestures: analysis of the National Comorbidity Survey. *J Abnorm Psychol.* Aug 2006;115(3):616-623.

4. Baca-Garcia E, Perez-Rodriguez MM, Keyes KM, et al. Suicidal ideation and suicide attempts in the United States: 1991-1992 and 2001-2002. *Mol Psychiatry.* Sep 9 2008.

5. Nock MK, Borges G, Bromet EJ, Cha CB, Kessler RC, Lee S. Suicide and suicidal behavior. *Epidemiol Rev.* 2008;30:133-154.

6. Conner KR. A call for research on planned vs. unplanned suicidal behavior. *Suicide Life Threat Behav.* Summer 2004;34(2):89-98.

7. Simon OR, Swann AC, Powell KE, Potter LB, Kresnow MJ, O'Carroll PW. Characteristics of impulsive suicide attempts and attempters. *Suicide Life Threat Behav.* 2001;32(1 Suppl):49-59.

8. Beautrais AL, Joyce PR, Mulder RT, Fergusson DM, Deavoll BJ, Nightingale SK. Prevalence and comorbidity of mental disorders in persons making serious suicide attempts: a case-control study. *Am J Psychiatry.* Aug 1996;153(8):1009-1014.

9. Astruc B, Torres S, Jollant F, et al. A history of major depressive disorder influences intent to die in violent suicide attempters. *J Clin Psychiatry.* May 2004;65(5):690-695.

10. Brown GK, Henriques GR, Sosdjan D, Beck AT. Suicide intent and accurate expectations of lethality: predictors of medical lethality of suicide attempts. *J Consult Clin Psychol.* Dec 2004;72(6):1170-1174.

11. Dhossche DM, Meloukheia AM, Chakravorty S. The association of suicide attempts and comorbid depression and substance abuse in psychiatric consultation patients. *Gen Hosp Psychiatry.* Jul-Aug 2000;22(4):281-288.

12. Hasley JP, Ghosh B, Huggins J, Bell MR, Adler LE, Shroyer AL. A review of "suicidal intent" within the existing suicide literature. *Suicide Life Threat Behav.* Oct 2008;38(5):576-591.

13. Kumar CT, Mohan R, Ranjith G, Chandrasekaran R. Characteristics of high intent suicide attempters admitted to a general hospital. *J Affect Disord.* Mar 2006;91(1):77-81.

14. Suominen K, Isometsa E, Henriksson M, Ostamo A, Lonnqvist J. Hopelessness, impulsiveness and intent among suicide attempters with major depression, alcohol dependence, or both. *Acta Psychiatr Scand.* Aug 1997;96(2):142-149.

15. Bolton JM, Belik SL, Enns MW, Cox BJ, Sareen J. Exploring the correlates of suicide attempts among individuals with major depressive disorder: findings from the national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry.* Jul 2008;69(7):1139-1149.

16. Dube SR, Anda RF, Felitti VJ, Chapman DP, Williamson DF, Giles WH. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: findings from the Adverse Childhood Experiences Study. *Jama.* Dec 26 2001;286(24):3089-3096.

17. Garroutte EM, Goldberg J, Beals J, Herrell R, Manson SM. Spirituality and attempted suicide among American Indians. *Soc Sci Med.* Apr 2003;56(7):1571-1579.

18. Ialongo N, McCreary BK, Pearson JL, et al. Suicidal behavior among urban, African American young adults. *Suicide Life Threat Behav.* Fall 2002;32(3):256-271.

19. Cohen J. A coefficient of agreement for nominal scales. *Educational Psychological Measures.* 1960;20:37-46.

20. Brezo J, Paris J, Barker ED, et al. Natural history of suicidal behaviors in a population-based sample of young adults. *Psychol Med.* Nov 2007;37(11):1563-1574.

21. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T. The accuracy of prevalence estimations for suicide attempts. how reliably do adolescents and young adults report their suicide attempts? *Arch Suicide Res.* 2006;10(3):253-263.

22. Fendrich M, Warner V. Symptom and substance use reporting consistency over two years for offspring at high and low risk for depression. *J Abnorm Child Psychol.* Aug 1994;22(4):425-439.

23. Nock MK, Holmberg EB, Photos VI, Michel BD. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. *Psychol Assess.* Sep 2007;19(3):309-317.

24. Shaffer D, Scott M, Wilcox H, et al. The Columbia Suicide Screen: validity and reliability of a screen for youth suicide and depression. *J Am Acad Child Adolesc Psychiatry.* Jan 2004;43(1):71-79.

25. Goldney RD, Winefield AH, Winefield HR, Saebel J. The benefit of forgetting suicidal ideation. *Suicide Life Threat Behav.* Feb 2009;39(1):33-37.

26. Klimes-Dougan B. Screening for suicidal ideation in children and adolescents: methodological considerations. *J Adolesc.* Aug 1998;21(4):435-444.

27. Schaeffer NC. *Asking Questions About Threatening Topics: A Selective Overview. In The Science of Self-Report.* London: Lawrence Erlbaum Associates; 2000.

28. Klimes-Dougan B, Safer MA, Ronsaville D, Tinsley R, Harris SJ. The value of forgetting suicidal thoughts and behavior. *Suicide Life Threat Behav.* Aug 2007;37(4):431-438.

29. Blaney PH. Affect and memory: a review. *Psychol Bull.* Mar 1986;99(2):229-246.

30. Bower GH. Mood and memory. *Am Psychol.* Feb 1981;36(2):129-148.

31. Antikainen R, Hanninen T, Honkalampi K, et al. Mood improvement reduces memory complaints in depressed patients. *Eur Arch Psychiatry Clin Neurosci.* 2001;251(1):6-11.

32. Calev A. Affect and memory in depression: evidence of better delayed recall of positive than negative affect words. *Psychopathology.* 1996;29(2):71-76.

33. Ellwart T, Rinck M, Becker ES. Selective memory and memory deficits in depressed inpatients. *Depress Anxiety.* 2003;17(4):197-206.

34. Gonzalez HM, Bowen ME, Fisher GG. Memory decline and depressive symptoms in a nationally representative sample of older adults: the Health and Retirement Study (1998-2004). *Dement Geriatr Cogn Disord.* 2008;25(3):266-271.

35. Ilsley JE, Moffoot AP, O'Carroll RE. An analysis of memory dysfunction in major depression. *J Affect Disord.* Oct 9 1995;35(1-2):1-9.

36. Kindermann SS, Brown GG. Depression and memory in the elderly: a meta-analysis. *J Clin Exp Neuropsychol.* Oct 1997;19(5):625-642.

37. Livner A, Berger AK, Karlsson S, Backman L. Differential effects of depressive symptoms on prospective and retrospective memory in old age. *J Clin Exp Neuropsychol.* Apr 2008;30(3):272-279.

38. Murray LA, Whitehouse WG, Alloy LB. Mood congruence and depressive deficits in memory: a forced-recall analysis. *Memory.* Mar 1999;7(2):175-196.

39. Raphael KG, Cloitre M. Does mood-congruence or causal search govern recall bias? A test of life event recall. *J Clin Epidemiol.* May 1994;47(5):555-564.

40. Sachs-Ericsson N, Joiner T, Blazer DG. The influence of lifetime depression on self-reported memory and cognitive problems: results from the National Comorbidity Survey-Replication. *Aging Ment Health.* Mar 2008;12(2):183-192.

41. Kuyken W, Dalgleish T. Autobiographical memory and depression. *Br J Clin Psychol.* Feb 1995;34 (Pt 1):89-92.

42. Pollock LR, Williams JM. Effective problem solving in suicide attempters depends on specific autobiographical recall. *Suicide Life Threat Behav.* Winter 2001;31(4):386-396.

43. Williams JM, Barnhofer T, Crane C, et al. Autobiographical memory specificity and emotional disorder. *Psychol Bull.* Jan 2007;133(1):122-148.

44. Williams JM, Scott J. Autobiographical memory in depression. *Psychol Med.* Aug 1988;18(3):689-695.

45. Wisco B, Nolen-Hoeksema S. The interaction of mood and rumination in depression: effects on mood maintenance and mood-congruent autobiographical memory. *J Rat-Emo Cognitive-Behav Ther.* 2009;27.

Chapter 1

A systematic review of unstable self-reporting of suicidality over time: How reliable are our estimates?

Introduction

Suicidal behaviors, or "suicidality"[1-5], include suicidal ideation, plans, gestures, and attempts. Reliably capturing suicidality in the population is essential for both descriptive and analytic estimates. Reliable estimates are necessary to accurately quantify suicidality incidence and prevalence, which in turn, are required for trend surveillance and resource allocation. Reliable estimates are also critical for the assessment of predictors and correlates of suicidality, which inform prevention and intervention efforts. There is evidence however, that many of the current suicidality estimates suffer from a moderate to substantial degree of unreliability.[6-18]

Epidemiologic estimates of these complex and diverse behaviors primarily derive from respondent self-reports, often elicited via a single survey question.[19-26] While self-reports are in many ways advantageous for etiologic inquiry, they are also prone to reporting errors, which to varying degrees, may compromise their reliability. That is, respondents may report a history of suicidal behaviors, and when queried again, not report such history. Conversely, respondents may not report past suicidality at baseline, but report a history at follow-up. In large epidemiologic surveys, it is infeasible to determine the validity of any one report and therefore it is unknown if unstable reporting over time results from underreporting true attempts, or falsely reporting attempts at one time point. However, a measure can only be as valid as it is reliable, therefore quantifying the reliability of suicidality self-reports will shed light on the measures' maximum degree of validity.

Reliability estimates may vary across studies for two main reasons. First, estimates may vary according to factors that influence the likelihood of reporting suicidality, such as the particular suicide construct (e.g. ideation/attempts, period of observation.), sample (e.g. sample type, demographics) and/or suicide question characteristics (e.g. question format, wording).

These traits could be considered factors that affect the probability of a respondent reporting suicidality at any one time point, and over time. However, reliability estimates may also vary according to factors that statistically influence the reliability estimate itself, but not respondent reporting, such as sample size, sampling scheme, and prevalence of suicidality. This review will only consider the former, as these are factors that reflect true underlying causal mechanisms, and could potentially be manipulated by investigators to better ensure valid responses to suicidality questions.

Researchers in the field have posited causal mechanisms underlying "unstable reporting" but none have been rigorously assessed. The most common explanations proposed are: 1) recall failure; 2) reinterpretation; and 3) social desirability effects. For example, respondents may fail to recall a past attempt because they suppress the memory of it as an adaptive coping mechanism, allowing the individual to heal and move on.[12] It is also possible that unstable reporting is a result of simple recall failure,[8,14] particularly when questions refer to events that took place many years ago. This recall failure may be influenced by factors such as the time since suicidality, context in which the question is embedded, severity of suicidality, history of a mental disorder, or even current mood or functioning of the respondent at the time of reporting. A second causal mechanism for unstable reporting may be "reinterpretation".[8,12,13] That is, a respondent may have reported a history of suicidality at baseline, but later, reinterpreted that event as not of suicidal nature, and therefore not reported it at follow-up. In other words, the nature of the event as the respondent sees it has changed over time. Like recall, the likelihood of reinterpretation may also be affected by factors such as severity of the suicidality and the respondent's mood at the time of reporting. Lastly, respondents may remember the suicidality, interpret it as suicidal, but decide not to report it during an interview[8,13] for reasons of shame or

social desirability.[8,12,27] The first step in elucidating causal mechanisms is assessing potential correlates, or predictors, of unstable reports and reporters.

A systematic review of the literature is necessary to examine test-retest reliability estimates of self-report suicidality measures, as well as correlates of reliable reports within and across studies. This literature has yet to be examined and synthesized systematically, and may have important implications for future research. Reviewing these estimates will potentially allow us to ascertain characteristics of the suicide construct, study, and/or sample, which may yield the most reliable estimates. This will provide researchers with insight into the reliability of suicidality measures they currently use, and may help guide them in the development of suicidal measures for future studies.

Methods

Articles in this systematic review were identified through Ovid Medline/Pubmed and PsycInfo databases. The search criteria comprised any combination of the following, within the Title and/or Abstract: (1) "suicid$" (suicide, suicides, suicidal, suicidal ideation, suicidality) or "suicide attempt$" (suicide attempt, suicide attempts) and (2) "reliability" or "recall" or "test-retest". This search yielded 622 unique publications whose abstracts were selected for further review to determine eligibility. Articles were included in the final analysis if they met the following criteria: (1) peer-reviewed; (2) original research; (3) published in the last 40 years (1970 -2010); (4) written in English; (5) contained a measure of past suicidality; and (6) included an analytic or descriptive measure of test-retest reliability of a suicide construct. Articles were excluded if they: (1) reported an alternate type of reliability estimate (i.e. internal consistency of scales); (2) reported the test-retest reliability of entire scale, rather than of individual items; (3)

were not self-reported measures of suicidal constructs (i.e. medical record studies); and (4) were measures of current suicidality. All decisions on inclusion and exclusion criteria were based on the parameters that characterize suicidality incidence, prevalence, and predictor estimates reported in the literature. That is, because the estimates utilized for surveillance and suicidality prevention and intervention efforts are most often self-report measures, elicited via one question and regarding past suicidality, these are the types of measures which are most pertinent and relevant to review in regards to their reliability. Review of paper citations and review articles yielded 3 additional articles that met inclusion/exclusion criteria.

Results

Thirteen papers met the eligibility criteria for this systematic review.[6-18] This review will first delineate the specific constructs measured, their respective reliabilities, and any correlates of unstable reports, as reported *within* studies. Then, I will explore ways in which the studies differed according to suicide construct, sample, study design, and research instrument characteristics that may have led to heterogeneity of reliability estimates; gleaning potential correlates of reliability through examination *across* studies.

I. Reliability estimates of suicidality

The thirteen papers measured five different suicide constructs (i.e. ideation, plans, gestures, attempts, "suicidal content"), examined three observation periods (i.e. lifetime, 3-month, and 12-month), and assessed reliability with three types of estimates (i.e. traditional Kappa coefficient or conditional Kappa coefficient, total percent agreement, and percent denial of suicidality at follow-up).

Suicide constructs

Nine papers reported the reliability of suicidal ideation; three reported 12-month estimates; one reported a 3-month estimate[17]; and five reported lifetime estimates (Table 1A).[6,7,9-12,15-18] Seven papers reported the reliability of suicide attempts; two reported 12-month estimates and five reported lifetime estimates (Table 1B).[6-10,16,17] Lastly, four papers reported the reliability of other suicide constructs (e.g. plans, gestures, and "suicidal content"); of these, one reported 12-month estimates and three reported lifetime estimates (Table 1C).[6,13,14,16]

Measures of reliability

The widely accepted index of test-retest reliability, the traditional Kappa coefficient,[28] was used by seven studies.[6,9,13-18] Kappa is the proportion of agreement corrected for chance agreement.[28] Landis and Koch suggested useful "benchmarks" for evaluating the degree of reliability according to the Kappa coefficient: < 0.00 = poor; 0.00-0.20 = slight; 0.21-0.40 = fair; 0.41-0.60 = moderate; 0.61-0.80 = substantial; 0.81-1.00 = almost perfect.[29]

Eight studies reported other indicators of reliability[*].[7-12,14,18] One paper[10] reported "percent agreement" as their main indicator of reliability, which is the percentage of respondents who reported suicidality at time 1 (T1) and time 2 (T2), as well as those who did not report suicidality at both time points. Percent agreement however, unlike the kappa coefficient, does not take into account the proportion of agreement that may have occurred due to chance alone. One paper[12] calculated a "conditional kappa coefficient".[30] This kappa coefficient, while corrected for chance agreement, was conditional on reporting suicidality at T1, and therefore gave no weight to those who did not report suicidality at T1. Hence, respondents who reported past suicidality at T1, but not at T2, were considered in the reliability estimate; those who did not

[*] Studies did not provide raw data with which to calculate a Kappa coefficient

report past suicidality at T1, but reported it at T2, were not considered. Lastly, six studies[7-9,11,12,18] simply reported the percent of respondents reporting a history of suicidality at T1, who did not report a history of suicidality at T2. This descriptive estimate is often described as the "percent recall" or conversely, the "percent denial" or "percent of recanting"[**]. As with the conditional kappa, these indicators only take into consideration respondents who reported a history of suicidality at T1, however, unlike the conditional kappa coefficient, this estimate does not correct for chance agreement.

Reliability of suicidal ideation

Nine papers reported a measure of reliability for suicidal ideation (Table 1A). [6,7,9-12,15-18] Six[6,9,15-18] reported a kappa or a conditional kappa coefficient, hence correcting for chance agreement. The lowest kappa coefficient was reported for *lifetime* ideation. In a general population sample of Australian adult twins, Statham et al.[18] reported the reliability for men and women separately; men: K= 0.10, women: K=0.11. The second lowest kappa was a conditional kappa (K= 0.46; 0.25-0.67) and was reported by Fendrich et al.[9]. Fendrich estimated the reliability of *lifetime* suicidal ideation, assessed 2 years apart, and in a mostly youth (<18yr) and mixed risk (low and high) sample. Shaffer et al.[17] reported the third lowest kappa (K= 0.48), an estimate of *3-month* ideation, in a convenience sample with assessments only 8 days apart. The highest kappa coefficient reported for ideation was by Koziol-McLain et al.[15] (K= 0.79, 0.59-0.99). Koziol-McLain et al. estimated the reliability for *12-month* suicidal ideation, and in a population-based sample, with assessments 7-28 days apart. The second highest kappa coefficient (K=0.74) was also for *12-month* ideation, assessed by Brener et al.[6] among a convenience sample with

[**] Studies either reported "percent denial of suicidality at follow-up" or "percent recall at follow-up". For consistency and comparative purposes, all results in this review are presented in terms of "percent denial" at follow-up.

assessments approximately 16 days apart. Lastly, the third highest kappa coefficient (K=0.70), was reported for *lifetime* ideation by Nock et al.[16] in a mixed (clinical and non-clinical) sample assessed 6 months apart.

Five[9-12,18] of the suicidal ideation studies reported other types of reliability measures, which did not take into account chance agreement. Flisher et al.[10] reported an observed agreement of 92.1% for a *12-month* ideation estimate, among a convenience sample, with assessments 10-14 days apart. The remaining four studies reported "percent denial", all regarding lifetime ideation, and with relatively comparable degrees of denial. Goldney et al. [11] was the first to report unstable reporting of lifetime ideation, and found that 40% of ideators at T1, did not re-report (i.e. denied) ideation at T2, 4 years later. When Goldney et al.[12] reanalyzed their data using a looser definition of ideation, they reported an even greater percent of denial (50%). Similarly, Fendrich et al.[9] reported a percent denial of 50%. Lastly, Statham et al. [18] reported percent denial separately for men and women, and found that 44% of female T1 ideators, and only 32% of male T1 ideators denied ideation at T2. In summary, the reliability of suicidal ideation reports, as measured by Kappa and reported by six studies, ranged from slight to moderate (K = 0.10- 0.79); the reliability of suicidal ideation reports as measured by "percent of denial" and reported by four studies, ranged from 32-50%.

Reliability of suicide attempts

Seven studies reported reliability estimates for suicide attempts (Table 1B)[6-10,16,17] Two of these studies reported comparable coefficients; Fendrich reported a conditional kappa coefficient of 0.58 (0.19-0.97) and Shaffer et al. reported a kappa of 0.58, both in regards to *lifetime* attempts. Brener et al. reported two reliability estimates, one regarding any attempts in the past *12 months*

(K=0.73), and one regarding attempts that specifically led to injury in the past *12 months* (K=0.52). Lastly, Nock et al.[16] reported a kappa of 0.80, for *lifetime* attempts "that carried at least some intent to die". Four studies chose other indicators of reliability, which did not adjust for chance agreement [7-10] Flisher et al.[10] reported an observed agreement of 94.7% for attempts in the past *12 month*s. Three of the four studies reported "percent denial", to comparable degrees, and all of *lifetime* estimates; Brezo et al. (31%), Christ et al. (33%); and Fendrich et al. (40%). In summary, the reliability of suicide attempt reports, as measured by Kappa and reported by four studies, ranged from fair to moderate (K= 0.52 – 0.80); the reliability of suicide attempt reports as measured by "percent of denial" and reported by three studies, ranged from 31-40%.

Reliability of other suicide constructs

Four studies[6,13,14,16] examined the reliability of other suicide constructs; namely, plans, gestures, and "suicidal content" (a construct composed of ideation, thoughts of death, wishes for death, and attempts) (Table 1C). Nock et al. found the *lifetime* reliability of suicide attempt plans to be moderate (K=0.71), and the *lifetime* reliability of suicidal gestures (something done to lead others to believe that you want to kill yourself when you really had no intention of doing so) to be slight (K=0.25). Brener et al. examined the *12-month* reliability of suicidal plans, and found it to be moderate (K=0.67). Lastly, Klimes-Dougan reported the reliability of "suicidal content" in two studies. The first study was conducted in a sample with an average age of 15, and reported a K=0.42 among the younger adolescents, and a K=0.60 among the older adolescents. Six and a half years later, when the average age of the participants was 22, the Kappa was 0.57. In summary, the reliability of other suicidal constructs, as measured by Kappa and reported by four studies, ranged from moderate to substantial (K=0.42-0.71).

Across the various suicide constructs, there was an overall fair to moderate degree of reliability, suggesting a true need for improvement in the measurement and measures of suicidality. The logical next step is to determine what evidence exists concerning factors that might lead to unreliable reporting.

II. Correlates of stable reporting examined within reviewed studies

Only five studies examined at least one correlate of report stability.[8,12-14,18] Klimes-Dougan et al. (1998) found five factors related to greater stability in reporting lifetime "suicidal content" 3-6 years after baseline report; 1) high-risk for depression (higher stability among children of depressed mothers vs. children of non-depressed mothers); 2) female gender; 3) older age; 4) current psychological distress (Symptom Checklist-90, Global Symptom Index); and 5) low self-concept/esteem (Perceived Competence Scale for Children). Klimes-Dougan et al. (2007) re-interviewed the cohort 7 years later and found that stable reporters were more likely to have a current depressed mood (Beck Depression Inventory), and overall lower functioning (Global Assessment of Functioning; Responses to Depression Scale (rumination and coping). Goldney et al.[12] assessed the reporting of lifetime suicidal ideation, and found that stable reporting (higher reliability) was greater amongst those with more ideation at baseline and at follow-up, lower self-esteem, and greater hopelessness, depressive affect and negative mood (GHQ; Srole Anomia Scale; the Negative Mood Scale). Christ et. al.[8] assessed the stability of lifetime attempt reports separated by 4 years, and found that stable reporting was associated with 5 factors; 1) male gender; 2) older age at baseline; 3) any lifetime depressive disorder (M-CIDI) (OR, 8.4; 95% CI, 1.1-61.5), 4) any lifetime somatoform disorder (M-CIDI) (OR, 16.1; 95% CI, 1.5-173.2); and 5)

greater mental health disorder comorbidity (3.5 disorders vs. 2.8 disorders on average). Lastly, a study by Statham el al[18] reported "percent denial" and kappa coefficients separately for men and women (Percent denial; women: 44%, men: 32% and Kappa; women: 0.11, men: 0.10), however did not show tests for significance across gender strata. According to the "percent denial", it appears that males were more reliable than females; however, according to the Kappa coefficients, which take into account change agreement, there appears to be no difference by gender.

In summary, only five of the thirteen studies examined at least one correlate of stable reporting. Three studies reported contradictory findings in regards to gender, two studies found older age to be associated with more stable reports, and four studies found that respondents who had lower levels of functioning, particularly greater mental distress (either current or lifetime), were more stable in their reports.

III. Sources of heterogeneity across studies

This section will examine variation in suicide construct, sample, study design, and suicide question characteristics across the thirteen studies. Examining variations across, rather just within studies, may illuminate other potential correlates of unstable reporting. Potential correlates of interest, as described in the literature, include the severity of suicidality, the time since suicidality, and social desirability effects. For example, if we were to find that lifetime estimates were less reliable than 12-month estimates, time since suicidality would be implicated as a potential correlate of unstable reporting. These influential factors and potential underlying causal mechanisms, such as recall and reinterpretation, will be explored in the discussion section. In this section however, we first describe variations in construct, sample, study design, and

suicide question characteristics across the thirteen studies, and their potential association with reliability estimates, before drawing upon them as a whole to determine overarching correlates and underlying mechanisms.

Suicide construct characteristics

Construct

The degree of unstable reporting differed according to suicide construct. Across the different constructs, the average degree of reliability from highest to lowest ranked as follows: plans (avg K= 0.69), attempts (avg K=0.64), "suicidal content" (avg K= 0.53), ideation (avg. K= 0.48), and gestures (avg. K= 0.25).

Observation period

The degree of unstable reporting also differed according to period of observation, dependent on suicidal construct. Among the six studies that assessed ideation reliability with a Kappa coefficient, there was an apparent pattern; three of the four lowest coefficients were *lifetime* estimates (range: 0.10-0.70), and the two highest coefficients were *12-month* estimates (0.74 and 0.79). Among studies that assessed attempt reliability however, there was no clear pattern with regards to observation period; both the highest and lowest reliability estimates reported were *lifetime* estimates. Among the studies that examined the reliability of other suicide constructs (i.e. plans, gestures, "suicidal content"), again, no pattern was evident. In summary, period of observation only appeared to influence the reliability estimates of ideation, across which 12-month estimates were more reliable than lifetime estimates.

Sample Characteristics

Age

There was very little variation in age across reviewed studies. With the exception of two studies, all reviewed reliability estimates were among youth or young adult populations. The age at follow-up among the youth studies ranged from 14-23; six studies[6,9,10,13,16,17] in populations of <18 years of age and five studies[7,8,11,12,14] in populations between the ages of 18-23. The two studies that assessed reliability among adults, were regarding reporting of suicidal ideation. Statham et al.[18] assessed the reliability of lifetime ideation in a population with an average age of 43 at follow-up; Koziol et al.[15] assessed 12-month suicidal ideation in a population with an average age of 45 at follow-up. There were no reliability estimates of suicide attempts or other suicide constructs among adult populations captured within this literature review. Given the very limited age range, any patterns with regard to age and degree of reliability were difficult to assess.

Gender

Eleven[6-8,10-17] out of the thirteen studies reported the gender ratio for their samples. With the exception of one[10], all studies had a higher percentage of female respondents in their sample, with 'percent female' ranging from 53-77%. When stratified by type of reliability statistic, there was no apparent relationship between gender ratio and average kappa coefficient across studies. However, among those studies who reported "percent denial"[7,8,11,12], there appeared to be a negative linear trend between percent female and percent denial; that is, it appears that women were more consistent in their reports.

Sample type

There were three main sample types across the thirteen reviewed papers. Three studies[6,15,18] used population-based samples (avg. K=0.52); 2 of them US samples, and 1 Australian. Five studies[8,10-12,17] used community samples (avg. K=0.53); 1 in Munich, Germany, 2 in Adelaide, Australia, 1 in Capetown, South Africa, and 1 in the New York metropolitan region. Lastly, five studies[7,9,13,14,16] used mixed samples (i.e. low and high-risk; clinical and non-clinical) (avg. K=0.59); 1 in Quebec, Canada, and 4 in northeastern metropolitan US cities. There was no strong pattern with regard to type of sample and degree of reliability; however, the reliability of suicidality among mixed samples may be slightly higher than among population and community samples.

Study design characteristics

Interval between assessments

There was a wide range in time between assessments across the thirteen studies. Four studies[6,10,15,17] had a time interval of less than one month (avg. K= 0.67); three studies[8,9,16] had a time interval of 6 months to 2 years (avg. K= 0.58); and six studies[7,11-14,18] had a time interval of greater than 2 years (avg. K=0.39). Across the three time intervals, the average kappa coefficient from highest to lowest ranked as follows: less than one month, six months to two years, more than two years. There appeared to be a negative linear trend between the time between assessments and reliability; that is, as the time between interviews increases, the likelihood of unstable reporting increases, thereby decreasing the reliability.

Interview format (in-person self-administered, interviewer-administered face-to-face, mail, telephone)

There was a range of interview formats used across the thirteen studies (Table 1D). Three studies[6,10,17] used in-person self-administered surveys (avg. K=0.64); six studies[7-9,13,14,16] used interviewer-administered face-to-face interviews (avg. K=0.57); three studies[11,12,18] used self-administered mail surveys (avg. K=0.11); and three studies[15,16,18] used a interviewer-administered telephone survey (avg. K=0.56). Two of the studies[16,18] used two different formats across assessments. Across the four interview formats, the average kappa coefficient from highest to lowest ranked as follows: in-person self-administered, interviewer-administered face-to-face, interview-administered telephone survey, self-administered mail surveys.

Suicidality instruments and questions (Table 1D)

Instrument

There were three main contexts in which the suicidality question was presented to the respondents. In five studies[6,10-12,15], the suicidality question was embedded within either a general health questionnaire, or a risk behavior survey (avg. K=0.73, avg. percent denial = 45%); in six studies[7-9,13,14,18], the suicidality question was embedded within a diagnostic module (usually depression module) of a structured or semi-structured interview (avg. K=0.42, avg. percent denial=36%); and in three studies[7,16,17], the suicidality question was part of a larger suicide assessment scale or interview (avg. K=0.63, avg. percent denial=31%). Across the three contexts, the average kappa coefficient from highest to lowest ranked as follows: General health questionnaire, suicide scale or interview, and diagnostic module. The average percent denial

from highest to lowest ranked as follows: General health questionnaire, diagnostic module, and suicide scale or interview.

Suicidality Questions

There was large variation in phrasing of questions across constructs, and within constructs. The questions varied in four main ways. First, they varied in the language used to describe the suicidal behavior. For example three of the studies used the term "suicide" [6-8] (e.g. Have you ever considered/ attempted suicide?); three studies[11,12,18] used the phrase "thoughts of doing away with yourself"; while the majority referred to "hurting or killing yourself". Second, questions varied with regard to severity of suicidality elicited. For example, while most questions did not specify a level of severity, four studies[6,10,15,16] asked the respondent about "serious" ideation or attempts (e.g. "Have you ever seriously though about trying to hurt yourself in a way that might have resulted in your death?"; "If you attempted suicide during the past 12 months, did any attempt result in an injury, poisoning, or overdose that had to be treated by a doctor or nurse?"; and "Have you ever made an actual attempt to kill yourself in which you had a least some intention to die?"). Third, studies varied with regard to the context in which the respondent was asked to consider the questions. For example, most questions did not require the respondent to have past mood symptoms in order to answer the suicidality questions. However, two studies[8,9] first required positive endorsements to at least one of two Major Depression diagnostic "gateway" questions in order to be asked the suicide questions, and therefore be included in the reliability sample; one of these studies[9] went further to specify that when asked about the suicidality question, they should think about their worst-past episode of depressive

mood. Similarly, one study[18] asked the respondent to endorse the statement "*Recently I have been so depressed* that I have thought of doing away with myself".

In summary, suicidality questions varied widely across and within constructs. Due to the large variation, it is difficult to draw conclusions or patterns regarding question format and reliability. However, it should be noted that the two highest reliability coefficients across the thirteen reviewed studies ("Have you ever made an actual attempt to kill yourself in which you had a least some intention to die?", K=0.80[16] and "In the past year, have you ever seriously thought about trying to hurt yourself in a way that might have resulted in your death?", K=0.79[15]) reported on questions that contained some reference to severity and the respondent's intention to die.

Discussion

Drawing upon the correlates reported within studies, as well as sources of heterogeneity across studies, we posit five main correlates of unstable reporting; 1) severity of suicidality, 2) time since the suicidal event, 3) social desirability, 4) mood context, and 5) suicide construct validity. Further, we believe that these five correlates lead to unstable reporting by way of four underlying causal mechanisms; recall failure, reinterpretation, denial*, and lack of construct comprehension. Below, we describe the findings from this literature review that implicate each of the five main posited correlates, as well as potential underlying causal mechanisms associated with each correlate.

* Denial in this context refers to "conscious denial". That is, the respondent recalls the suicidality, correctly interprets it as suicidal, but chooses not to report it. This is not to be conflated with "percent denial" which is simply a generic label placed on respondents who report suicidality at baseline, but not at follow-up. "Percent denial" does not indicate that the respondent necessarily "consciously denied" the suicidality.

Severity of suicidality

There are several indications that the severity of the suicidality may affect the likelihood of reporting suicidality at any one time point, and therefore the stability of reporting over time. First, the review suggests that more serious, or advanced suicide constructs (i.e. plans and attempts) may be more reliable compared with less serious constructs (i.e. ideation and gestures). Second, multiple studies found that respondents with lower levels of functioning, particularly greater mental distress or a mood disorder (either current or lifetime), were more stable in their reports. It has been shown that suicidality that occurs within the context of a depressive disorder is often more medically serious.[31-34] Therefore depressed mood, current or lifetime, may be a marker for severity of suicidality. Third, this review provided some evidence that mixed samples, comprising clinical (i.e. treated) or high-risk individuals, may produce slightly higher reliability estimates compared with community or general population samples; a pattern corroborated by findings in the MDD literature.[35,36] Clinical or high-risk individuals are more likely to have higher severity of suicidality compared with non-clinical individuals, therefore sample type may also be a marker for severity of suicidality. Lastly, we found that questions specifying *serious* ideation or attempts had amongst the highest reliability estimates.

Multiple studies examining the reliability of depressive symptoms and disorders have found that the severity of the respondent's depressed mood, and history of psychiatric treatment (also a proxy for severity) predicted stability in reporting[9,37-43]. It is therefore plausible that severity of suicidality is a predictor of stable suicidality reporting. If this is indeed true, we posit that the underlying causal mechanism of this relationship is recall failure and/or reinterpretation. That is, less severe suicidality is likely to be forgotten more easily or reinterpreted as non-suicidal, and consequently less consistently reported.[9]

Time since suicidality

There are several indications from this review, and the extant literature[40], that the amount of time passed between the suicidal behavior and reporting may influence the likelihood of reporting the suicidality. First, this review revealed that in general, the reliability of ideation was higher for ideation occurring within the last year, compared with measures of lifetime ideation. Studies have shown that reliability estimates of lifetime diagnoses for Major Depression and other mood disorders are often lower than current, past month, or 12-month estimates,[35,44-47] lending credence to our finding. Second, this review found that as time between assessments increased, reliability decreased; a phenomenon also documented in the MDD literature[46,48,49]. Time span can be considered a proxy for time since suicidality since as the time span increases, inevitably, so does the time between the suicidal behavior and reporting.

If increased time since suicidality did indeed decrease the reliability of reports, it is plausible that recall failure and reinterpretation could once again be underlying causal mechanisms. That is, as more time passes between a respondent's suicidal behavior and reports thereof, the more likely this behavior is to be forgotten or reinterpreted.[40]

Social desirability effects

It is plausible that social desirability may influence the likelihood of reporting suicidality; the underreporting of psychiatric symptoms, such as depression, has been shown in the literature to be positively associated with respondent social desirability per the Crowne-Marlowe scale.[40] An indication that social desirability effects may influence reporting of suicidality is our finding that, with the exception of the mail survey[*], the in-person self-administered surveys appeared to elicit

[*] The average kappa coefficient associated with the mail surveys was only based on one study, since the other two mail surveys utilized 'percent denial', and thus may not be representative estimate.

more stable responses compared with interviewer-administered face-to-face and interviewer-administered telephone surveys. Perhaps because the latter formats require respondents to admit aloud actions that they may feel embarrassed, shameful, or perhaps regretful about, they choose to deny it to the interviewer. A well-established literature that suggests that self-administration (vs. administration by an interviewer) increases reporting of socially undesirable behaviors,[50,51] particularly reports of mental health symptoms,[52] corroborates our findings.

Mood context

There are several ways in which the context that suicidality questions are worded, or answered, specifically in relation to respondent depression, may impact the likelihood of reporting. For example, in some studies, respondents were first asked about past experiences of depressive symptoms ("gateway questions"), which they had to first endorse before being asked the suicidality question. Other studies worded their questions to specifically elicit suicidal events that took place during a depressive episode. It is plausible that being forced to reflect upon prior depressive symptoms, primes the respondent, helping cue recall of the suicidal event.[8,40] Alternatively, perhaps being required to first acknowledge experiences of depression decreases the probability of reinterpreting the event as non-suicidal.

The context in which the respondent answers the question may also affect the likelihood of reporting the suicidality. The most compelling and consistent finding of this review, four studies reported that respondents with *current* depressed mood or poor functioning at the time of the interview were more likely to report suicidality at that interview. These findings are supported by mood-congruent theory[53] which posits that currently depressed persons are more likely to recall previous depressed states or negatively-valenced events. Current distress may

evoke reminders of past suicidality, or amplify the perceived severity of it, thereby facilitating recall, and resulting in an increased probability of reporting.[40] It is also plausible that the respondent's current depressed mood may make them more likely to interpret an event as suicidal, and therefore report it as such. These findings have also been well documented within the MDD reliability literature.[39-41,43,46,54]

Suicide construct validity

Lastly, respondent comprehension of suicidal constructs may vary; there is variation, particularly among laypersons, about what constitutes suicidality, especially a suicide attempt. For example, some respondents may conflate constructs such as suicidal gestures (meant as a cry for help, but without intention to die) and non-suicidal self-harm (e.g. cutting, burning), with attempts, which is supposed to carry an intention to die. Therefore, it is highly probable that wording used to describe the suicidal construct influences the likelihood of reporting. A more explicit description, such as "killing yourself", could be considered less ambiguous, and therefore may be answered more consistently over time. Some studies have shown that questions asking about specific actions and behaviors rather than labels, elicit more endorsements because they do not require the respondent to make the judgment as to whether their actions fit the label.[55] Three studies in this review found that an older youth age was predictive of stable reporting of suicidality or depressive symptoms, compared with younger youth age. It is plausible that older youth may have a clearer, and therefore more consistent understanding of what constitutes different suicidality constructs.

Summary, limitations, and recommendations

In summary, our review indicates that estimates of suicidality suffer from an overall moderate degree of unreliability, but dependent on suicide construct, sample, study design, and instrument/wording characteristics. After examining heterogeneity of these characteristics within and across studies, we posit four plausible underlying reasons for unstable reporting; recall failure, reinterpretation, conscious denial, and lack of construct comprehension. Further, we propose that the likelihood of these mechanisms are influenced by factors such as the severity of suicidality, time since the suicidal event, social desirability, mood context, and suicide construct validity.

The main limitation of this review was the relatively small number of studies available that examined and reported the test-retest reliability of suicidality estimates. Specifically, the wide variations across a small number of studies made it difficult to isolate which characteristics were truly influencing the reliability estimate. For example, when examining potential patterns between reliability and factor X (e.g. observation period) across studies, all other factors (e.g. construct, age, sample type) were not evenly distributed, hence not held constant. While studies that examined correlates of stable reporting were optimal for controlling other covariates, only five of the thirteen studies examined at least one correlate of unstable reporting. Studies were also heterogeneous in regards to reliability statistics used. Since the majority of studies used the Kappa or conditional Kappa statistic, and they accounted for chance agreement, we tended to rely more on kappa coefficients, and therefore those studies, to examine correlates across studies.

Another limitation of this study was the lack of a few likely important correlates. We used the variables at hand to construct possible influential factors, such as time since suicidality, severity of suicidality, and social desirability. However, actual measures of these factors would

have been ideal, and allowed us to test the concurrent validity of some of our conclusions. Lastly, we did not consider factors that influence the reliability estimate, but not respondent reporting behaviors, such as the sample size, sampling scheme, and prevalence of suicidality. In particular, Kappa is strongly influenced by the prevalence of the outcome; specifically, as the prevalence of suicidality increases, given the same frequency of unstable reporting, Kappa decreases.[56,57] This occurs because as the prevalence of an outcome increases, the probability for agreement simply due to chance increases, which Kappa in essence, deducts from the final reliability coefficient. Therefore differing prevalences of suicidality across studies may have led to variation in reliability estimates that was not necessarily reflective of differing report stability.

Based on this review, we recommend that future research consider the following issues to better characterize the reliability of suicidality measures, and comprehensively understand the correlates of unstable reporting. There is of yet, no reliability estimate reported for suicide attempts in an adult population. Based on reports that the lifetime prevalence of suicide attempts among adults (1.9%-4.6%) is lower than even the 12-month prevalence among adolescents (7.3%-10.6%)[25], there is ancillary evidence suggesting that adult reports also suffer from a substantial degree of unreliability. Furthermore, future studies should include a larger baseline sample of individuals with suicidality. Since suicidality, particularly attempts, is a rare phenomenon, the reliability estimates were almost all based on extremely small samples; with the exception of one study, all samples contained fewer than 69 baseline attempts on which to base their estimates. Such small sample sizes make estimates particularly vulnerable to random error, and hence imprecision. In addition, it is difficult to explore correlates of unstable reporting due to limited power capabilities.

There are also methodological issues to consider for future studies. The kappa coefficient, a reliability estimate that accounts for chance agreement, was not calculated for five of the thirteen studies, and is critical for all future studies. Secondly, many studies calculated a "percent denial" or a conditional kappa, which only considers those who report suicidality at baseline. However, respondents also "newly endorse" suicidality. That is, respondents may not report a history of suicidality at baseline, but report it at follow-up, and state that the suicidality did not occur between interviews, thereby appropriately excluding new-onset suicidality. Both "deniers" and "new endorsers" at follow-up are unstable reporters and should be included in estimate calculation. Lastly, a wider array of potential correlates of unstable reporting would benefit future studies. A factor that consistently emerged from our reviewed studies was respondent history of, or current depressed mood. Our understanding of causal mechanisms underlying unstable reporting would clearly benefit from a more in-depth investigation of this seemingly mood-congruent reporting pattern. We believe future research should include measures of severity (including treatment history), recency of suicidality, and social desirability in order to further test speculated reasons for unstable reporting.

In conclusion, this review found considerable unreliability in reports of suicidality, which varied depending on construct, sample, study design, and suicide question characteristics. This unreliability is likely contributing to the underestimation of the prevalence and incidence of suicidality reported in cross-sectional surveys, and may also bias estimates between predictors and attempts in the population. We posit that factors related to the severity and recency of suicidality, social desirability, mood context, and construct comprehension may influence the likelihood of reporting suicidality, and should be considered when utilizing suicidality measures in secondary data sources or future data collection.

References

1. Goldblatt MJ, Maltsberger JT. Self-Harming Behavior and Suicidality: Suicide Risk Assessment. Suicide Life Threat Behav 2011.

2. Johnson J, Wood AM, Gooding P, Taylor PJ, Tarrier N. Resilience to suicidality: The buffering hypothesis. Clin Psychol Rev 2011.

3. Reeves RR, Ladner ME. Antidepressant-induced suicidality: an update. CNS Neurosci Ther 2010;16:227-34.

4. Fawcett JA, Baldessarini RJ, Coryell WH, Silverman MM, Stein DJ. Definition and management of suicidality in psychiatric patients. J Clin Psychiatry 2009;70:e38.

5. Bursztein C, Apter A. Adolescent suicide. Curr Opin Psychiatry 2009;22:1-6.

6. Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EC, Ross JG. Reliability of the 1999 youth risk behavior survey questionnaire. J Adolesc Health 2002;31:336-42.

7. Brezo J, Paris J, Barker ED, et al. Natural history of suicidal behaviors in a population-based sample of young adults. Psychol Med 2007;37:1563-74.

8. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T. The accuracy of prevalence estimations for suicide attempts. how reliably do adolescents and young adults report their suicide attempts? Arch Suicide Res 2006;10:253-63.

9. Fendrich M, Warner V. Symptom and substance use reporting consistency over two years for offspring at high and low risk for depression. J Abnorm Child Psychol 1994;22:425-39.

10. Flisher AJ, Evans J, Muller M, Lombard C. Brief report: Test-retest reliability of self-reported adolescent risk behaviour. J Adolesc 2004;27:207-12.

11. Goldney RD, Smith S, Winefield AH, Tiggeman M, Winefield HR. Suicidal ideation: its enduring nature and associated morbidity. Acta Psychiatr Scand 1991;83:115-20.

12. Goldney RD, Winefield AH, Winefield HR, Saebel J. The benefit of forgetting suicidal ideation. Suicide Life Threat Behav 2009;39:33-7.

13. Klimes-Dougan B. Screening for suicidal ideation in children and adolescents: methodological considerations. J Adolesc 1998;21:435-44.

14. Klimes-Dougan B, Safer MA, Ronsaville D, Tinsley R, Harris SJ. The value of forgetting suicidal thoughts and behavior. Suicide Life Threat Behav 2007;37:431-8.

15. Koziol-McLain J, Brand D, Morgan D, Leff M, Lowenstein SR. Measuring injury risk factors: question reliability in a statewide sample. Inj Prev 2000;6:148-50.

16. Nock MK, Holmberg EB, Photos VI, Michel BD. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. Psychol Assess 2007;19:309-17.

17. Shaffer D, Scott M, Wilcox H, et al. The Columbia Suicide Screen: validity and reliability of a screen for youth suicide and depression. J Am Acad Child Adolesc Psychiatry 2004;43:71-9.

18. Statham DJ, Heath AC, Madden PA, et al. Suicidal behaviour: an epidemiological and genetic study. Psychol Med 1998;28:839-55.

19. Baca-Garcia E, Perez-Rodriguez MM, Keyes KM, et al. Suicidal ideation and suicide attempts in the United States: 1991-1992 and 2001-2002. Mol Psychiatry 2008.

20. Bolton JM, Belik SL, Enns MW, Cox BJ, Sareen J. Exploring the correlates of suicide attempts among individuals with major depressive disorder: findings from the national epidemiologic survey on alcohol and related conditions. J Clin Psychiatry 2008;69:1139-49.

21. Dube SR, Anda RF, Felitti VJ, Chapman DP, Williamson DF, Giles WH. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: findings from the Adverse Childhood Experiences Study. Jama 2001;286:3089-96.

22. Garroutte EM, Goldberg J, Beals J, Herrell R, Manson SM. Spirituality and attempted suicide among American Indians. Soc Sci Med 2003;56:1571-9.

23. Ialongo N, McCreary BK, Pearson JL, et al. Suicidal behavior among urban, African American young adults. Suicide Life Threat Behav 2002;32:256-71.

24. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999;56:617-26.

25. Nock MK, Kessler RC. Prevalence of and risk factors for suicide attempts versus suicide gestures: analysis of the National Comorbidity Survey. J Abnorm Psychol 2006;115:616-23.

26. Substance Abuse and Mental Health Services Administration OoAS. The NSDUH Report: Suicidal Thoughts and Behaviors among Adults. Rockville, MD; 2009.

27. Schaeffer NC. Asking Questions About Threatening Topics: A Selective Overview. In *The Science of Self-Report* London: Lawrence Erlbaum Associates; 2000.

28. Cohen J. A coefficient of agreement for nominal scales. Educational Psychological Measures 1960;20:37-46.

29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

30. Bishop YMM, Fienberg SE, Holland PW. Discrete multivariate analysis: Theory and Practice. Cambridge, MA: MIT Press; 1975.

31. Astruc B, Torres S, Jollant F, et al. A history of major depressive disorder influences intent to die in violent suicide attempters. J Clin Psychiatry 2004;65:690-5.

32. Dhossche DM, Meloukheia AM, Chakravorty S. The association of suicide attempts and comorbid depression and substance abuse in psychiatric consultation patients. Gen Hosp Psychiatry 2000;22:281-8.

33. Feinstein A. An examination of suicidal intent in patients with multiple sclerosis. Neurology 2002;59:674-8.

34. Kumar CT, Mohan R, Ranjith G, Chandrasekaran R. Characteristics of high intent suicide attempters admitted to a general hospital. J Affect Disord 2006;91:77-81.

35. Williams JB, Gibbon M, First MB, et al. The Structured Clinical Interview for DSM-III-R (SCID). II. Multisite test-retest reliability. Arch Gen Psychiatry 1992;49:630-6.

36. Keller MB, Lavori PW, McDonald-Scott P, et al. Reliability of lifetime diagnoses and symptoms in patients with a current psychiatric disorder. J Psychiatr Res 1981;16:229-40.

37. Fendrich M, Weissman MM, Warner V, Mufson L. Two-year recall of lifetime diagnoses in offspring at high and low risk for major depression. The stability of offspring reports. Arch Gen Psychiatry 1990;47:1121-7.

38. Foley DL, Neale MC, Kendler KS. Reliability of a lifetime history of major depression: i mplications for heritability and co-morbidity. Psychol Med 1998;28:857-70.

39. Rice JP, Rochberg N, Endicott J, Lavori PW, Miller C. Stability of psychiatric diagnoses. An application to the affective disorders. Arch Gen Psychiatry 1992;49:824-30.

40. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. J Health Soc Behav 1987;28:232-46.

41. Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. Psychol Med 2004;34:1001-11.

42. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. The lifetime history of major depression in women. Reliability of diagnosis and heritability. Arch Gen Psychiatry 1993;50:863-70.

43. Kendler KS, Gardner CO, Prescott CA. Are there sex differences in the reliability of a lifetime history of major depression and its predictors? Psychol Med 2001;31:617-25.

44. Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend 2003;71:7-16.

45. Keller MB, Klein DN, Hirschfeld RM, et al. Results of the DSM-IV mood disorders field trial. Am J Psychiatry 1995;152:843-9.

46. Bromet EJ, Dunn LO, Connell MM, Dew MA, Schulberg HC. Long-term reliability of diagnosing lifetime major depression in a community sample. Arch Gen Psychiatry 1986;43:435-40.

47. Prusoff BA, Merikangas KR, Weissman MM. Lifetime prevalence and age of onset of psychiatric disorders: recall 4 years later. J Psychiatr Res 1988;22:107-17.

48. Andreasen NC, Grove WM, Shapiro RW, Keller MB, Hirschfeld RM, McDonald-Scott P. Reliability of lifetime diagnosis. A multicenter collaborative perspective. Arch Gen Psychiatry 1981;38:400-5.

49. Mazure C, Gershon ES. Blindness and reliability in lifetime psychiatric diagnosis. Arch Gen Psychiatry 1979;36:521-5.

50. Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007;133:859-83.

51. Hochstim J. A critical comparison of three strategies of collecting data from households. Journal of the American Statistical Association 1967;62:976-89.

52. Richman WL, Kiesler S, Weisband S, Drasgow GC. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interview. Journal of Applied Psychology 1999:754-75.

53. Blaney PH. Affect and memory: a review. Psychol Bull 1986;99:229-46.

54. Thompson R, Bogner HR, Coyne JC, Gallo JJ, Eaton WW. Personal characteristics associated with consistency of recall of depressed or anhedonic mood in the 13-year follow-up of the Baltimore Epidemiologic Catchment Area survey. Acta Psychiatr Scand 2004;109:345-54.

55. Zeitler MS, Paine AD, Breitbart V, et al. Attitudes about intimate partner violence screening among an ethnically diverse sample of young women. J Adolesc Health 2006;39:119 e1-8.

56. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990;43:543-9.

57. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46:423-9.

Table 1A: Studies assessing the reliability of suicidal ideation

Study	Study design	Setting	Sampling/ Sample type	N	Avg age at f/u (yrs)	Avg. interval time	Period of obs	Reliability estimate (suicidal ideation)	Reliability correlates
Brener et al. (2002)	Cross-sectional, test-retest	US	Convenience – pop-based sample similar to national distribution of 9th-12th graders	4619	15	15.6 days	12-mon	Ideation: Kappa: 0.74	None tested
Fendrich et al. (1994)	Longitudinal	US (New Haven)	Convenience - Mixed (low and high risk) sample	150	17	2 years	Lifetime	Ideation: a) Conditional K = 0.46 (0.25-0.67) b) Recall/Deny % = 0.50 (0.24-0.73)	None tested
Flisher et al. (2004)	Cross-sectional, test-retest	Cape Town, South Africa	Convenience - community sample of 8th & 11th graders	358	15	10-14 days	12-mon	Ideation Observed agreement$^{\partial}$ =92.1%	None tested
Goldney et al. (1991)	Longitudinal	Adelaide, Australia	Random cluster – Community sample	432	23.6	4 years	Lifetime	Ideation 40% (n=16)* of T1 ideators, denied at T2	None tested

Goldney et al. (2009)	Longitudinal	Adelaide, Australia	Random cluster – Community sample	432	23.6	4 years	Lifetime	<u>Ideation</u> 50% (n=68)* of T1 ideators, denied at T2	<u>Greater reliability</u>: a)↑ T1 ideation And at T2: b) ↓self-esteem c)↑hopeless d) ↑current depressive mood e) ↓functioning f) ↑ negative mood
Koziol-McLain et al. (2000)	Cross-sectional, test-retest	US-Colorado	Random - Population-based sample (Colorado)	229	45.2	7-28 days	12 months	<u>Ideation</u> Kappa= 0.79 (0.59-0.99)	None tested
Nock et al. (2007)	Cross-sectional, test-retest	US-Boston	Convenience - Mixed (clinical and non-clinical) sample	67	17.1	6 months	Lifetime	<u>Ideation</u> Kappa[#] = 0.70	None tested
Shaffer et al. (2004)	Cross-sectional, test-retest	US-NY metro region	Convenience-community sample of 9th-12th graders	85	15	8 days	3-month	<u>Ideation</u> Kappa[#]= 0.48	None tested

Statham et al. (1998)	Longitudinal	Australia	Convenience-pop-based sample of adult twins registered with volunteer Australian Twin Registry	5995	43.4	11-12 years	T1: Current T2: Lifetime	<u>Ideation</u> <u>Women</u>: a) 44% of T1 ideators, denied at T2 b) Kappa= 0.11 <u>Men</u>: a) 32% of T1 ideators, denied at T2 b) Kappa= 0.10	Stratified on gender → homogenous Kappas

$^{\partial}$ = Investigators did not compute Kappa due to significantly different prevalences across the 2 waves
* = denominator is attempters at T1, who ALSO were present for F/U at T2
$^{\#}$ = Did not assess, and appropriately subtract out suicidality "since last interview" (new onset attempts)
note: "denying" does not literally mean, denying…just not re-reporting

Table 1B: Studies assessing the reliability of suicide attempts

Study	Study design	Setting	Sampling/ Sample type	N	Avg age at f/u (yrs)	Avg. interval time	Period of obs	Reliability estimate (suicide attempt)	Reliability correlates
Brener et al. (2002)	Cross-sectional, test-retest	US	Convenience – pop-based sample similar to national distribution of 9th-12th graders	4619	15	15.6 days	12-mon	1) Attempt: Kappa= 0.73 2) Injurious Attempt: Kappa= 0.52	None tested
Brezo et al. (2007)	Longitudinal	Quebec	Random and non-random-mixed (high and low risk) sample	1684	21.4	5 years	Lifetime	1) Attempt: 31% (n=32)* of T1 attempters, denied at T2	None tested
Christl et al. (2006)	Longitudinal	Munich	Random-community sample	2548	23	4 years	Lifetime	Attempt: 33% (n=15)* of T1 attempters, denied at T2	Higher Stability: a) Male b) Older at baseline c) T2 lifetime mental Dx: *Any depressive dx* (OR: 8.4; 1.1-61.5); *Any somatoform dx* (OR:16.1; 1.5-173.2); *comorbidites* (3.5 vs. 2.8 on average)
Fendrich et al. (1994)	Longitudinal	US (New Haven)	Convenience - Mixed (low and high risk) sample	150	17	2 years	Lifetime	Attempt: a) Conditional K = 0.58 (0.19-0.97) b) Recall % = 0.60 (0.17-0.85) = 40% Deny	None tested

Flisher et al. (2004)	Cross-sectional, test-retest	Cape Town, South Africa	Convenience - community sample of 8th & 11th graders	358	15	10-14 days	12-month	Attempt Observed agreement$^{\Phi}$= 94.7%	None tested
Nock et al. (2007)	Cross-sectional, test-retest	US-Boston	Convenience - Mixed (clinical and non-clinical) sample	67	17.1	6 months	Lifetime	Attempt Kappa$^{\#}$ = 0.80	None tested
Shaffer et al. (2004)	Cross-sectional, test-retest	US-NY metro region	Convenience-community sample of 9th-12th graders	85	15	8 days	Lifetime	Attempt Kappa$^{\#}$= 0.58	None tested

* = denominator is attempters at T1, who also were present for F/U at T2

$^{\Phi}$ = Investigators did not computer Kappa because of low (<5%) prevalence at Wave 1.

$^{\#}$ = Did not assess, and appropriately subtract out suicidality "since last interview" (new onset attempts)

note: "denying" does not literally mean, denying…just not re-reporting

Table 1C: Studies assessing the reliability of other suicidal constructs (i.e. plans, gestures, "suicidal content")

Study	Study design	Setting	Sampling /Sample type	N	Avg age at f/u (yrs)	Avg. interval time	Period of obs	Reliability estimate (other suicidal constructs)	Reliability correlates
Brener et al. (2002)	Cross-sectional, test-retest	US	Convenience – pop-based sample similar to national distribution of 9^{th}-12^{th} graders	4619	15	15.6 days	12-mon	Plans: Kappa = 0.67	None tested
Klimes-Dougan (1998)	Longitudinal	NY, NY	Convenience - Mixed (low and high risk) sample	192	15	3-6 years (waves 3 years apart)	Lifetime	"Suicidal content" (i.e. ideation, thoughts of death, wishes for death, attempts): Younger: Kappa=0.42 Older: Kappa=0.60	Greater reliability: a) children of well mothers (vs. children of depressed mothers) b) older age c) females d) current distress e) less likely to report high self-concept
Klimes-Dougan (2007)	Longitudinal	NY, NY	Convenience - Mixed (low and high risk) sample	78	21.8	6.5 years	Lifetime	"Suicidal content": a) 88% agreement b) Kappa = 0.57	Greater reliability: a) current depressed mood b) lower functioning (i.e. higher BDI, lower GAF, etc)

Nock et al. (2007)	Cross-sectional, test-retest	US-Boston	Convenience - Mixed (clinical and non-clinical) sample	67	17.1	6 months	Lifetime	Plans Kappa[#] = 0.71 Gestures Kappa[#]= 0.25	None tested

[#] = Did not assess, and appropriately subtract out suicidality "since last interview" (new onset attempts)

Table 1D: Suicide instrument and question characteristics

Study	Interview Format	Suicidality instrument	Prevalence of suicidality	Questions
Brener et al. (2002)	**T1/T2:** Self-administered pencil/paper	1999 Youth Risk Behavior Survey questionnaire	**T1*:** Ideation=17% (n=785); plan=13% (n=600); attempt=8.4% (n=388); injurious attempt=2.1% (n=97) **T2***: Ideation=16% (n=739); plan=12.9% (n=596); attempt=8.5% (n=393); injurious attempt=2.7% (n=125)	**T1/T2:** 1) Have you ever seriously considered suicide during the past 12 months? 2) Have you planned a suicide attempt during the past 12 months? 3) Have you had one or more suicide attempts during the past 12 months? 4) If you attempted suicide during the past 12 months, did any attempt result in an injury, poisoning, or overdose that had to be treated by a doctor or nurse?
Brezo et al. (2007)	**T1:** Attempt = Structured Diagnostic **T2:** Screening interview	**T1:** Attempt = DISC (Depression module) **T2:** Attempt= 2 Qs from Scale for Suicidal Ideation	**T1:** attempt= 3.5% (n= 60) **T2:** attempt=6.2% (n=104)	**T1:** Have you already attempted suicide? **T2:** Have you already attempted suicide?
Christl et al. (2006)	**T1/T2:** Structured diagnostic; computer-assisted, face-to-face	**T1/T2:** M-CIDI	**T1:** Attempt = 2% (n=69) **T2:** Attempt = 3.5% (n=88) Note: T1 had gateway Qs, T2 had no gateway Qs	**T1/T2:** Have you ever attempted suicide? Note: T1 had gateway Qs, T2 had no gateway Qs
Fendrich et al. (1994)	Structured diagnostic	**T1/T2:** K-SADS-E	**T1:** Attempt= not reported Ideation = not reported	**T1/T2:** $\wedge^{\delta}$ Note: question asked about their "worst-past" episode of depressed mood.

			T2: Attempt=not reported Ideation=not reported	1) Did you ever think about hurting or killing yourself? 2) Did you ever try and kill yourself
Flisher et al. (2004)	Self-report, pencil/paper survey	Unknown source	**T1*:** Ideation =18.4% (n=66) Attempts= 4.7% (n=17) **T2*:** Ideation = 14% (n=50) Attempts= 7.1% (n=25)	**T1/T2:** 1) During the past 12 months, did you ever seriously think about harming yourself in a way that may result in death? 2) During the past 12 months, did you actually ever try to put an end to your life?
Goldney et al. (1991)	Self-report mail survey	**T1:** 4 Qs from the General Health Questionnaire (GHQ) **T2:** unknown source	**T1*:** Ideation = 9% (n=40) **T2*:** Ideation=18.7% (n=81)	**T1:** 1) Have you recently felt that life is not worth living? 2) Have you recently found yourself wishing you were dead and away from it all? 3) Have you recently had thoughts of the possibility that you might do away with yourself 4) Have you recently found the idea of taking your own life kept coming into your mind? **T2:** Have you ever had any thoughts of killing yourself?
Goldney et al. (2009)	Self-report mail survey	**T1:** 4 Qs from the General Health Questionnaire **T2:** unknown source	**T1:** Ideation = 30% (n=302) **T2:** Ideation = not reported	**T1:** 1) Have you recently felt that life is not worth living? 2) Have you recently found yourself wishing you were dead and away from it all? 3) Have you recently had thoughts of the possibility that you might do away with yourself

				4) Have you recently found the idea of taking your own life kept coming into your mind? **T2:** Have you ever had any thoughts of killing yourself?
Klimes-Dougan (1998)	**T1-T3:** Structured diagnostic	**T1:** The Child Assessment Schedule (CAS) **T2/T3:** the Diagnostic Interview for Children/Adolescents (DICA)	**T1:** "Suicidal content": Younger= 12.6% Older = 16% **T2:** "Suicidal content" Younger=16.5% Older=14.9% **T3:** Younger=10.4% Older=22.8 <u>Cumulative lifetime</u>: Younger=29% Older=43.6%	**T1:** 3 questions on "suicidal content" 1) Do you ever think of hurting yourself? If yes, even killing yourself? 2) Do you ever think of how you would do it 3) Did you ever try to hurt or kill yourself? **T2/T3:** 3 questions on "suicidal content" 1) Have you ever thought about killing yourself? 2) Did you ever have a plan about how you are going to kill yourself? 3) Have you ever tried to kill yourself?
Klimes-Dougan et al. (2007)	**T1:** Structured diagnostic **T2:** unknown	**T1:** DICA **T2:** unspecified	**T1*:** "Suicidal content" = 17% (n=13) **T2*:** "Suicidal content" = 15% (n=12)	**T1:** 3 questions on "suicidal content" 1) Have you ever thought about killing yourself? 2) Did you ever have a plan about how you are going to kill yourself? 3) Have you ever tried to kill yourself? **T2:** Question asking them to Recall how they responded to suicidal content questions at T1 interview
Koziol-McLain et al. (2000)	**T1/T2:** Random digit dial telephone survey	**T1/T2:** Behavioral Risk Factor Surveillance System (BRFSS)	**T1*:** Ideation = 5.3% (n=12) **T2*:**	**T1/T2:** In the past year, have you ever seriously thought about trying to hurt yourself in a way that might have resulted in

			Ideation = 3.5% (n=8)	your death?
Nock et al. (2007)	**T1:** Structured Interview, in-person **T2:** Structured interview, telephone	**T1/T2:** Self-Injurious Thoughts and Behaviors Interview (SITBI)	**T1:** Ideation= 70.2% (n=66) Plans = 37.2% (n=35) Gestures= 22.3% (n=21) Attempt = 28.7% (n=27) **T2:** not reported	**T1/T2:** 1) Have you ever had thoughts of killing yourself? 2) Have you ever actually made a plan to kill yourself? 3) Have you ever done something to lead others to believe you wanted to kill yourself when you really had no intention of doing so? 4) Have you ever made an actual attempt to kill yourself in which you had at least some intention to die?
Shaffer et al. (2004)	**T1/T2:** self-administered screening questionnaire	**T1/T2:** Columbia Suicide Screen (CSS)	**T1:** Girls a) Ideation = 18.8% b) Attempt=8.7% Boys a) Ideation=10.7% b) Attempt= 3.5% **T2:** not reported	**T1/T2:** 1) During the past 3 months, have you thought about killing yourself? 2) Have you ever tried to kill yourself
Statham et al. (1998)	**T1:** self-report, mail survey **T2:** semi-structured, telephone survey	**T1:** "State Anxiety-Depression" Subscale of the Delusions-Symptoms-State Inventory **T2:** Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA-OZ): suicidal behavior section	**T1:** Ideation = not reported **T2:** Ideation Women = 22.2% Men = 23.8%	**T1:** "Recently I have been so depressed that I have thought of doing away with myself" **T2:** Have you ever thought about killing yourself?

*= denominator is individuals who were present at both waves 1 and 2
^ in depression module
δ gateway Questions

note: wave number designations do not neceesarily correspond to the wave number reported in paper (i.e., some papers had T0….and others had multiple waves, but the relability estimate was based on only some of them). For consistency purposes, all waves in this review paper are labeled T1 as baseline assesment for the reliability estimate.

Chapter 2

The test-retest reliability of adult self-reported lifetime suicide attempts: predicting discordant reports over time.

Introduction

An estimated 5% of the US adult population attempts suicide at least once in their lifetime.[1] Estimates of incidence, prevalence, and correlates of attempts are necessary for appropriate resource allocation, trend surveillance, and ultimately, for the prevention of future attempts. The reliability of adult lifetime attempt reports in the population is therefore of critical importance, yet is unknown to date.

Large population-based surveys, such as the National Comorbidity Study (NCS), the National Survey on Drug Use and Health (NSDUH), and the National Survey on Alcohol and Related Conditions (NESARC), provide our current estimates of suicidal behaviors. These estimates are based on *self-reports* of suicide attempts, which most often entail only a few, if not just a single question(s), regarding lifetime suicide attempts.[1-6] While these studies are able to capture a broad range of attempts (e.g., both non-medically serious and medically serious) and attempters, self-report measures are prone to reporting biases which may reduce the reliability and validity of these assessments. The respondents must a) comprehend what the investigator considers an "attempt" b) accurately recall making an attempt (which may be prone to recall bias), and c) be willing to reveal their attempt to the interviewer (which may be prone to social desirability bias given the sensitive nature of suicidal behaviors). Each step provides an opportunity for inaccurate ascertainment of suicide attempt estimates at the time of the interview. Without a gold standard (e.g., medical report), there is no way to assess the accuracy (i.e., validity) of attempt reports at any one time point. However, the attempt reports' reliability can shed light on their validity since a measure can only be as valid as it is reliable. An assessment of the test-retest reliability, or concordance, of reports over separate time points is feasible, and necessary.

There is ancillary evidence suggesting that adult reports may suffer from a significant degree of unreliability. A comparison of prevalence estimates of reported suicide attempts between youth and adults in a review by Nock et al.[7] revealed that the lifetime prevalence of suicide attempts among adults (1.9%-4.6%) is actually lower than the 12-month prevalence among adolescents (7.3%-10.6%).[7] One potential explanation for this incongruence is that youth attempts are increasing with more recently born cohorts, while adult attempts are remaining constant or decreasing. However, according to the Youth Risk Behavior Survey (YRBS), youth rates have remained constant from 1991-2001, and have actually been on the decline since 2001.[8] A second plausible explanation for the higher youth rates could be that adolescents who report attempts in their youth complete before adulthood, and are therefore excluded from adult samples, leading to lower lifetime rates in adults. However, according to US Centers for Disease Control and Prevention data, youth suicide is rare with approximately 8 deaths per 100,000 persons aged 10-24, and therefore could not account for the attempt prevalence discrepancy.[9] A third possible explanation could be that youth are overreporting attempts, while adults are reporting accurately. However, social desirability effects tend to cause individuals to underreport rather than overreport sensitive experiences,[10] particularly related to mental health,[11,12] therefore overreporting of youth attempts is unlikely to fully explain the discrepancy. The most plausible explanation, and one suggested by experts in the field,[7] is that adults are underreporting their past lifetime attempts.

Studies of *youth* attempts also lend credence to the plausibility of unreliable adult lifetime attempt reports. Five studies have found evidence of unreliable reporting of lifetime suicide attempts in adolescents and young adults,[13-17] with kappa coefficients ranging from 0.58-0.80. In a longitudinal study, Christl et al.[14] found that among a community sample of adolescents in

Munich, 33% (n=15) of those who reported a lifetime attempt at baseline recanted 3 years later. Also utilizing a community sample, Shaffer et al.[17] administered the Columbia Suicide Screen to 9th-12th graders, and reported a kappa of 0.58 for the 8-day test-retest reliability. Three other studies examined the reliability of lifetime attempts within mixed risk (low and high) samples. Fendrich et al.[15] assessed attempts two years apart and reported a 40% recanting frequency and a conditional kappa coefficient of 0.58 (0.19-0.97). Brezo et al.[13] found that among a sample of Quebec youth, 31% of those who reported attempts at baseline, recanted them 5 years later. Lastly, Nock et al.[16] examined the 6-month reliability of lifetime attempts, resulting in a kappa coefficient of 0.80. These five studies demonstrate varying degrees of attempt report reliability among youth, and while unexamined to date, lend supporting evidence to possible unreliability of attempt reports among adults.

Very little is known about individuals who inconsistently report attempts over time. Are they true attempters who underreport their attempt at one time point or true non-attempters who falsely report an attempt at one time point? Further, if they are indeed true attempters, why do they underreport their attempts? Only one of the five aforementioned studies examined correlates of discordant reporters. Christl et al. (2006) found that younger age, female gender, and a lack of past depression predicted recanting an attempt three years later. There is some evidence that in the absence of depression, attempters make less medically and psychologically (i.e. intent to die) serious attempts,[18-23] and it's plausible that respondents may be less likely to remember less serious attempts.[11,14,24] It is also possible that less serious attempters may feel uncertain whether the incident was a true suicide attempt or not. Respondents who feel this uncertainty may reinterpret the incident as not a true attempt at another time point, and accordingly, not report it.[14] Regardless of mechanism— recall or reinterpretation, it is possible that respondents without

a history of depression would be more likely to discordantly report compared with those with a history of depression.[14] Christl et al.'s findings are notable because they not only demonstrate the occurrence of discordant attempt reporting, but also are the first to illuminate potential predictors. Nonetheless, Christl et al. examined a limited number of correlates, with findings based on a small non US-based youth sample, therefore there is need to further this line of inquiry within a US adult sample, with a robust size and diversity of predictors.

It is plausible that adult lifetime suicide attempt reports suffer from a significant degree of unreliability, hindering our understanding of the incidence, prevalence, and correlates of attempts. However, there has yet to be an inquiry focused on the quantification and correlates of unreliable reporting. Examining predictors of discordant reporting may also shed light on the validity of discordant reports. Based on the assumption that respondents who consistently report an attempt over time (Concordant yes respondents) are true attempters, examining similarities between discordant and Concordant yes, and between discordant and Concordant no respondents (those who consistently do not report attempts) may allow for inferences regarding whether the discordant responders are themselves likely true attempters or not. If discordant responders are indeed true attempters, then the prevalence of adult lifetime suicide attempts reported in the literature is likely to be an underestimate of the true prevalence. Furthermore, examining predictors of discordant reporters may also provide clues as to reasons why respondents under or overreport attempts (e.g. issues related to recall or reinterpretation).

Against this background, this paper reports results from a longitudinal study in which a large, nationally representative sample of US adults reported on lifetime suicide attempts, at two waves, 3 years apart. This test-retest research design allows us to examine our first aim, to conduct the first reliability assessment of adult-reported attempts over time. Our second aim is

to examine the extent to which discordant reporting over time is related to various sociodemographic and psychiatric characteristics. This analysis will demonstrate if discordant reporters appear overall more similar to individuals who consistently report an attempt over time (Concordant yes respondents), or to those who consistently report no attempt over time (Concordant no respondents). We hypothesize that 1) discordant responders will be *overall* more similar to Concordant yes individuals, given that in general, individuals tend to underreport, rather than falsely report sensitive experiences, however specifically, 2) discordant reporters will be less likely to have a history of depressive disorders than Concordant yes respondents, since depression is inversely associated with attempt severity, and less serious attempts may be more likely to be forgotten or reinterpreted.

MATERIALS AND METHODS

Sample

The data used in this investigation was from the National Survey on Alcohol and Related Conditions (NESARC), developed by the National Institute on Alcohol Abuse and Alcoholism (NIAAA). An overview of the study design has been described elsewhere.[25] In brief, it is a longitudinal survey with its first wave of interviews fielded in 2001-2002 and second wave in 2004-2005. The NESARC is a representative sample of the non-institutionalized US population 18 years of age and older. Respondents were informed in writing about the nature of the survey, the statistical uses of the survey data, the voluntary aspect of participation and the federal laws that protect the confidentiality of the identifiable survey information. Those respondents consenting to participate after receiving this information were interviewed in person. A sample of 43,093 participants was obtained for Wave 1 with a response rate of 81%. A total of 34,653

respondents were re-interviewed at Wave 2, providing a follow-up rate of 87% and a cumulative response rate of 70.2%.

The NESARC used a multistage sampling design that oversampled specific ethnic groups, including African Americans and Hispanics, and also oversampled young adults. Data were weighted to account for oversampling of specific groups, clustered sampling, and non-response. The data were weighted to be representative of the US civilian population on the basis of the 2000 Decennial Census of Population and Housing. The specific aspects of the sampling design of the NESARC are described elsewhere in detail.[26]

Measures

At each wave, the NESARC collected detailed information on basic sociodemographic characteristics, the presence and timing of mental health symptoms and diagnoses, and information on lifetime suicide attempts. Participants were interviewed with computer-assisted face-to-face interviews using the Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV) diagnostic survey.[25] This survey took approximately 1 hour, and was administered by either lay interviewers or clinicians.

Lifetime suicide attempts

Lifetime suicide attempts, our outcome of interest, were assessed at both Waves 1 and 2. At Wave 1, the suicide attempt question was only asked of individuals who answered "yes" to at least one of the two "gateway" or "stem" questions for a diagnosis of Major Depressive Disorder (MDD). Respondents were asked the following MDD stem questions: 1) "In your entire life, have you ever had a time when you felt sad, blue, or down most of the time for at least 2

weeks?", and 2) "In your entire life, have you ever had a time, lasting at least 2 weeks, when you didn't care about the things that you usually cared about, or when you didn't enjoy the things you usually enjoyed?" (see Appendix 2A). Any respondent who answered "yes" to at least one of the two questions was then asked the following suicide attempt question: "During that time when your mood was at its lowest or you enjoyed or cared the least about things, did you attempt suicide?" At Wave 2, all respondents, regardless of their response to the MDD gateway questions at Wave 1, were asked the following suicide attempt question: "In your ENTIRE life did you EVER attempt suicide?" The question regarding a lifetime suicide attempt was therefore present in the survey at both Waves 1 and 2. However, because individuals were not asked this question at Wave 1 if they did not respond "yes" to at least one of the two MDD stem questions, this study necessarily focuses only on the subgroup of individuals who were asked the suicide question at Wave 1 (Figure 2A).

Individuals who reported a lifetime attempt at Wave 1 and reported no lifetime attempt at Wave 2 were labeled ***"Recanters"*** (Figure 2A). Individuals that reported no lifetime attempt at Wave 1, but newly reported a lifetime attempt at Wave 2, and which had not taken place between Waves 1 and 2, were labeled as ***"New endorsers"***. Together, the "Recanters" and the "New endorsers" comprised the ***"Discordant responders".*** Individuals who consistently reported a lifetime suicide attempt at both Waves 1 and 2 were labeled as ***"Concordant yes"***, and those who consistently reported no lifetime suicide attempt at both Waves 1 and 2 were labeled as ***"Concordant no"*** (Figure 2A).

Suicidal ideation and suicide attempt characteristics

Other aspects of suicidality, namely ideation and characteristics of attempts, were broadly assessed in this survey. Lifetime suicidal ideation was only assessed at Wave 1. Individuals who answered "yes" to at least one of the two gateway questions for MDD, were asked the following: 1) Did you ever think about committing suicide?; 2) Did you ever feel like you wanted to die?; and 3) Did you ever think a lot about your own death? At Wave 2, respondents who endorsed at least one of the two MDD gateway questions regarding their experience "since the last interview", were asked about suicidal ideation that occurred since the last interview. Therefore, if respondents answered 'yes' to at least one of the three ideation questions, at either wave, they were categorized as having had *lifetime suicidal ideation* at Wave 2 (see Appendix 2B). Attempt characteristics were only assessed at Wave 2. Information regarding these variables is only available for New endorsers and Concordant yes individuals since these questions were only asked if the individual reported an attempt at Wave 2. Attempt characteristics were assessed through the following questions: 1) How old were you the first time that happened? (*age at first attempt*) and 2) How old were you the most recent time that happened? (*age at last attempt*). *Recency of attempt* was calculated based on current age and age of last attempt (see Table 2A).

Psychiatric disorders

The NESARC uses the Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS-IV). The AUDADIS-IV is a fully structured assessment administered by trained lay interviewers, and has demonstrated good test-retest reliability on mental health outcomes.[25] In the AUDADIS-IV, lifetime diagnoses are defined as disorders occurring in the past 12 months and/or prior to the past 12 months, as reported at Wave 2. All mental health diagnoses were

based on the DSM-IV, and those included in our analysis as potential predictors of attempt report stability were the following: Lifetime diagnoses of Major Depressive Disorder (MDD) and Dysthymia; lifetime diagnoses of Manic Episode, Hypomanic Episode, Generalized Anxiety Disorder (GAD), Panic Disorder, Social Phobia, Agoraphobia, Specific Phobia, PTSD, Borderline Personality Disorder, Schizotypal Personality Disorder, Narcissistic Personality Disorder; and lifetime diagnoses of Alcohol Abuse and Dependence Disorders. The psychiatric outcomes used in these analyses are the following: 1) *lifetime depressive disorder* (lifetime MDD or Dysthymia), 2) *lifetime anxiety disorder* (lifetime GAD, Panic Disorder, Social Phobia, Agoraphobia, Specific Phobia, or PTSD), 3) *lifetime Substance Use Disorder* (lifetime alcohol, heroin, inhalant, cocaine, hallucinogen, cannabis, amphetamine, opioid, tranquilizer, or sedative abuse and/or dependence) and 4) *lifetime Personality Disorder* (lifetime Borderline, Antisocial, Schizotypal, or Narcisstic Personality Disorders). See Appendices 2C, 2D, 2E, 2F & 2G for descriptive and results relating to individual psychiatric disorders.

Sociodemographics

Sociodemographic variables known to be associated with adult lifetime suicide attempts were considered predictors of interest for discordant reporting, and included: Wave 2 *age, race/ethnicity, sex, marital status, income, and education level*. We chose these variables because if found to also predict reporting of attempts, there would be implications for the accuracy of predictor- attempt associations reported in the literature.

Data Analysis

The first aim of this study was to assess the reliability of adult lifetime attempt reports. We therefore calculated the frequency of Recanters, New endorsers, Concordant yes and Concordant no individuals, as well as the kappa coefficient for the reliability of attempt reports across waves. We also examined reliability across strata of our sociodemographic and psychiatric predictor variables. We used SAS 9.2 to conduct these analyses.

Our second aim was to examine the extent to which discordant reporting over time was related to various sociodemographic and psychiatric characteristics. Predictors of discordant responses were determined in three main analyses. First, bivariate predictors of discordant responders (i.e., Recanters & New endorsers) were compared with 1) Concordant yes individuals and 2) Concordant no individuals, in order to determine broad differences between concordant and discordant responders. This was assessed using a 3-category outcome variable consisting of discordant responders, Concordant yes and Concordant no individuals. Recanters were then compared with New endorsers to detect any significant differences that may warrant separation in further analyses. This was assessed using a 2-category outcome variable consisting of Recanters and New endorsers. Recanters and New endorsers were then each compared with Concordant yes individuals, to assess similarities between discordant responders and likely true attempters, and analyzed using 4-category outcome consisting of Recanters, New endorsers, Concordant yes and Concordant no individuals. Predictors considered included sociodemographic variables, suicidal ideation, psychiatric outcomes, and suicide attempt characteristics[*]. Model variables were parsimoniously chosen based on significant bivariate associations and theoretical salience. If a variable was significant in at least one of the bivariate

[*] For New endorsers and Concordant yes individuals only

comparisons, it was included in all models. All logistic regression models included age, sex, education, income, marital status, lifetime suicidal ideation, depressive disorders, anxiety disorders, substance use disorders, and personality disorders. Models were analyzed using SUDAAN version 10 to account for the complex sampling structure, and significance level was set at $\alpha = 0.05$.

RESULTS

Reliability Assessment

The kappa coefficient for the reliability of the attempt reports across the two waves was 0.51. As shown in Table 2B, of the individuals who reported a lifetime attempt at Wave 1, 377 individuals (42%) "recanted" their attempt at Wave 2. Of those individuals who reported no lifetime attempt at Wave 1, 187 individuals (1.8%) "newly endorsed" a lifetime attempt at Wave 2. Lastly, there were 516 "Concordant yes" and 10,010 "Concordant no" individuals. As shown in Table 2C, we examined the reliability across strata of our main predictors of interest (see Appendix 2F for Kappa coefficients of individual psychiatric disorders). Six of the ten variables had a low degree of reliability variation across strata (age, gender, education, marital status, and any lifetime anxiety and substance use disorder), with all stratum-specific coefficients within an average of 0.05 of one another. Three variables had moderate degree of reliability variation across strata (income, any lifetime depressive disorders, and any lifetime personality disorders), with coefficients across strata being within an average of 0.15 of one another. Specifically, respondents who earned 35-69K, and did not have a history of depressive or personality disorders, had lower reliability of reporting. Lastly, there was a high degree of reliability

variation for suicidal ideation; individuals with a history of lifetime suicidal ideation had a kappa coefficient of 0.52, while those without a history of ideation had an extremely low coefficient of 0.13. In summary, the overall reliability of lifetime attempts was fair, and the reliability across strata of the main predictors of interest varied from slight to moderate (according to criteria by Landis and Koch[27]).

Predictors of discordant reporting

Our second aim was to examine if discordant reporting was associated with sociodemographic and psychiatric characteristics. Our purpose was two-fold. First, to determine if discordant responders were on average more similar to Concordant yes or Concordant no individuals. Second, to determine if discordant individuals were significantly different from Concordant yes individuals regarding history of depressive disorders. Predictors of discordant responses were determined in three main analyses (Tables 2D-F) (see Appendix 2G for results for individual psychiatric disorders).

Discordant vs. Concordant yes and Discordant vs. Concordant no (Table 2D)

We first examined discordant responders (Recanters + New endorsers) by comparing them separately with 1) Concordant no and 2) Concordant yes individuals. In the final model, we found that discordant responders appeared more similar to Concordant yes than Concordant no individuals, both in regard to sociodemographic factors as well as psychiatric outcomes. *Sociodemographics:* Discordant responders were less likely than Concordant no individuals to be 65+ (AOR, 0.33; 95% CI, 0.20-0.54), to have earned more than a high school degree (AOR,

0.72; 95% CI, 0.52-0.99), and have an income over 35K (AORs, 0.52-0.68). There were no significant sociodemographic differences between discordant and Concordant yes individuals.

Mental Health: Discordant responders were more likely than Concordant no individuals to have a history of suicidal ideation (AOR, 10.53; 95% CI, 6.71-16.53), and any depressive, anxiety, substance use or personality disorders (AORs: 1.29-1.74). In comparison, in the final model, discordant responders only differed with Concordant yes individuals on two outcomes: discordants were less likely than Concordant yes individuals to have had lifetime suicidal ideation (AOR, 0.13; 95% CI, 0.05-0.37) and a personality disorder (AOR, 0.39; 95% CI, 0.28-0.55). In summary, discordant responders appeared more similar to Concordant yes than Concordant no individuals, and specifically, discordant responders did not differ from Concordant yes individuals with regard to history of depressive disorders.

New endorsers vs. Recanters (Table 2E)

We tested for significant differences between New endorsers and Recanters to determine if they should be compared separately to Concordant yes individuals. Overall, there were some significant differences between the two groups, mostly with regards to psychiatric characteristics. *Sociodemographics*: In the final model, New endorsers were similar to Recanters, with the exception of one stratum of age; New endorsers were 2.10 (95% CI, 1.04-4.28) times as likely to be 30-64 compared with Recanters. *Mental Health:* In the final model, New endorsers were less likely to have a history of suicidal ideation (AOR, 0.21; 95% CI, 0.08-0.55), but more likely to have any anxiety (AOR, 2.48; 95% CI, 1.42-4.35) or personality disorder (AOR, 2.71; 95% CI, 1.48-4.96) compared with Recanters.

Recanters vs. Concordant yes (Table 2F)

Since New endorsers and Recanters differed on some predictors, we compared them to Concordant yes individuals separately. We first examined differences between Recanters and Concordant yes individuals. *Sociodemographics*: In the final model, compared to Concordant yes responders, Recanters were younger (30-64 vs. <30; AOR, 0.52; 95% CI, 0.31-0.86), less likely to have more than a high school degree (AOR, 0.56; 95% CI, 0.34-0.91) and more likely to be male (AOR, 1.53; 95% CI, 1.04-2.25). *Mental health*: In the final model, compared to Concordant yes individuals, Recanters were less likely to have a history of suicidal ideation (AOR, 0.25; 95% CI, 0.07-0.82) or a personality disorder (AOR, 0.28; 95% CI, 0.19-0.42). In summary, while there were some sociodemographic and psychiatric differences between Recanters and Concordant yes individuals, there were no differences regarding history of depressive disorders.

New endorsers vs. Concordant yes (Table 2F):

We then examined differences between New endorsers and Concordant yes individuals. In terms of both sociodemographic and mental health characteristics, New endorsers appeared similar to Concordant yes individuals. *Sociodemographics*: No significant differences. *Mental Health:* There were two significant differences between New endorsers and concordant individuals; New endorsers were much less likely to have a history of suicidal ideation (AOR, 0.05; 95% CI, 0.02-0.15), but more likely to have any lifetime anxiety disorder (AOR, 1.90; 95% CI, 1.09-3.30). *Suicide attempt characteristics:* New endorsers differed from Concordant yes individuals to some extent regarding all the suicide attempt characteristics examined. New endorsers were younger (18-44 vs. < 18 years) at their first attempt (AOR, 0.60; 95% CI 0.40-0.90) and at their

last attempt (AOR, 0.52; 95% CI, 0.33-0.82), more likely to have attempted more than 20 years ago (AOR, 2.45; 95% CI 1.38-4.32), and less likely to have attempted more than once (AOR, 0.54; 95% CI, 0.33-0.89) compared with Concordant yes individuals. Therefore, there were no significant sociodemographic differences, some psychiatric differences, and multiple attempt characteristic differences between New endorsers and Concordant yes individuals; however there were no differences regarding history of depressive disorders.

To summarize our results, as a whole, discordant individuals appeared more similar to Concordant yes than Concordant no individuals. When New endorsers and Recanters were each compared with Concordant yes individuals, sociodemographic differences emerged; namely, Recanters were younger, more likely to be male, and less educated than Concordant yes individuals, while New endorsers were sociodemographically identical to Concordant yes individuals. In regards to mental health characteristics, while New endorsers and Recanters were no less likely to have a history of depressive disorders compared with Concordant yes individuals, they were much less likely to have a history of suicidal ideation.

DISCUSSION

This is the first study to examine the reliability of suicide attempt measures in adults. We found the reliability over a 3-yr period to be moderate (0.51). This is to say, individuals' responses at each wave were consistent with each other only 51% of the time greater than would be expected by chance alone. There is no study in the literature to serve as a direct comparison for this estimate; however the reliability of attempts in *youth* has been examined by five prior studies.[13-17] The reliability obtained in this study, among adults, is lower than any of those obtained in the youth samples. This finding is difficult to interpret given the varying time intervals between

assessments, suicide constructs, and ages across samples; however if recall failure was indeed a reason for unreliable reporting, adults would likely be less reliable in their reports since on average more time has passed since their attempt, compared with youth.

The second aim of this study was to determine predictors of discordant responses. We hypothesized that on average, discordant responders would be more similar to Concordant yes than Concordant no individuals. Discordants did indeed appear more similar to Concordant yes than Concordant no individuals. Under the assumption that Concordant yes individuals are likely true attempters, and that Concordant no individuals are true non-attempters, our findings would imply that discordant reporters are also likely true attempters, and therefore underreporting, rather than falsely reporting their attempts.

We further hypothesized that discordant responders would be less likely than Concordant yes individuals to have a history of depressive disorders. Contrary to our hypothesis however, while depression was only marginally non-significant in the bivariate model, it fell away entirely in the final model after adjusting for other predictors. The null finding remained even when Recanters and New endorsers were analyzed separately. This finding was surprising since the reliability of attempts among individuals without a history of depressive disorders was shown (Table 2C) to be lower than the reliability among those with such history (κ = 0.39 and 0.52, respectively). However, as shown in Appendix 2C, approximately 90% of respondents who reported an attempt at any wave (Concordant yes= 91.5%; Recanter=86.9%; New Endorser=87.5%), had a depressive disorder by Wave 2. Therefore, we may not have had enough power to detect small differences such as these. Notably, when we examined differences in lifetime MDD and Dysthymia separately, across discordant and concordant groups (see Appendix 2C and 2G), we found some significant differences. In a bivariate assessment,

discordants were less likely than Concordant yes individuals to have had Dysthmia (OR, 0.54; (95% CI, 0.39-0.75); Recanters were less likely than Concordant yes individuals to have had MDD (OR, 0.60; 95% CI, 0.37-0.97) and Dysthymia (OR, 0.53; 95% CI, 0.37, 0.76); and New endorsers were less likely than Concordant yes individuals to have had Dysthymia (OR, 0.56; 95% CI, 0.36-0.89). Therefore, while there were no significant differences in depressive disorders as a combined category, individually, findings regarding lifetime MDD and Dysthymia provide partial support for our hypothesis that discordant responders are less likely to have a history of depressive disorders.

Surprisingly, lack of lifetime suicidal ideation emerged as the predominant predictor of discordant reports in all final models. This finding corroborated the low reliability found amongst those who had no history of ideation (Table 2C, $\kappa = 0.13$). While these findings may appear to contradict those regarding depression, we posit a plausible explanation. Attempts made in the context of little pre-meditation or planning, are referred to as impulsive attempts. Therefore, our results signify that while discordant reporters may indeed be true attempters, they might be characterized by more impulsive attempts than Concordant yes responders. This explanation is consistent with many findings suggesting impulsive attempts are less psychologically and medically serious (i.e. less intent to die and lower lethality).[21,28-31] Following the same logic described in our introduction, attempts that are less serious may be more vulnerable to both reinterpretation and recall failure. Hence, suicidal ideation may serve as a marker for attempt severity, rather than depression, as originally hypothesized. Therefore, while discordant attempters are likely true attempters, they may differ on the characteristics of the attempt, namely impulsivity and severity. More detailed data on the attempts themselves would be needed to explore this hypothesis in greater depth.

We examined possible differences between Recanters and New endorsers to determine if a more refined analysis was indicated. These findings were intriguing because while these two groups of discordant reporters were overall sociodemographically similar, they differed on three mental health predictors in the final model. New endorsers were much less likely than Recanters to have lifetime ideation, but much more likely to have any lifetime anxiety or personality disorder. While reasons behind these differences are unclear, it is possible that the varied positioning of the attempt question within the survey across the two waves may have led to differences between the two groups.

When we compared Recanters and New endorsers each with Concordant yes responders, we once again found no relationship between discordant reporting and history of depressive disorders. However, some additional differences emerged in these more refined analyses, some of which may provide clues to possible underlying mechanisms of the underreporting of attempts. Comparing Recanters and Concordant yes individuals, we found Recanters to be younger and less educated. This finding was not predicted, but may indicate a lack of construct comprehension such that those that were discordant interpreted the incident as an attempt at one wave, and not an attempt at another. "Suicide attempt" is an undefined construct in the survey, and there is often confusion among lay audiences regarding whether other forms of self-harm (e.g. cutting) or suicidal gestures ("cries for help" which do not carry the intention of death), are also considered attempts. It's plausible that this confusion would be more common among younger and less educated individuals. In addition, Recanters were found to be more likely than the Concordant yes individuals to be male. Males may be more reluctant to report a suicide attempt due to social desirability effects rooted in gender norms and stereotypes.

Comparing New endorsers and Concordant yes individuals, we were able to examine various attempt characteristics which add yet another dimension to the picture of discordant reporters. New endorsers were more likely to be younger at the time of their first and last attempt, have attempted more than 20 years ago, and have had only one attempt. There are a few plausible interpretations of these findings. First, it is possible that New endorsers failed to recall the attempt at Wave 1 due to its lack of recency and its single occurrence. It is also plausible that the respondent remembered the event at Wave 1, however interpreted it as not an attempt. Due to the lack of recency of the event, the details and context surrounding it may be ambiguous, and hence the new endorser may have interpreted the event as an attempt at Wave 2, which they did not label as such at another Wave 1. In summary, since New endorsers overall appear similar to Concordant yes individuals regarding sociodemographic and mental health indicators, yet different regarding suicidal ideation and attempt characteristics, we believe that New endorsers are likely true attempters, but may differ in regards to certain characteristics or perhaps circumstances (e.g. impulsive) of the attempts themselves.

Comparing our findings to Christl et al. [14] who examined predictors of discordant attempt reports among a sample of adolescents and young adults in Munich, we note similarities as well as differences. Like Christl et al., we found that younger age predicted recanting, however in contrast, we found recanting to be more common among males than females. When mental health associations were examined, Christl et al. found that a history of depression predicted response concordance over time, while we observed no such difference. Our study differed from that of Christl et al. in ways that may have contributed to the disparate findings. First, the NESARC is a rich and detailed dataset, allowing for a greater breadth of potential predictors. For instance, one predictor of note that was not reported in Christl's study, but was very prominent in

our findings, was suicidal ideation. Secondly, Christl et al. had only 69 attempts at baseline, while our sample contained 893 baseline attempts. Our larger sample size provided the power to examine predictors in adjusted models, rather than a bivariate analysis alone. If Christl et al. had assessed ideation and depression together in an adjusted analysis, perhaps we would have had similar findings. Finally, our sample comprised US adults over the age of 18, as opposed to adolescents and young adults from the Munich area; there may be different underlying mechanisms, and therefore predictors relating to both age and region, driving discordant reporting in these two populations.

This investigation had a number of limitations. This investigation could only assess the reliability of attempt reports among individuals who passed the MDD gateway questions at Wave 1, thereby excluding 29,340 individuals from our analysis. Unfortunately, due to the survey structure at Wave 2, we are unable to assess how many attempts were missed due to the gateway questions at Wave 1. The NESARC's predecessor survey, the National Longitudinal Alcohol Epidemiologic Survey (NLAES), examined the gateway question issue and found that the number of individuals who did not screen into the depression section and reported a suicide attempt was very low (< 0.1% of the sample).[32] Unlike at Wave 1, the attempt question at Wave 2 was not preceded by gateway questions, which could also influence our findings. The fact that respondents were not "primed" by the depression questions at Wave 2 may explain the high percentage of Recanters, and the comparatively low percentage of New endorsers. This lack of priming may have diminished recall or influenced reinterpretation among the respondents. Additionally, this could explain some of the observed differences between Recanters and New endorsers. As with all prospective studies, this study had some loss to follow up (Figure 2A). Seventeen percent of our sample was lost to follow-up before Wave 2. Individuals lost to follow

up were similar to the final sample on most major baseline predictors, but were slightly less likely to have a history of depressive or anxiety disorders. If those lost to follow up were also more likely to be discordant reporters, it could perhaps partially explain our null finding. Lastly, many characteristics of suicide attempts, particularly the medical and psychological seriousness of the attempts, were not included in the NESARC, but could certainly influence reporting reliability.

The investigation has a number of public health implications. First and foremost, the reliability of attempt reports is critical for its impact on the observed prevalence of attempts in the population. For example, as observed, 2.1% of respondents reported an attempt at Wave 1. If those that newly endorsed at Wave 2 were true attempters at Wave 1, then the true prevalence would be 2.5%. Likewise, as observed, 2.1% of respondents reported an attempt at Wave 2. If those that recanted at Wave 2 were true attempters, then the true prevalence at Wave 2 would be 3.1%. Therefore, the only moderate reliability is likely contributing to the underestimation of the prevalence of attempts in the population. Accurate estimates are necessary for the surveillance of time-related trends[32,33] and appropriate resource allocation. Second, given that discordant and concordant responders differ on varying sociodemographic and mental health indicators, the effect estimates between predictors and attempts in the population may be also biased to some extent. For example, in our study Recanters were more likely to be male compared with Concordant yes responders. Therefore, if males are less likely to report their suicide attempts, then the commonly reported increased odds of attempts among women[1,34,35] is potentially an overestimate. Third, determining predictors of discordant reporting may allow for the refinement of attempt measures used in epidemiologic investigations. For example, if in fact impulsive attempts tend to be underreported in surveys, then attempt questions could be modified to

address this (e.g., “Have you ever attempted suicide in your life? By suicide attempt, I am referring both to impulsive and premeditated events.”). Lastly, understanding the unreliability of suicide attempt reports may have broader implications for many other areas of psychiatric epidemiology. Unreliability exists to varying degrees in the measurement of all psychiatric diagnoses and conditions, and examination of it within the suicide literature may aid in explaining this phenomenon within other psychiatric disorders.

In conclusion, our findings indicate that adult lifetime suicide attempt reports suffer from a moderate degree of unreliability. Further, our results suggest that discordant reporters are likely to be true attempters, and therefore underreporting, rather than falsely reporting their attempts. While we did not find discordant individuals to be less likely to have a lifetime depressive disorder than Concordant yes responders, we did find them to have less suicidal ideation. We propose this may indicate that discordant reporters had more impulsive, and therefore perhaps less serious attempts. Future studies should aim to further unravel the mechanisms underlying the underreporting of attempts, considering possibilities of recall failure, event reinterpretation, and perhaps even social desirability, or lack of suicide construct comprehension. Furthermore, studies should consider how the respondent’s mood at the time of interview may affect the reporting of past attempts, specifically by affecting recall or reinterpretation. This may aid in elucidating why a discordant responder may recall or reinterpret an attempt at one time point, and not another.

References

1. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry.* Jul 1999;56(7):617-626.

2. Bolton JM, Belik SL, Enns MW, Cox BJ, Sareen J. Exploring the correlates of suicide attempts among individuals with major depressive disorder: findings from the national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry.* Jul 2008;69(7):1139-1149.

3. Dube SR, Anda RF, Felitti VJ, Chapman DP, Williamson DF, Giles WH. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: findings from the Adverse Childhood Experiences Study. *Jama.* Dec 26 2001;286(24):3089-3096.

4. Garroutte EM, Goldberg J, Beals J, Herrell R, Manson SM. Spirituality and attempted suicide among American Indians. *Soc Sci Med.* Apr 2003;56(7):1571-1579.

5. Ialongo N, McCreary BK, Pearson JL, et al. Suicidal behavior among urban, African American young adults. *Suicide Life Threat Behav.* Fall 2002;32(3):256-271.

6. Nock MK, Kessler RC. Prevalence of and risk factors for suicide attempts versus suicide gestures: analysis of the National Comorbidity Survey. *J Abnorm Psychol.* Aug 2006;115(3):616-623.

7. Nock MK, Borges G, Bromet EJ, Cha CB, Kessler RC, Lee S. Suicide and suicidal behavior. *Epidemiol Rev.* 2008;30:133-154.

8. CDC. Trends in the Prevalence of Suicide-Related Behaviors. National YRBS: 1991-20092010.

9. National Center for Injury Prevention and Control C. *WISQARS fatal injuries: mortality reports.* Atlanta, GA: Centers for Diease Control and Prevention;2008.

10. Schaeffer NC. *Asking Questions About Threatening Topics: A Selective Overview. In The Science of Self-Report.* London: Lawrence Erlbaum Associates; 1999.

11. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. *J Health Soc Behav.* Sep 1987;28(3):232-246.

12. Robins LN. Epidemiology: reflections on testing the validity of psychiatric interviews. *Arch Gen Psychiatry.* Sep 1985;42(9):918-924.

13. Brezo J, Paris J, Barker ED, et al. Natural history of suicidal behaviors in a population-based sample of young adults. *Psychol Med.* Nov 2007;37(11):1563-1574.

14. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T. The accuracy of prevalence estimations for suicide attempts. how reliably do adolescents and young adults report their suicide attempts? *Arch Suicide Res.* 2006;10(3):253-263.

15. Fendrich M, Warner V. Symptom and substance use reporting consistency over two years for offspring at high and low risk for depression. *J Abnorm Child Psychol.* Aug 1994;22(4):425-439.

16. Nock MK, Holmberg EB, Photos VI, Michel BD. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. *Psychol Assess.* Sep 2007;19(3):309-317.

17. Shaffer D, Scott M, Wilcox H, et al. The Columbia Suicide Screen: validity and reliability of a screen for youth suicide and depression. *J Am Acad Child Adolesc Psychiatry.* Jan 2004;43(1):71-79.

18. Astruc B, Torres S, Jollant F, et al. A history of major depressive disorder influences intent to die in violent suicide attempters. *J Clin Psychiatry.* May 2004;65(5):690-695.

19. Dhossche DM, Meloukheia AM, Chakravorty S. The association of suicide attempts and comorbid depression and substance abuse in psychiatric consultation patients. *Gen Hosp Psychiatry.* Jul-Aug 2000;22(4):281-288.

20. Kumar CT, Mohan R, Ranjith G, Chandrasekaran R. Characteristics of high intent suicide attempters admitted to a general hospital. *J Affect Disord.* Mar 2006;91(1):77-81.

21. Suominen K, Isometsa E, Henriksson M, Ostamo A, Lonnqvist J. Hopelessness, impulsiveness and intent among suicide attempters with major depression, alcohol dependence, or both. *Acta Psychiatr Scand.* Aug 1997;96(2):142-149.

22. Brown GK, Henriques GR, Sosdjan D, Beck AT. Suicide intent and accurate expectations of lethality: predictors of medical lethality of suicide attempts. *J Consult Clin Psychol.* Dec 2004;72(6):1170-1174.

23. Hasley JP, Ghosh B, Huggins J, Bell MR, Adler LE, Shroyer AL. A review of "suicidal intent" within the existing suicide literature. *Suicide Life Threat Behav.* Oct 2008;38(5):576-591.

24. Simon GE, VonKorff M. Recall of psychiatric history in cross-sectional surveys: implications for epidemiologic research. *Epidemiol Rev.* 1995;17(1):221-227.

25. Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend.* Jul 20 2003;71(1):7-16.

26. Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry.* Apr 2004;61(4):361-368.

27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* Mar 1977;33(1):159-174.

28. Baca-Garcia E, Diaz-Sastre C, Basurte E, et al. A prospective study of the paradoxical relationship between impulsivity and lethality of suicide attempts. *J Clin Psychiatry.* Jul 2001;62(7):560-564.

29. Baca-Garcia E, Diaz-Sastre C, Garcia Resa E, et al. Suicide attempts and impulsivity. *Eur Arch Psychiatry Clin Neurosci.* Apr 2005;255(2):152-156.

30. Groholt B, Ekeberg O, Haldorsen T. Adolescents hospitalised with deliberate self-harm: the significance of an intention to die. *Eur Child Adolesc Psychiatry.* Dec 2000;9(4):244-254.

31. Brent DA. Correlates of the medical lethality of suicide attempts in children and adolescents. *J Am Acad Child Adolesc Psychiatry.* Jan 1987;26(1):87-91.

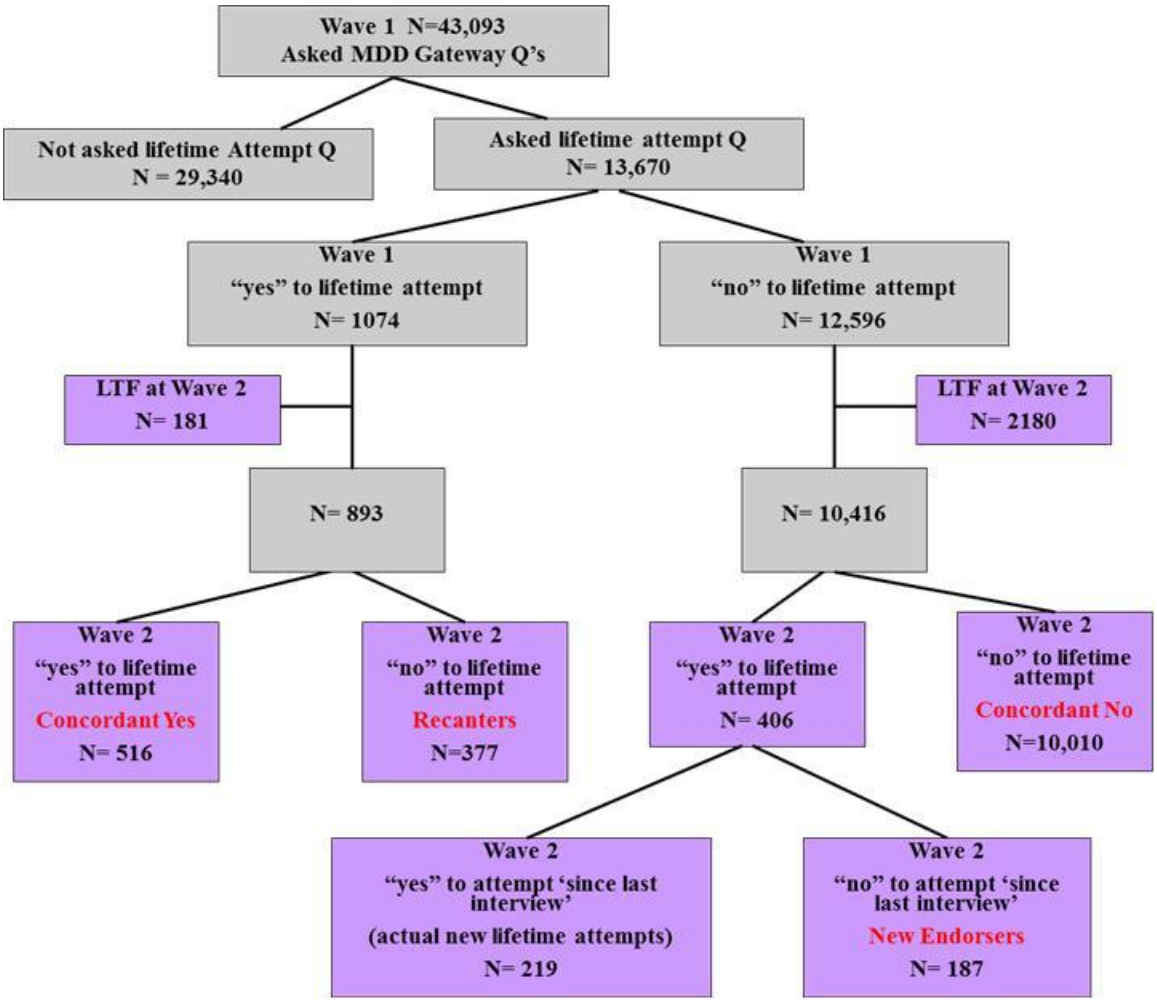
32. Baca-Garcia E, Perez-Rodriguez MM, Keyes KM, et al. Suicidal ideation and suicide attempts in the United States: 1991-1992 and 2001-2002. *Mol Psychiatry.* Sep 9 2008.

33. Kessler RC, Berglund P, Borges G, Nock M, Wang PS. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990-1992 to 2001-2003. *Jama.* May 25 2005;293(20):2487-2495.

34. Borges G, Angst J, Nock MK, Ruscio AM, Walters EE, Kessler RC. A risk index for 12-month suicide attempts in the National Comorbidity Survey Replication (NCS-R). *Psychol Med.* Dec 2006;36(12):1747-1757.

35. Kuo WH, Gallo JJ, Tien AY. Incidence of suicide ideation and attempts in adults: the 13-year follow-up of a community sample in Baltimore, Maryland. *Psychol Med.* Oct 2001;31(7):1181-1191.

<u>Figure 2A</u>: Flow chart detailing ascertainment of suicide attempts and groups of interest: "Recanters", "Concordant yes", "Concordant no", and "New endorsers" responders.



MDD = Major Depressive Disorder
LTF = Lost to follow-up

MDD Gateway Questions: 1) "In your entire life, have you ever had a time when you felt sad, blue, or down most of the time for at least 2 weeks?"; 2) "In your entire life, have you ever had a time, lasting at least 2 weeks, when you didn't care about the things that you usually cared about, or when you didn't enjoy the things you usually enjoyed?"

Table 2A: Sample characteristics at Wave 2

		Sample of interest (Concordant yes, Concordant no, Recanters, New endorsers) (n = 11090) N (Weighted %)	NESARC sample (n =34,653) N (Weighted %)
Demographics			
Age	18-29	1550 (15.4)	4913 (15.7)
	30-64	7531 (69.2)	22563 (66.3)
	65 +	2009 (15.4)	4913 (18.0)
Sex	Female	7377 (62.4)	20089 (52.8)
	Male	3713 (37.6)	14564 (47.2)
Education	≤ HS	1617 (12.6)	5514 (13.5)
	> HS	9473 (87.4)	29139 (86.5)
Income	0-34,999	7296 (66.9)[∂]	21886 (63.0) *
	35 -69	2070 (20.2)	6826 (21.4)
	70 +	1267 (12.9)	4551 (15.6)
Marital status	Never	2145 (16.8)	6638 (17.0)
	Married	5354 (58.9)	18866 (64.3)
	Divorced	3591 (24.3)	9149 (18.7)
Mental Health			
Lifetime Suicidal ideation	No	5780 (52.7)	5780 (48.0) **
	Yes	5270 (47.3)	6409 (52.0)
Lifetime Depressive D/O	No	4459 (39.7)	26375 (76.4)
	Yes	6631 (60.3)	8278 (23.6)
Lifetime Anxiety disorder	No	5961 (53.3)	24786 (71.9)
	Yes	5129 (46.7)	9867 (28.1)
Lifetime Substance use D/O	No	6366 (54.6)	22569 (62.1)
	Yes	4724 (45.4)	12084 (37.9)
Lifetime Personality D/O	No	8639 (78.3)	29686 (86.3)
	Yes	2451 (21.7)	4967 (13.7)
Attempt Characteristics[Ω]			
Age at first attempt	< 18	27 (43.2)	519 (44.79)
	18-44	387 (53.6)	661 (50.63)
	>44	25 (3.2)	60 (4.58)
Age at last attempt	< 18	172 (28.7)	347 (31.06)
	18-44	456 (63.5)	786 (61.37)
	>44	61 (7.8)	103 (7.57)
Attempt Recency (yrs)	0-10	262 (38.4)	483 (39.72)
	11-20	184 (29.6)	324 (27.63)
	>20	244 (32.0)	434 (32.65)
Number of Attempts	1	216 (31.2)	424 (34.35)
	>1	475 (68.8)	819 (65.65)

* N= 33,263

** N= 12,189 reduced because at Wave 1, only individuals who endorsed one of the two MDD gateway questions were asked about suicidal ideation

† N = 1,031

∂ N = 10,633

Ω Attempt characteristics are only available for Concordant yes and New endorser individuals within "Sample of interest" (n~703). In the "NESARC sample", attempt characteristics are only available for those who reported a lifetime attempt at Wave 2 (n~1240).

Table 2B: Reliability of suicide attempt reports across two waves, 3 years apart

		Wave 2 (W2) Attempt Report			
		Yes	**No**	**Total**	**%**
Wave 1 (W1) Attempt Report	**Yes**	**Concordant yes (a) 516**	**Recanters (b) 377**	**893**	**b/ (a + b) = 42%**
	No	**New endorsers (c) 187**	**Concordant no (d) 10,010**	**10, 197**	**c/ (c + d) = 1.8%**
	Total	**713**	**10,444**		

Kappa = 0.51
(weighted for cluster sampling)

Table 2C: Reliability across levels of predictor variables

		Kappa
Age	18-29	0.47
	30-64	0.53
	65+	0.46
Gender	Male	0.51
	Female	0.52
Education	≤ HS	0.49
	> HS	0.52
Income ($)	0 – 34,999	0.53
	35-69	0.37
	70+	0.54
Marital Status	Never	0.51
	Married	0.50
	Divorced	0.56
W2 Lifetime suicidal ideation	Yes	0.52
	No	0.13

W2 Lifetime depressive disorder	Yes	0.52
	No	0.39
W2 Lifetime anxiety disorder	Yes	0.53
	No	0.46
W2 Lifetime substance use disorder	Yes	0.54
	No	0.49
W2 Lifetime personality disorder	Yes	0.58
	No	0.40

Table 2D: Multinomial logistic regression models examining odds of being a) Discordant vs. a Concordant no individual and b) Discordant vs. a Concordant yes individual

	OR (95% C.I.) Discordant vs. Concordant no	**Adjusted OR (95% C.I.)**	**OR (95% C.I.) Discordant vs. Concordant yes**	**Adjusted OR (95% C.I.)**
Demographics				
Age 18-29	----	---	---	
30-64	* 0.63 (0.48, 0.83)	0.72 (0.52, 1.00)	0.75 (0.50, 1.13)	0.64 (0.40, 1.04)
65 +	* 0.24 (0.16, 0.36)	* 0.33 (0.20, 0.54)	1.20 (0.64, 2.26)	0.76 (0.37, 1.55)
Sex Female	---		---	
Male	0.80 (0.63, 1.01)	0.83 (0.64, 1.07)	1.16 (0.84, 1.61)	1.37 (0.98, 1.93)
Edu ≤ HS	---		---	
> HS	* 0.60 (0.44, 0.81)	* 0.72 (0.52, 0.99)	0.78 (0.52, 1.18)	0.64 (0.41, 1.00)
Inc 0-34,999	---		---	
35 -69	* 0.58 (0.42, 0.80)	* 0.68 (0.49, 0.96)	1.77 (0.98, 2.81)	1.67 (0.99, 2.67)
70 +	* 0.36 (0.23, 0.55)	* 0.52 (0.33, 0.80)	1.06 (0.60, 1.88)	0.96 (0.52, 1.79)
Marital Never	---		---	
Married	* 0.72 (0.54, 0.96)	1.08 (0.78, 1.50)	1.16 (0.79, 1.73)	1.19 (0.75, 1.88)
Divorced	0.91 (0.65, 1.27)	1.29 (0.89, 1.87)	0.87 (0.56, 1.37)	0.99 (0.59, 1.65)
Mental Health (W2 lifetime)				
Suicidal ideation	*14.21 (9.35,21.60)	*10.53(6.71, 16.53)	*0.08 (0.03, 0.23)	* 0.13 (0.05, 0.37)
Depressive D/0	*5.02 (3.68, 6.86)	* 1.74 (1.24, 2.44)	0.63 (0.39, 1.01)	0.88 (0.54, 1.46)
Anxiety disorder	* 2.63 (2.12, 3.25)	* 1.29 (1.01, 1.64)	* 0.68 (0.49, 0.94)	0.97 (0.69, 1.36)
Substance use D/O	* 2.47 (1.98, 3.09)	* 1.39 (1.08, 1.78)	* 0.62 (0.45, 0.85)	0.78 (0.56, 1.09)
Personality D/O	* 3.23 (2.59, 4.03)	*1.51 (1.17, 1.94)	* 0.38 (0.28, 0.52)	* 0.39 (0.28, 0.55)

* $p < 0.05$

Table 2E: Binomial logistic regression model examining odds of being a New Endorser vs. Recanter

		OR (95% C.I.) New Endorser vs. Recanter	Adjusted OR (95% C.I)
Demographics			
Age	18-29	----	---
	30-64	*2.46 (1.39, 4.37)	*2.10 (1.04, 4.28)
	65 +	0.66 (0.22, 2.04)	0.75 (0.21, 2.60)
Sex	Female	---	---
	Male	0.65 (0.39, 1.08)	0.71 (0.40, 1.26)
Education	≤ HS	---	---
	> HS	1.32 (0.70, 2.49)	1.58 (0.88, 2.83)
Income	0-34,999	---	---
	35 - 69	1.06 (0.57, 1.99)	1.04 (0.53, 2.03)
	70 +	0.88 (0.32, 2.38)	0.79 (0.26, 2.35)
Marital status	Never	---	---
	Married	* 2.11 (1.12, 3.97)	1.46 (0.70, 3.04)
	Divorced	* 2.02 (1.05, 3.90)	1.52 (0.71, 3.27)
Mental Health (W2 lifetime)			
Suicidal ideation		*0.23 (0.10, 0.57)	*0.21 (0.08, 0.55)
Depressive disorder		1.05 (0.58, 1.91)	0.79 (0.42, 1.49)
Anxiety disorder		* 2.96 (1.78, 4.93)	*2.48 (1.42, 4.35)
Substance use disorder		1.10 (0.65, 1.86)	1.01 (0.59, 1.72)
Personality D/O		* 2.75 (1.69, 4.45)	*2.71 (1.48, 4.96)

* $p < 0.05$

Table 2F: Multinomial logistic regression models examining odds of being a) Recanter vs. a Concordant yes individual and b) New Endorser vs. a Concordant yes individual

		OR (95% C.I.) Recanter vs. Concordant yes	Adjusted OR (95% C.I.)	OR (95% C.I.) New Endorser vs. Concordant yes	Adjusted OR (95% C.I.)
Demographics					
Age	18-29	---		---	
	30-64	*0.58 (0.38, 0.89)	*0.52 (0.31, 0.86)	1.43 (0.79, 2.60)	1.09 (0.52, 2.28)
	65 +	1.29 (0.67, 2.48)	0.79 (0.38, 1.66)	0.86 (0.27, 2.73)	0.59 (0.16, 2.12)
Sex	Female	---		---	
	Male	1.33 (0.92, 1.92)	*1.53 (1.04, 2.25)	0.87 (0.54, 1.38)	1.09 (0.65, 1.83)
Education	≤ HS	---		---	
	> HS	0.72 (0.45, 1.13)	*0.56 (0.34, 0.91)	0.95 (0.52, 1.73)	0.88 (0.50, 1.55)
Income	0-34,999	---		---	
	35 - 69	*1.74 (1.01, 2.98)	1.65 (0.96, 2.86)	* 1.84 (1.04, 3.25)	1.72 (0.94, 3.15)
	70 +	1.11 (0.61, 2.03)	1.04 (0.53, 2.04)	0.97 (0.37, 2.54)	0.82 (0.29, 2.32)

Marital Never	---		---	
Married	0.94 (0.61, 1.45)	1.08 (0.66, 1.76)	* 1.99 (1.10, 3.60)	1.57 (0.77, 3.20)
Divorced	0.72 (0.43, 1.20)	0.88 (0.49, 1.60)	1.45 (0.81, 2.61)	1.34 (0.67, 2.66)
Mental Health (W2 lifetime)				
Suicidal ideation	*0.16 (0.05, 0.52)	*0.25 (0.07, 0.82)	*0.04 (0.01, 0.11)	*0.05 (0.02, 0.15)
Depressive D/O	0.62 (0.37, 1.03)	0.94 (0.54, 1.61)	0.65 (0.35, 1.22)	0.74 (0.39, 1.41)
Anxiety D/O	*0.50 (0.36, 0.70)	0.77 (0.53, 1.10)	1.48 (0.90, 2.46)	*1.90 (1.09, 3.30)
Substance use D/O	*0.60 (0.43, 0.84)	0.78 (0.55, 1.12)	0.66 (0.4, 1.08)	0.79 (0.47, 1.33)
Personality D/O	*0.27 (0.19, 0.39)	*0.28 (0.19, 0.42)	0.74 (0.48, 1.15)	0.77 (0.45, 1.31)
Attempt Characteristics§				
First attempt < 18			---	
18-44			* 0.60 (0.40, 0.90)	
>44			1.35 (0.51, 3.55)	
Last attempt < 18			---	
18-44			* 0.52 (0.33, 0.82)	
>44			0.82 (0.33, 2.02)	
Recency (yrs) <10			---	
11-20			1.52 (0.85, 2.71)	
>20			* 2.45 (1.38, 4.32)	
# of Attempts 1			---	
>1			* 0.54 (0.33, 0.89)	

* $p < 0.05$

§ Since data on suicide characteristics were only available for New endorsers and Concordant yes responders, they were not included in the adjusted models due to empty cells in the 4-category outcome

Chapter 3

Mood Matters:

Investigating the effects of respondent mood on the reporting of suicide attempts

Introduction

Full ascertainment of suicide attempts in the population is critical for the accurate assessment of incidence, prevalence, and etiology; all of which inform suicide prevention efforts. However, as established in the first two chapters of this dissertation, self-reports of suicide attempts are plagued by a fair to moderate degree of unreliability. [1-7] When surveyed about past suicide attempts, approximately 45% of adult respondents give discordant responses over time. The question that remains however, is *why* might a respondent report or not report a past attempt? A commonly posited reason for the differential reporting is recall failure;[3,8,9] an explanation supported by the much greater frequency of recanting (i.e. reporting an attempt at Time 1, but not reporting at Time 2) compared with newly reporting attempts over time, as was found in Chapter 2. Further, some findings suggest that recall failure may not be random, yet systematically influenced by the mood of the respondent at the time of the interview.[8-10] There are three ways in which a currently depressed mood may affect suicide attempt reporting; a depressed mood may increase, decrease, or have no effect on the likelihood that a respondent recalls, and hence reports an attempt. By examining competing hypotheses, this paper will attempt to elucidate the most prominent mood effect.

A depressed mood at the time of the interview may *increase* the likelihood that a respondent reports a past attempt. Three decades of research have been devoted to the phenomenon of 'mood congruent recall', which purports that mood influences memory by priming content that is congruent with one's mood, which in turn leads to selective retrieval of mood-congruent information. [11,12] Therefore, when a respondent is currently depressed, they may have an increased likelihood of recalling past negatively-valenced events, such as a suicide attempt. This theory has been repeatedly supported experimentally, in which either the mood or

memory was manipulated by the investigator.[13-20] There have also been supportive findings within studies involving autobiographical memories and naturally occurring mood states;[21-26] many examining mood-congruent recall of past depressive episodes or symptoms[21,22,27-31] and three[8-10] which examined mood-congruent recall of past suicidal ideation. While there is therefore a strong foundation for potential mood-congruent reporting of attempts, there are some limitations of the literature. Evidence for mood-congruent recall theory is certainly not unequivocal, with many studies finding no effect of depressed mood on recall of negative events.[32-36] In addition, the few studies that found mood-congruent reporting of suicidal ideation were in small, non population-based youth samples and perhaps most importantly, did not control for any other known predictors of discordant reporting, for example, age, gender, and suicidal ideation.

Alternatively, a depressed mood at the time of the interview may *decrease* the likelihood that a respondent reports a past attempt. General memory deficit effects, particularly with regards to episodic memory, are very typical in individuals with depression.[37-46] Over the last two decades, research has demonstrated that depressed patients have "overgeneral" autobiographical memories, meaning, they recall broad categories of events, rather than specific events. A range of mechanisms may be involved in this memory impairment, including reduced executive capabilities,[47]as well as affect regulation by way of preventing individuals from recalling events that evoke painful, negative emotions.[48] The link between depression and overgeneral memory was first reported among suicidal patients[49], but has since been widely documented in numerous samples and settings.[25,50-55] However, we believe there are limits regarding the application of this theory to the reporting of suicide attempts. First, a large majority of the research has been conducted using a word-cuing task (e.g. "Think of a time

recently that made you sad") rather than through standard survey methods, and overgeneral memory may not be enacted when probed about specific events. Second, many studies show that overgeneral memory operates predominantly regarding positive and neutral-valenced events, rather than negative-valenced events such as a suicide attempt.[44,49,54-56] Lastly, whether a suicide attempt is "specific" enough to be forgotten by an overgeneral memory is debatable; a respondent could remember attempting (and hence, report it), but simply not recall the specific details. Together, we believe the aforementioned reasons cast reasonable doubt on the potential for memory deficit effects influencing the reporting of attempts.

Finally, a depressed mood at the time of interview may have no effect on the likelihood of recalling a past attempt. One mood-independent factor that has been posited to influence attempt recall, is attempt severity.[3] Salient (von Restorff Effect)[57,58] and intense events (Affective Intensity Effect)[59,60] are more memorable than less salient and intense ones, regardless of mood at the time of memory retrieval. Therefore, the severity of the attempt, comprising both the psychological intent to die and medical consequences, may lead to differential recall; specifically, more severe attempts are more likely to be remembered, and therefore less likely to be inconsistently reported. Yet, this particular mood-independent explanation lacks an important element present in the mood-dependent theories. Unlike respondent mood, which may vary over time, the severity of the attempt is constant. Therefore, why would a respondent remember a non-severe attempt at one time point, and not another? Because of the invariance in attempt severity within individuals, we do not believe it can fully explain discordant reports over time.

Against this background, we will examine in what manner and to what extent, a respondent's mood at the time of the interview affects their likelihood of reporting a past suicide attempt. There are three potential mood-recall effects; 1) depressed mood enhancing recall (i.e.

mood-congruent effects); 2) depressed mood inhibiting recall (i.e. memory deficit effects); and 3) no effect of mood on recall of past attempts (i.e. mood-independent effects). Considering the strength of each body of literature, and respective theoretical applicability, we hypothesize that there will be mood-congruent effects on the reporting of past suicide attempts; that is, respondents who are currently depressed will be more likely to report a past attempt at that interview.

Current mood has yet to be examined as a potential factor influencing the reporting of suicide attempts in adults. To investigate this critically important issue, we will use a longitudinal study design in which a large, population-based sample of adult respondents were queried about their past attempts and current mood at two waves, separated by 3 years.

MATERIALS AND METHODS

Sample

The data used in this investigation was from the National Survey on Alcohol and Related Conditions (NESARC), developed by the *National Institute on Alcohol Abuse and Alcoholism* (NIAAA). An overview of the study design has been described elsewhere.[61] In brief, it is a longitudinal survey with its first wave of interviews fielded in 2001-2002 and second wave in 2004-2005. The NESARC is a representative sample of the non-institutionalized US population 18 years of age and older. Respondents were informed in writing about the nature of the survey, the statistical uses of the survey data, the voluntary aspect of participation and the federal laws that protect the confidentiality of the identifiable survey information. Those respondents consenting to participate after receiving this information were interviewed in person. A sample of 43,093 participants was obtained for Wave 1 with a response rate of 81%. A total of 34,653

respondents were re-interviewed at Wave 2, providing a follow-up rate of 87% and a cumulative response rate of 70.2%.

The NESARC used a multistage sampling design that oversampled specific ethnic groups, including African Americans and Hispanics, as well as young adults. Data were weighted to account for oversampling of specific groups, clustered sampling, and non-response. The data were weighted to be representative of the US civilian population on the basis of the 2000 Decennial Census of Population and Housing. The specific aspects of the sampling design of the NESARC are described elsewhere in detail.[62]

Measures

At each wave, the NESARC collected detailed information on basic sociodemographic characteristics, the presence and timing of mental health symptoms and diagnoses, and information on lifetime suicide attempts. Participants were interviewed with computer-assisted face-to-face interviews using the Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV) diagnostic survey.[61] This survey took approximately 1 hour, and was administered by either lay interviewers or clinicians.

Lifetime suicide attempts

Lifetime suicide attempts, our outcome of interest, were assessed at both Waves 1 and 2. As shown in Figure 3A, at Wave 1, the suicide attempt question was only asked of individuals who answered "yes" to at least one of the two "gateway" or "stem" questions for a diagnosis of Major Depressive Disorder (MDD). Respondents were asked the following MDD stem questions: 1) "In your entire life, have you ever had a time when you felt sad, blue, or down most of the time

for at least 2 weeks?", and 2) "In your entire life, have you ever had a time, lasting at least 2 weeks, when you didn't care about the things that you usually cared about, or when you didn't enjoy the things you usually enjoyed?" Any respondent who answered "yes" to at least one of the two questions was then asked the following suicide attempt question: "During that time when your mood was at its lowest or you enjoyed or cared the least about things, did you attempt suicide?" At Wave 2, all respondents, regardless of their response to the MDD gateway questions at Wave 1, were asked the following suicide attempt question: "In your ENTIRE life did you EVER attempt suicide?" The question regarding a lifetime suicide attempt was therefore present in the survey at both Waves 1 and 2. However, because individuals were not asked this question at Wave 1 if they did not respond "yes" to at least one of the two MDD stem questions, this study necessarily focuses only on the subgroup of individuals who were asked the suicide question at Wave 1.

Attempter profiles

See Figure 3A. Individuals who reported a lifetime attempt at Wave 1 and reported no lifetime attempt at Wave 2 were labeled ***"Recanters"*** (n=377). Individuals who reported no lifetime attempt at Wave 1, but newly reported a lifetime attempt at Wave 2, and reported that the attempt had not taken place between Waves 1 and 2, were labeled as ***"New endorsers"*** (n=187). Individuals who consistently reported a lifetime suicide attempt at both Waves 1 and 2 were labeled as ***"Concordant yes"*** (n=516), and those who consistently reported no lifetime suicide attempt at both Waves 1 and 2 were labeled as ***"Concordant no"*** (n=10,010).

Respondent current depressed mood

The NESARC uses the Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS-IV). The AUDADIS-IV is a fully structured assessment administered by trained lay interviewers, and has demonstrated good test-retest reliability on mental health outcomes.[61] Our current depressed mood predictor was **"*Depressed mood*",** based on one question from the 12-Item Short-Form Survey Instrument (SF-12)[63]; How much of the time during the past 4 weeks have you felt downhearted and depressed? The response options were the following: all of the time; most of the time; some of the time; a little of the time; none of the time (see Appendix 3A). We chose to dichotomize this variable as (Yes = all or most of the time) vs. (No= some, little or none of the time) for two reasons; 1) relative homogeneity of effect within dichotomized strata, and parsimony across strata (see Appendix 3B) and 2) if a respondent was depressed at least most of the time within the past 4 weeks, they were likely to have been depressed during or near the time of the interview, which was our intended construct of interest. Current mood was assessed at both Waves 1 and 2. Using the mood assessments at each wave, we created four ***"Wave 1:Wave 2 depressed mood profiles"***, in the format of D_{xy}; where x= current depression status at Wave 1 (1= depressed, 0= not depressed), and y= depression status at Wave 2 (1= depressed, 0= not depressed). For example D_{10} denotes a respondent that was depressed at Wave 1 and not depressed at Wave 2 (see Appendix 3C).

Covariates

Sociodemographic and mental health variables found to be associated with the reporting of suicidality (ideation, attempts, plans, gestures) in Chapters 1 and 2, were considered potential covariates, including: *age* (18-29, 30-64, 65+)*; sex; education level (*≤ HS, >HS); and *lifetime*

suicidal ideation (yes/no).* (see Appendices 3D and 3E). Lifetime suicidal ideation was assessed at both Waves 1 and 2. At Wave 1, respondents in our sample were asked the following: 1) Did you ever think about committing suicide?; 2) Did you ever feel like you wanted to die?; and 3) Did you ever think a lot about your own death? At Wave 2, respondents were asked about their ideation "since the last interview". Therefore, if respondents answered 'yes' to at least one of the three suicidal ideation questions, at either wave, they were classified as having had *lifetime suicidal ideation* at Wave 2 (see Appendix 3F).

Data Analysis

The sample for all analyses consisted of respondents who endorsed at least one of the two MDD gateway questions at Wave 1, and therefore were asked the lifetime suicide attempt questions at Wave 1, and were not lost to follow-up prior to Wave 2 (n = 11,090). Sample characteristics can be found in Table 3A. We analyzed Recanters and New endorsers separately, using the same analytic approach. The analyses with regards to Recanters will be presented first.

Recanters

We first assessed the odds of being a Recanter (*ref* = Concordant yes), given respondent depressed mood at *i*) Wave 1 and *ii*) Wave 2, in separate binomial logistic regression models, to obtain the bivariate associations. We then adjusted each model for 1) demographic covariates (age, sex, education) (see AOR^1) and 2) demographic covariates and lifetime suicidal ideation (see AOR^2). We considered AOR^2 models as our final adjusted models. We conducted this analysis to assess which mood recall theory or theories, were implicated with regards to recanting an attempt over time. For example, as shown in Table 3B, if Recanters were found to

* All covariates were W2 measurements

be less likely than Concordant yes individuals to have a depressed mood at Wave 2 (OR < 1.0) then mood-congruent theory would be supported; however if they were found to be more likely to have a depressed mood at Wave 2 (OR > 1.0), then mood-deficit theory would be supported. Any findings not predicted by either mood-congruent or mood-deficit theory would be by definition mood-independent effects; that is, knowing a respondent's mood at a particular wave would not allow one to predict if they reported an attempt at that wave.

We also assessed the odds of being a Recanter (*ref* = New endorsers, Concordant yes & Concordant No)[*], given specific "Wave 1:Wave 2 mood profiles" (D_{xy}). In separate binomial logistic regression models, we examined the odds of a Recanter being *i*) D_{10} (*ref* = $D_{01,}$ $D_{00,}$ D_{11})[**]; *ii*) D_{01} (*ref* = $D_{10,}$ $D_{00,}$ D_{11}) [**]; and *iii*) D_{10} (*ref* = D_{01}). We then adjusted each model for 1) demographic covariates (age, sex, education) (see AOR^1) and 2) demographic covariates and lifetime suicidal ideation (see AOR^2). We considered AOR^2 models as our final adjusted models. This set of analyses was to assess the *degree* to which each mood theory was supported. For example, as shown in Table 3C, if Recanters were more likely than any other attempter profile to be depressed at Wave 1 and not depressed at Wave 2, then it would indicate that mood-congruent theory was supported, and the odds ratio would reflect the strength of that theory; however, if Recanters were more likely to be not depressed at Wave 1, and depressed at Wave 2, then it would indicate that mood-deficit theory was supported, and to what degree. Finally, by comparing D_{10} vs. D_{01}, we are able to assess how likely Recanters are to respond in a mood-congruent or mood-deficit manner, compared to one another.

[*] Analyses were also conducted using only the Concordant yes and Concordant no respondents as the reference group, see Appendix 3G

[**] Analyses were also conducted using only D_{00} and D_{11} as the reference group, see Appendix 3G

New endorsers

We then conducted the same two sets of analyses for New endorsers. We first assessed the odds of being a New endorser (*ref* = Concordant yes), given respondent depressed mood at *i*) Wave 1 and *ii*) Wave 2, in separate binomial logistic regression models, and adjusting for all covariates in the final models. As shown in Table 3B, if New endorsers were found to be less likely than Concordant yes individuals to have a depressed mood at Wave 1 (OR < 1.0), then mood-congruent theory would be supported; however if they were found to be more likely to have a depressed mood at Wave 1 (OR > 1.0), then mood-deficit theory would be supported. Once again, anything not predicted by either mood-congruent or mood-deficit theory would indicate by definition, mood-independent effects.

Lastly, we assessed the odds of being a New endorser (*ref* = Recanter, Concordant yes, & Concordant No)*, given specific "Wave 1:Wave 2 mood profiles" (D_{xy}). In separate binomial logistic regression models, we examined the odds of a New endorser being *i*) D_{10} (*ref*= $D_{01,}$ $D_{00,}$ D_{11})**; *ii*) D_{01} (*ref* = $D_{10,}$ $D_{00,}$ D_{11})**; and *iii*) D_{01} (*ref* = D_{10}). As shown in Table 3C, if New endorsers were more likely than any other attempter profile to be depressed at Wave 1 and not depressed at Wave 2, then it would indicate that mood-deficit theory was supported, and the odds ratio would reflect the strength of that theory; however if New endorsers were more likely to be not depressed at Wave 1, and depressed at Wave 2, then it would indicate that mood-congruent theory was supported, and to what degree. Finally, by comparing D_{01} vs. D_{10}, we are able to

* Analyses were also conducted using only the Concordant yes and Concordant no respondents as the reference group, see Appendix 3G

** Analyses were also conducted using only D_{00} and D_{11} as the reference group, See Appendix 3G

assess how likely Recanters are to respond in a mood-congruent or mood-deficit manner, compared to one another.

All analyses were conducted using SUDAAN version 10 to account for the complex sampling structure, and significance level was set at $\alpha = 0.05$.

RESULTS

Our aim in this study was to examine possible effects of respondent mood at the time of reporting on the likelihood of reporting an attempt at that interview. We tested for effects of mood on reporting of attempts separately for Recanters and New endorsers in order to detect potential differences in reporting across the two groups of discordant responders.

Recanters

As shown in Table 3D, we found that Recanters were as likely as Concordant yes individuals to have been depressed at Wave 1 (AOR, 0.91; 95% CI, 0.61-1.37), however, were less likely to have been depressed at Wave 2 (AOR, 0.59; 95% CI, 0.39-0.90). Comparing these findings with what would have been expected according to each theory (as shown in Table 3B), indicates clear mood-congruent effects.

To determine how likely Recanters were to report in a mood-congruent manner, we examined Recanters' odds of specific Wave 1: Wave 2 mood profiles. As shown in Table 3E, we found that Recanters were more likely than any other attempter profile (i.e. New endorsers, Concordant yes, and Concordant no individuals) to be D_{10}; that is, depressed at Wave 1 and not depressed at Wave 2, compared to all other mood profiles combined (AOR, 1.62; 95% CI, 1.13-2.33). Further, we found that Recanters were as likely as all other attempter profiles to be D_{01};

that is, not depressed at Wave 1 and depressed at Wave 2 (AOR, 0.61; 95% CI, 0.36-1.02). Together, when comparing these findings to those expected according to each mood theory (as shown in Table 3C), we once again find clear mood-congruent effects. Specifically, Recanters were 62% more likely to report attempts in a mood-congruent manner compared with any other attempter and mood profile, and no more likely to report in a mood-deficit manner. Lastly, we found that Recanters were two and a half times as likely to report in a mood-congruent (i.e. D_{10}) manner, than a mood-deficit manner (D_{01}) (AOR, 2.50; 95% CI, 1.39-4.49).

New endorsers

As shown in Table 3D, we found that New endorsers were as likely as Concordant yes individuals to be depressed at Wave 1 (AOR, 1.12; 95% CI, 0.69-1.81), however more likely to have been depressed at Wave 2 (AOR, 1.68; 95% CI, 1.05-2.67). Comparing these findings with what would have been expected according to each theory (as shown in Table 3B), since not consistent with either mood-congruent nor mood-deficit theory, by definition, these findings indicate mood-independent reporting effects. That is, simply knowing that a respondent was depressed at Wave 1 does not help one determine if they reported an attempt at Wave 1, since New endorsers and Concordant yes individuals were equally likely to be depressed. Accordingly, simply knowing that a respondent was depressed at Wave 2 does not help one determine if they reported an attempt at Wave 2, since both New endorsers and Concordant yes individuals reported an attempt at that wave.

To determine the strength of support for each theory, we examined New endorsers' odds of specific Wave 1: Wave 2 mood profiles. As shown in Table 3E, we found that New endorsers were as likely as any other attempter profile, to be D_{10}; that is, depressed at Wave 1 and not

depressed at Wave 2, compared to all other mood profiles combined (AOR, 1.59; 95% CI, 0.91-2.78), however more likely to be to be D_{01}; that is, not depressed at Wave 1 and depressed at Wave 2 (AOR, 2.71; 95% CI, 1.71-4.31). Together, when comparing these findings to those expected according to each mood theory (as shown in Table 3C), we find clear mood-congruent effects. Specifically, New endorsers were almost 3 times as likely to report attempts in a mood-congruent manner compared with any other attempter and mood profile, and no more likely to report in a mood-deficit manner. However, when we compared the strength of each mood effect theory with one another, we found that New endorsers were no more likely to report in a mood-congruent manner (i.e. D_{01}), than a mood-deficit manner (i.e. D_{10}) (AOR, 1.56; 95% CI, 0.80-3.06). Therefore, once again, New endorsers appear to report attempts in a mood-independent manner.

In summary, we found clear mood-congruent reporting effects among Recanters, however, mood-independent reporting effects among New endorsers. These disparate findings across the two groups of discordant reporters indicate potential distinctions between the groups, which will be explored further in our discussion.

DISCUSSION

Our aim was to assess how, and to what degree, a respondent's mood at the time of the interview influences their likelihood of reporting a past suicide attempt. We hypothesized that we would detect a mood-congruent pattern of reporting, meaning, respondents who had a depressed mood at the time of the interview would be more likely to report a past attempt than respondents who were not currently depressed. Our findings were only partially consistent with this hypothesis; specifically, Recanters reported in a way most consistent with mood-congruent theory, while

New endorsers appeared to report in a mood-independent manner. However, given that discordant respondents are over 20 times more likely to recant than newly endorse an attempt, thereby making recanting a far more likely reporting pattern, we believe that discordant reporters on a whole, can be typified as following a mood-congruent pattern of reporting attempts.

The mood-congruent findings found amongst Recanters are consistent with three studies that examined mood and reporting of past suicidal ideation. [8-10] Aside from being the first study to examine this issue in regards to suicide attempt reporting, these findings add to the broader body of evidence in a number of other ways. First, this study sample comprised adults (18 yrs and older), while all prior examinations have been conducted only among youth samples. Secondly, this study was able to test these theories among not just Recanters, but also New endorsers; the majority of studies only capture the former discordant group. Lastly, and perhaps most importantly, unlike prior studies, these findings controlled for other known predictors of discordant reporting, hence reducing the possibility for confounding. We believe consistent mood-congruent findings across varying suicidal constructs, ages, types of discordant responders, and assessments of current depressed mood, lends robustness to our findings, and mood-congruent recall theory as a whole.

While mood-congruence theory was originally founded in regards to recall, we believe it may have a broader application to the reporting of attempts. Mood may have an impact beyond memory retrieval; it may influence respondents' reconstruction or subjective evaluation of details about recalled experiences.[3,45,64] For example, a respondent may report an attempt at Wave 1, but at Wave 2, reinterpret that event as not a true attempt (e.g. an accident, cry for help, lack of intent to die, etc.) and accordingly, not report it.[3] The likelihood for reinterpretation could plausibly be influenced by respondent mood, and in a mood-congruent fashion;

respondents might interpret an event as suicidal when depressed, but as non-suicidal when not depressed. Respondents who are uncertain about what exactly constitutes a suicide attempt may be particularly vulnerable to reinterpretation. Therefore, we believe our mood-congruent patterns of discordant reporting found among Recanters may indicate issues with recall or reinterpretation.

Unlike Recanters however, New endorsers did not follow a mood-congruent pattern of reporting, and in fact, reported in a mood-independent manner. Our inconsistent findings across the two discordant groups were unexpected, and lead us to believe that New endorsers may be a discordant group quite distinct from Recanters. Intuitively, this supposition is supported by the fact that respondents were far more likely to recant than newly endorse an attempt. Since respondents are more likely to forget attempts over time, and hence recant, individuals who have enhanced recall of an attempt over time are likely unique in some way. By examining differences between Recanters and New endorsers, drawn from both this chapter and Chapter 2, below, we attempt to gain a clearer picture of New endorsers.

Our finding that New endorsers were more likely than Concordant yes individuals to be currently depressed at Wave 2, when both groups reported an attempt, was unexpected. As shown in Appendix 3H, New endorsers were also more likely to have had a depressive disorder (MDD or Dysthymia) within the past year. Together, we believe these findings could indicate that some New endorsers were actually incident (ie. new-onset) attempters, and hence were more likely to have had recent depression. In order to be correctly classified as a New endorser, respondents had to accurately determine if their attempt was “since the last interview” or “before the last interview”. Therefore, some New endorsers may have actually attempted soon after the first interview (hence, an incident attempt), but misremembered it as before the last interview

(misclassifying them as a New endorser). However, contrary to this supposition, in Chapter 2, we found that New endorsers were younger at their first and last attempts, were more likely to have attempted more than 20 years ago, and have attempted only once, compared with Concordant yes individuals. These findings would therefore greatly challenge the plausibility that many New endorsers were in fact incident attempters. In Chapter 2, we also reported that New endorsers were much more likely than Recanters to have a lifetime personality or anxiety disorder, but much less likely than Recanters to have had past suicidal ideation. We are unsure of what these findings indicate, but propose some interpretations. It's plausible that a respondent must be currently or recently depressed (and more likely than Concordant yes individuals), and perhaps severely so, in order to either remember a past attempt that they did not recall at Wave 1, or interpret an event as an attempt at Wave 2, which they did not categorize as such at Wave 1. Further, New endorsers' greater likelihood of a past anxiety or personality disorder, coupled with their decreased likelihood of prior ideation when compared with Recanters, may indicate that New endorsers are more unstable and impulsive individuals than Recanters. Individuals who are currently depressed, with impulsive and anxious traits, may be more likely to reinterpret an event as suicidal over time. This however is only speculative, and one possible interpretation given the data at hand. While differences in the nature of these discordant groups may have led to our disparate findings between Recanters and New endorsers, it is also possible that they are artifactual and arose from differences in power. We in fact did find that New endorsers were significantly likely to report in a mood-congruent fashion, however simply not significantly more than a mood-deficit manner. Given that the sample of Recanters was more than twice as large as that of New endorsers, the latter insignificant finding may have resulted from reduced power.

Aside from the limited power within the New endorsers group, this investigation had some other limitations. We were only able to assess mood effects on attempt reports among individuals who endorsed at least one of the two MDD gateway questions at Wave 1, thereby excluding 29,340 individuals from our analysis. Unfortunately, due to the survey structure at Wave 2, we were unable to assess how many attempts were not captured due to these gateway questions. However, the NESARC's predecessor survey, the National Longitudinal Alcohol Epidemiologic Survey (NLAES), examined the gateway question issue and found that the number of individuals who did not screen into the depression section and yet still reported a suicide attempt, was very low (< 0.1% of the sample).[65] Therefore, we do not believe the gateway questions affected our external validity substantially. As with all prospective studies, we experienced loss to follow up (Figure 3A). Seventeen percent of our baseline sample was lost to follow-up before Wave 2, and these respondents differed from our final sample in two ways (see Appendix 3J). Those lost to follow-up (LTF) were less likely to have had more than a high school education, however more likely to have had a current depressed mood at Wave 1, compared with out final sample. Since there is no way of assessing LTF up differential by attempter profile, it is difficult to determine the effect of the differential LTF by depressed mood at Wave 1. However, if those individuals who would have been Recanters at Wave 2 were also more likely to be LTF, then our mood-congruent findings would be an underestimate of the true strength of the theory; but if New endorsers were more likely to be LTF, the our findings would represent an overestimate of the mood-congruent findings. Lastly, many characteristics of suicide attempts, particularly the medical and psychological severity of the attempts, were not included in the NESARC, but could certainly influence reporting reliability.

The main strength of this study is that it is the first to examine how respondent mood may affect the recall and reporting of suicide attempts. Further, we had a large and rich dataset with which to do so, allowing for a viable sample size of attempters, control for other potential covariates, and comprehensive mood and mental health measures. As our measure of current mood, we used respondents' endorsements of a downhearted and depressed mood most or all of the time during the past four weeks. A respondent who is depressed "most or all of the time" during the past four weeks is likely to be depressed during the interview. However, we also appreciate that this measure captures a relatively steady mood state for a month up until the interview. A depressed mood most or all of the time for the month prior to the interview indicates a slightly less transient, and more persistent depressed mood, in which rumination during, could also enhance recall even if a respondent was not depressed at the exact moment of the interview. Therefore, we believe this measure perhaps offers a broader perspective relative to momentary measures of mood. As a way of assessing the robustness of our findings, we also classified current depressed mood using three other measures; past year MDD, past year Dysthymia, and either past year MDD or Dysthymia (see Appendices 3I, 3J and 3K). The results from these analyses not only support our main findings, but also strengthen them; both Recanters and New endorsers followed a mood-congruent pattern of reporting. We believe the consistent results given varying measures of current mood as well as reference groups (see Appendix 3G), lends robustness to our findings.

Suicide researchers often use past reports of attempts in etiologic investigations, and it would be prudent for them to consider that respondents may be less likely to report a past attempt if they are currently not depressed. Knowledge of mood-congruent biases may allow for the refinement of attempt measures, by including measures of momentary mood as well as

mood-relevant probes. Finally, an understanding of mood effects on reporting may have broader implications for psychiatric epidemiology, as a substantial portion of the field has been devoted to elucidating how past negative events may lead to depressive disorders. Based on the quantifiable mood-related underreporting of behaviors, psychiatric researchers could potentially use this information to adjust risk estimates pertaining to negative-valenced life events and depression, as well as to adapt future survey questions.

References

1. Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EC, Ross JG. Reliability of the 1999 youth risk behavior survey questionnaire. *J Adolesc Health.* Oct 2002;31(4):336-342.

2. Brezo J, Paris J, Barker ED, et al. Natural history of suicidal behaviors in a population-based sample of young adults. *Psychol Med.* Nov 2007;37(11):1563-1574.

3. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T. The accuracy of prevalence estimations for suicide attempts. how reliably do adolescents and young adults report their suicide attempts? *Arch Suicide Res.* 2006;10(3):253-263.

4. Fendrich M, Warner V. Symptom and substance use reporting consistency over two years for offspring at high and low risk for depression. *J Abnorm Child Psychol.* Aug 1994;22(4):425-439.

5. Flisher AJ, Evans J, Muller M, Lombard C. Brief report: Test-retest reliability of self-reported adolescent risk behaviour. *J Adolesc.* Apr 2004;27(2):207-212.

6. Nock MK, Holmberg EB, Photos VI, Michel BD. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. *Psychol Assess.* Sep 2007;19(3):309-317.

7. Shaffer D, Scott M, Wilcox H, et al. The Columbia Suicide Screen: validity and reliability of a screen for youth suicide and depression. *J Am Acad Child Adolesc Psychiatry.* Jan 2004;43(1):71-79.

8. Goldney RD, Winefield AH, Winefield HR, Saebel J. The benefit of forgetting suicidal ideation. *Suicide Life Threat Behav.* Feb 2009;39(1):33-37.

9. Klimes-Dougan B, Safer MA, Ronsaville D, Tinsley R, Harris SJ. The value of forgetting suicidal thoughts and behavior. *Suicide Life Threat Behav.* Aug 2007;37(4):431-438.

10. Klimes-Dougan B. Screening for suicidal ideation in children and adolescents: methodological considerations. *J Adolesc.* Aug 1998;21(4):435-444.

11. Blaney PH. Affect and memory: a review. *Psychol Bull.* Mar 1986;99(2):229-246.

12. Bower GH. Mood and memory. *Am Psychol.* Feb 1981;36(2):129-148.

13. Bradley BP, Mogg K. Mood and personality in recall of positive and negative information. *Behav Res Ther.* Jan 1994;32(1):137-141.

14. Eich E, Macaulay D, Ryan L. Mood dependent memory for events of the personal past. *J Exp Psychol Gen.* Jun 1994;123(2):201-215.

15. Goodwin AH, Sher KJ. Effects of induced mood on diagnostic interviewing: Evidence for a mood and memory effect. *Psychological Assessment.* 1993;5(2):197-202.

16. McFarland C, Buehler R. The impact of negative affect on autobiographical memory: the role of self-focused attention to moods. *J Pers Soc Psychol.* Dec 1998;75(6):1424-1440.

17. Singer J, Salovey P. Mood and memory: Evaluating the network theory of affect. *Clinical Psychology Review.* 1988;8:211-251.

18. Teasdale JD, Fogarty SJ. Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *J Abnorm Psychol.* Jun 1979;88(3):248-257.

19. Teasdale JD, Russell ML. Differential effects of induced mood on the recall of positive, negative and neutral words. *Br J Clin Psychol.* Sep 1983;22 (Pt 3):163-171.

20. Wisco B, Nolen-Hoeksema S. The interaction of mood and rumination in depression: effects on mood maintenance and mood-congruent autobiographical memory. *J Rat-Emo Cognitive-Behav Ther.* 2009;27.

21. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. *J Health Soc Behav.* Sep 1987;28(3):232-246.

22. Ben-Zeev D, Young MA. Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: an experience sampling study. *J Nerv Ment Dis.* Apr 2010;198(4):280-285.

23. Clark DM, Teasdale JD. Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *J Abnorm Psychol.* Apr 1982;91(2):87-95.

24. Dalgleish T, Watts FN. Biases of attention and memory in disorders of anxiety and depression. *Clinical Psychology Review.* 1990;10(5):589-604.

25. Mathews A, MacLeod C. Cognitive vulnerability to emotional disorders. *Annu Rev Clin Psychol.* 2005;1:167-195.

26. Lloyd GG, Lishman WA. Effect of depression on the speed of recall of pleasant and unpleasant experiences. *Psychol Med.* May 1975;5(2):173-180.

27. Bromet EJ, Dunn LO, Connell MM, Dew MA, Schulberg HC. Long-term reliability of diagnosing lifetime major depression in a community sample. *Arch Gen Psychiatry.* May 1986;43(5):435-440.

28. Kendler KS, Gardner CO, Prescott CA. Are there sex differences in the reliability of a lifetime history of major depression and its predictors? *Psychol Med.* May 2001;31(4):617-625.

29. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. The lifetime history of major depression in women. Reliability of diagnosis and heritability. *Arch Gen Psychiatry.* Nov 1993;50(11):863-870.

30. Thompson R, Bogner HR, Coyne JC, Gallo JJ, Eaton WW. Personal characteristics associated with consistency of recall of depressed or anhedonic mood in the 13-year follow-up of the Baltimore Epidemiologic Catchment Area survey. *Acta Psychiatr Scand.* May 2004;109(5):345-354.

31. Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychol Med.* Aug 2004;34(6):1001-1011.

32. Banos RM, Medina PM, Pascual J. Explicit and implicit memory biases in depression and panic disorder. *Behav Res Ther.* Jan 2001;39(1):61-74.

33. Bazin N, Perruchet P, Feline A. Mood congruence effect in explicit and implicit memory tasks: a comparison between depressed patients, schizophrenic patients and controls. *Eur Psychiatry.* 1996;11(8):390-395.

34. Danion JM, Kauffmann-Muller F, Grange D, Zimmermann MA, Greth P. Affective valence of words, explicit and implicit memory in clinical depression. *J Affect Disord.* Jun 8 1995;34(3):227-234.

35. Brewin CR, Andrews B, Gotlib IH. Psychopathology and early experience: a reappraisal of retrospective reports. *Psychol Bull.* Jan 1993;113(1):82-98.

36. Watkins PC, Mathews A, Williamson DA, Fuller RD. Mood-congruent memory in depression: emotional priming or elaboration? *J Abnorm Psychol.* Aug 1992;101(3):581-586.

37. Airaksinen E, Larsson M, Lundberg I, Forsell Y. Cognitive functions in depressive disorders: evidence from a population-based study. *Psychol Med.* Jan 2004;34(1):83-91.

38. Ilsley JE, Moffoot AP, O'Carroll RE. An analysis of memory dysfunction in major depression. *J Affect Disord.* Oct 9 1995;35(1-2):1-9.

39. Kindermann SS, Brown GG. Depression and memory in the elderly: a meta-analysis. *J Clin Exp Neuropsychol.* Oct 1997;19(5):625-642.

40. Ellwart T, Rinck M, Becker ES. Selective memory and memory deficits in depressed inpatients. *Depress Anxiety.* 2003;17(4):197-206.

41. Calev A. Affect and memory in depression: evidence of better delayed recall of positive than negative affect words. *Psychopathology.* 1996;29(2):71-76.

42. Gonzalez HM, Bowen ME, Fisher GG. Memory decline and depressive symptoms in a nationally representative sample of older adults: the Health and Retirement Study (1998-2004). *Dement Geriatr Cogn Disord.* 2008;25(3):266-271.

43. Livner A, Berger AK, Karlsson S, Backman L. Differential effects of depressive symptoms on prospective and retrospective memory in old age. *J Clin Exp Neuropsychol.* Apr 2008;30(3):272-279.

44. Murray LA, Whitehouse WG, Alloy LB. Mood congruence and depressive deficits in memory: a forced-recall analysis. *Memory.* Mar 1999;7(2):175-196.

45. Raphael KG, Cloitre M. Does mood-congruence or causal search govern recall bias? A test of life event recall. *J Clin Epidemiol.* May 1994;47(5):555-564.

46. Sachs-Ericsson N, Joiner T, Blazer DG. The influence of lifetime depression on self-reported memory and cognitive problems: results from the National Comorbidity Survey-Replication. *Aging Ment Health.* Mar 2008;12(2):183-192.

47. Hermans D, de Decker A, de Peuter S, Raes F, Eelen P, Williams JM. Autobiographical memory specificity and affect regulation: coping with a negative life event. *Depress Anxiety.* 2008;25(9):787-792.

48. Williams J, Stiles W, Shapiro D. Cognitive mechanisms in the avoidance of painful and dangerous thoughts: Elaborating the assimilation model. *Cognitive Therapy Research.* 1999;23:285-306.

49. Williams JM, Broadbent K. Autobiographical memory in suicide attempters. *J Abnorm Psychol.* May 1986;95(2):144-149.

50. Kuyken W, Dalgleish T. Autobiographical memory and depression. *Br J Clin Psychol.* Feb 1995;34 (Pt 1):89-92.

51. Pollock LR, Williams JM. Effective problem solving in suicide attempters depends on specific autobiographical recall. *Suicide Life Threat Behav.* Winter 2001;31(4):386-396.

52. Van Vreeswijk MF, De Wilde EJ. Autobiographical memory specificity, psychopathology, depressed mood and the use of the Autobiographical Memory Test: a meta-analysis. *Behav Res Ther.* Jun 2004;42(6):731-743.

53. Wessel I, Meeren M, Peeters F, Arntz A, Merckelbach H. Correlates of autobiographical memory specificity: the role of depression, anxiety and childhood trauma. *Behav Res Ther.* Apr 2001;39(4):409-421.

54. Williams JM, Barnhofer T, Crane C, et al. Autobiographical memory specificity and emotional disorder. *Psychol Bull.* Jan 2007;133(1):122-148.

55. Williams JM, Scott J. Autobiographical memory in depression. *Psychol Med.* Aug 1988;18(3):689-695.

56. Burt DB, Zembar MJ, Niederehe G. Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychol Bull.* Mar 1995;117(2):285-305.

57. Hunt R. The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review.* 1995;2(1):105-112.

58. Tversky A, Kahneman D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology.* 1973;5:207-232.

59. Dutta S, Kanungo RN. *Affect and Memory: A reformulation*. London: Pergamon; 1975.

60. Matlin M, Stang D. *The Pollyanna Principle: Selectivity in language, memory, and thought*. Cambridge, MA.: :Schenkman; 1978.

61. Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend.* Jul 20 2003;71(1):7-16.

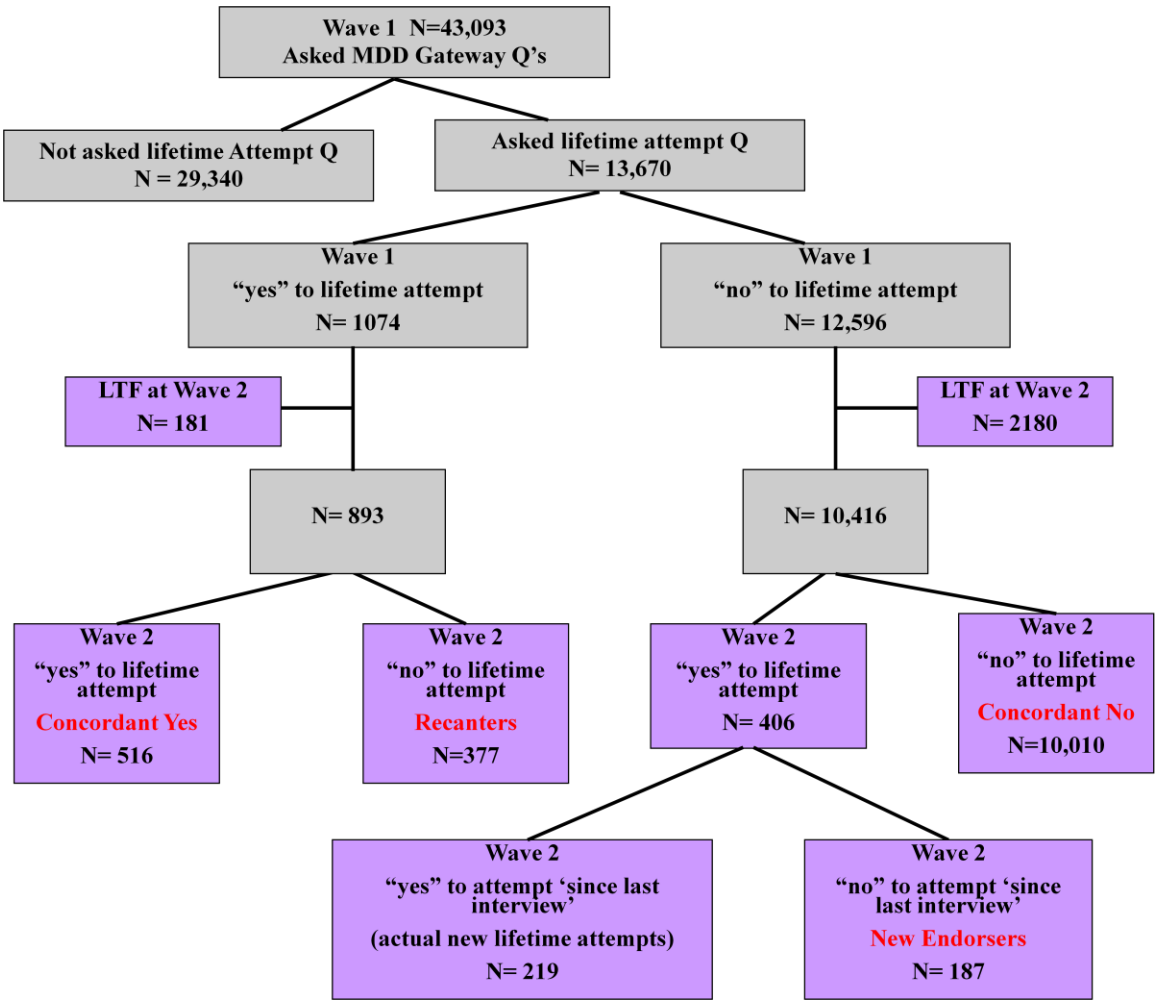
62. Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry.* Apr 2004;61(4):361-368.

63. Incorporated; MOTaQ. 12-Item Short-Form Survey Instrument (SF-12)2002.

64. Khilstrom J, Eich E, Sandbrand D, Tobias B. *Emotion and Memory: Implications for Self-Report: in The Science of Self-Report*. London: Lawrence Erlbaum Associates; 1999.

65. Baca-Garcia E, Perez-Rodriguez MM, Keyes KM, et al. Suicidal ideation and suicide attempts in the United States: 1991-1992 and 2001-2002. *Mol Psychiatry.* Sep 9 2008.

Figure 3A: Flow chart detailing ascertainment of suicide attempts and groups of interest: "Recanters", "Concordant yes", "Concordant no", and "New endorsers" responders.



MDD = Major Depressive Disorder
LTF = Lost to follow-up

MDD Gateway Questions: 1) "In your entire life, have you ever had a time when you felt sad, blue, or down most of the time for at least 2 weeks?"; 2) "In your entire life, have you ever had a time, lasting at least 2 weeks, when you didn't care about the things that you usually cared about, or when you didn't enjoy the things you usually enjoyed?"

Table 3A: Sample characteristics at Wave 2

		Sample of interest (Concordant yes, Concordant No, Recanters, New endorsers) **(n = 11090)** **N (Weighted %)**	**Sample of attempters** (Concordant yes, Recanters, New endorsers) **(n = 1080)** **N (Weighted %)**	**NESARC sample (n =34,653)** **N (Weighted %)**
Demographics				
Age	18-29	1550 (15.4)	214 (21.8)	4913 (15.7)
	30-64	7531 (69.2)	787 (72.9)	22563 (66.3)
	65 +	2009 (15.4)	79 (5.3)	4913 (18.0)
Sex	Female	7377 (62.4)	772 (68.4)	20089 (52.8)
	Male	3713 (37.6)	308 (31.6)	14564 (47.2)
Education	≤ HS	1617 (12.6)	206 (17.2)	5514 (13.5)
	> HS	9473 (87.4)	874 (82.8)	29139 (86.5)
Income	0-34,999	7296 (66.9)∂	847 (82.0) †	21886 (63.0) *
	35 -69	2070 (20.2)	127 (12.0)	6826 (21.4)
	70 +	1267 (12.9)	57 (6.0)	4551 (15.6)
Marital status	Never	2145 (16.8)	254 (20.6)	6638 (17.0)
	Married	5354 (58.9)	432 (50.3)	18866 (64.3)
	Divorced	3591 (24.3)	394 (29.1)	9149 (18.7)
Mental Health				
Lifetime Suicidal ideation	No	5780 (52.7)	41 (5.0)	5780 (48.0) **
	Yes	5270 (47.3)	1039 (95.0)	6409 (52.0)
Current Depressed Mood	No	9868 (90.2)	814 (77.2)	32328 (94.3)
	Yes	1213 (9.8)	265 (22.8)	2226 (5.7)
Lifetime Depressive D/O	No	4459 (39.7)	130 (10.8)	26375 (76.4)
	Yes	6631 (60.3)	950 (89.2)	8278 (23.6)
Lifetime Anxiety disorder	No	5961 (53.3)	315 (28.8)	24786 (71.9)
	Yes	5129 (46.7)	765 (71.2)	9867 (28.1)
Lifetime Substance D/O	No	6366 (54.6)	416 (35.8)	22569 (62.1)
	Yes	4724 (45.4)	664 (64.2)	12084 (37.9)
Lifetime Personality D/O	No	8639 (78.3)	510 (46.6)	29686 (86.3)
	Yes	2451 (21.7)	570 (53.4)	4967 (13.7)

* N= 33,263

** N= 12,189 reduced because at Wave 1, only individuals who endorsed one of the two MDD gateway questions were asked about suicidal ideation

† N = 1,031

∂ N = 10,633

Note: *lifetime depressive disorder* (lifetime MDD or Dysthymia); *lifetime anxiety disorder* (lifetime GAD, Panic Disorder, Social Phobia, Agoraphobia, Specific Phobia, or PTSD); *lifetime Substance Use Disorder* (lifetime alcohol use, alcohol abuse, heroin, inhalant, cocaine, hallucinogen, cannabis, amphetamine, opioid, tranquilizer, or sedative use); *lifetime Personality Disorder* (lifetime Borderline, Antisocial, Schizotypal, or Narcisstic Personality Disorders).

Table 3B: Expected odds of being a) Recanters or b) New endorser (ref=Concordant yes), predicted by mood-congruent and mood-deficit theory, given respondent depression status at *i*) Wave 1 and *ii*) Wave 2.

Recanters vs. Concordant yes	Mood Congruent (AOR)	Mood Deficit (AOR)
Wave 1		
Depressed mood	1.0	1.0
Wave 2		
Depressed mood	< 1.0	> 1.0
New endorsers vs. Concordant yes	**Mood Congruent (AOR)**	**Mood Deficit (AOR)**
Wave 1		
Depressed mood	< 1.0	> 1.0
Wave 2		
Depressed mood	1.0	1.0

Table 3C: Expected odds of being a) Recanter or b) New endorser (ref = any other attempter profile), predicted by mood-congruent and mood-deficit theory, given specific Wave 1: Wave 2 depressed mood profiles

Wave 1: Wave 2 Depressed Mood Profiles Recanters vs. (New endorsers + Concordant yes + Concordant No)	Mood Congruent (AOR)	Mood Deficit (AOR)
D_{10}†	> 1.0	---
D_{01}†	---	> 1.0
$D_{10\ vs.}\ D_{01}$	> 1.0	< 1.0
Wave 1: Wave 2 Depressed Mood Profiles **New endorsers vs.** **(Recanters + Concordant yes + Concordant No)**	**Mood Congruent (AOR)**	**Mood Deficit (AOR)**
D_{10}†	---	> 1.0
D_{01}†	> 1.0	---
$D_{01\ vs.}\ D_{10}$	> 1.0	< 1.0

† (reference = all other mood profiles)

Table 3D: Binomial logistic regression models examining odds of being a) Recanters vs. Concordant yes; and b) New endorsers vs. Concordant yes, given respondent depression status at *i*) Wave 1 and *ii*) Wave 2.

	OR (95% C.I.) Recanters vs. Concordant yes	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
Wave 1			
Depressed mood	0.90 (0.61, 1.32)	0.88 (0.59, 1.32)	0.91 (0.61, 1.37)
Wave 2			
Depressed mood	* 0.59 (0.39, 0.89)	* 0.57 (0.37, 0.87)	* 0.59 (0.39, 0.90)
	OR (95% C.I.) New endorsers vs. Concordant yes	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
Wave 1			
Depressed mood	1.04 (0.66, 1.64)	1.00 (0.63, 1.61)	1.12 (0.69, 1.81)
Wave 2			
Depressed mood	1.48 (0.94, 2.32)	1.45 (0.90, 2.33)	* 1.68 (1.05, 2.67)

AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation

Table 3E: Discordant responders (Recanters, New endorsers) odds of Wave 1:Wave 2 depressed mood profiles.

Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) Recanters** (vs. NE, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	* 2.25 (1.62, 3.12)	* 2.09 (1.45, 3.00)	* 1.62 (1.13, 2.33)
D_{01}†	0.94 (0.57, 1.57)	0.89 (0.53, 1.49)	0.61 (0.36, 1.02)
D_{10} vs. **D_{01}**	* 2.18 (1.23, 3.85)	* 2.15 (1.21. 3.83)	* 2.50 (1.39, 4.49)
Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) New endorsers** (vs. Rec, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	*2.11 (1.22, 3.67)	* 1.94 (1.14, 3.31)	1.59 (0.91, 2.78)
D_{01}†	* 3.75 (2.34, 6.01)	* 3.55 (2.22, 5.65)	* 2.71 (1.71, 4.31)
D_{01} vs. **D_{10}**	1.66 (0.85, 3.26)	1.64 (0.84, 3.23)	1.56 (0.80, 3.06)

† (reference = all other mood profiles)
AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation.

Conclusion

The challenge of capturing suicide attempts in the population, plague its examination. A richer, more nuanced understanding of suicide attempt estimation will allow for greater insight into the accuracy of reported prevalences and correlates; estimates which are vital for resource allocation, as well as intervention and prevention efforts. Further, establishing and characterizing factors that may affect the adequacy of our measures may guide future improvements in the measurement of attempts. The reliability of adult-reported lifetime suicide attempts had not been rigorously explored prior to this work, and therefore estimates have remained largely unchallenged. This dissertation explicitly sought to fill this research gap by utilizing a longitudinal study, comprising two waves of data collection, in which information on suicide attempts was obtained at both time points. In this conclusion, I synthesize findings presented in Chapters 1, 2 and 3, and posit potential public health implications and applications of this research, and suggest future research directions.

Summary of findings

Chapter 1 presented a systematic review of the literature depicting the state of the literature with regards to the reliability of suicidality measures (e.g. ideation, plans, and attempts). Specifically, I identified test-retest reliability estimates, as well as correlates of discordant reporting over time. Using these correlates, identified within, as well as across studies, I then posited potential causal mechanisms underlying discordant reporting over time. Descriptively, I found that the literature primarily concerned the reliability of suicidal ideation, and with the vast majority of studies conducted among youth or young adult populations. Further, few studies assessed correlates of discordant reporting, and no studies examined the reliability of adult-reported suicide attempts. Drawing upon the correlates reported within studies, as well sources of heterogeneity across

studies, I then posited four plausible causal mechanisms underlying discordant suicidality reporting; recall failure, reinterpretation, conscious denial, and lack of construct comprehension. Extending these findings, I proposed that the likelihood of each mechanism is influenced by factors such as the severity of the suicidality, amount of time passed since the suicidal event, social desirability effects, mood context, and suicide construct validity. This literature review found that estimates of suicidality on average, suffer from a moderate degree of unreliability. The degree of discordant reporting however, was dependent on suicide construct, sample, study design, and instrument/wording characteristics. In particular, it highlighted the lack of research concerning the reliability of attempts within adult populations, and deemed it a research priority. Further, it demonstrated the necessity for continued and in-depth assessment of potential correlates of discordant reporting.

In Chapter 2, I addressed research gaps identified in Chapter 1. I assessed the reliability of adult-reported lifetime attempts as reported in a large, population-based longitudinal study, and found reports to be moderately reliable, with a Kappa coefficient of 0.51. I hypothesized that discordant reporters would be more similar to individuals who reported a past attempt at both waves (Concordant yes responders), compared with individuals who reported no attempt at both waves (Concordant no responders). I found that indeed, discordant reporters were more similar to the former, potentially signifying that discordant reporters are true attempters who underreported their attempt at one time point. Further, I hypothesized that discordant reporters would be less likely to have a history of depressive disorders compared with Concordant yes responders; positing that this history would serve as a marker for attempt severity[1-6], and that discordants would have less severe attempts, which would therefore be more easily forgotten or

reinterpreted.[7-9] Contrary to this hypothesis however, discordants were as likely as Concordant yes individuals to have a history of depressive disorders, and unexpectedly, discordants were much less likely to have a history of suicidal ideation. It is therefore plausible that a history of suicidal ideation serves as a marker for attempt severity, and/or that discordant reporters are characterized by more impulsive attempts. I concluded by suggesting further examination of the detected association between suicidal ideation and discrepant suicide attempt reporting, and by calling for further exploration of potential correlates, particularly the role of a respondent's current mood at the time of reporting.

Finally, in Chapter 3, I built upon findings from Chapter 2 by examining how a respondent's current depressed mood may influence the recall, and hence reporting of attempts. Based on established mood-recall theories, I tested three competing hypotheses to determine if a current depressed mood would enhance (mood-congruent recall),[10,11]inhibit (mood memory deficit effect),[12-18] or have no effect on the recall (mood-independent recall)[19-22] and reporting of attempts. I hypothesized that discordant reporters would demonstrate a mood-congruent pattern of reporting, such that a depressed mood at the time of the interview would increase the likelihood that a respondent would report an attempt at that wave. There were in fact, distinct mood-congruent reporting effects among Recanters, yet mood-independent effects detected among New endorsers. This may indicate that New endorsers are a unique group of discordant responders, which warrant further examination. Still, because respondents in our sample were over 20 times more likely to recant than newly endorse, and comparatively, there was limited power within our New endorser group, I believe these results may be generalized to assert that overall, discordant responders report in a mood-congruent fashion.

Public health implications and applications of findings

This research agenda was innovative in that it was the first to address a significant and prominent gap in the suicidality literature; the reliability of adult-reported lifetime suicide attempts. These findings not only highlight the compromised reliability of these estimates, but also the validity of attempt reports, and accordingly, the accuracy of our estimates. I believe this research to have several potential public health implications and applications, which include the accuracy of prevalence and predictor estimates, the representativeness of attempter samples, the modification and improvement of suicide attempt measures, and a broader application of findings within the field of psychiatric epidemiology.

The finding that discordant reporters were more similar to Concordant yes, than to Concordant no individuals, likely signifies that respondents are underreporting their attempts to some extent. This underreporting results in underestimates of the prevalence of suicide attempts in the population, on which many public health actions, such as surveillance, resource allocation, and priority-setting are based. For example, as observed, 2.1% of respondents reported an attempt at Wave 1. If those that newly endorsed at Wave 2 were true attempters at Wave 1, then the true prevalence at Wave 1 would be 2.5%. Likewise, as observed, 2.1% of respondents reported an attempt at Wave 2. If those that recanted at Wave 2 were true attempters, then the true prevalence at Wave 2 would be 3.1%. In the case of the latter, an underestimate of approximately 1%, translates into 310,000 attempts in the U.S. not represented by the sample obtained. Therefore, if this underreporting of attempts is consistent over time, trend surveillance should be unaffected; however the determined prevalence, which serves as the basis for resource allocation would certainly be underestimated.

The underreporting of attempts may also have implications for assessed predictors of attempts. This research found that males were more likely to recant than females. If males are less likely to report their suicide attempts, then the commonly reported increased odds of attempts among women[23-25]is potentially an overestimate. In this sample, as observed at Wave 2, females are 52% (OR, 1.52; 95% CI, 1.22-1.85) more likely than males to make an attempt. However, if we assume that individuals that recanted at Wave 2 are true attempters, then females are 33% more likely (OR, 1.33; 95% CI, 1.12-1.59) than males to make an attempt. Therefore, disproportionate underreporting by factors that are also associated with suicide attempts, may bias estimates between assessed predictors and attempts. If future research supports the degree of underreporting and correlates of discordant reporting found in this dissertation, it is plausible that both prevalence and predictor estimates may be adjusted to reflect the likely true estimates.

If discordant reporters are indeed characterized by more impulsive attempts, as potentially indicated by their comparatively reduced likelihood of past suicidal ideation, impulsive attempts might be underrepresented in population estimates. Discordant responders were much less likely to have had past ideation compared with Concordant yes individuals (OR, 0.13; 95% CI, 0.05, 0.37), highlighting the high degree of potential underreporting of impulsive attempts; potentially reducing the representativeness of the attempter sample, and the generalizability of predictor findings. In addition, while relatively less is understood about impulsive attempts, some studies have shown that impulsive attempts may differ from premeditated attempts in regards to various mental health characteristics and attempt severity.[6,26-29] A disproportionate underreporting of impulsive attempts may obscure other important etiologic differences between premeditated and

impulsive attempts, such as the influence of recent stressful life events and substance use, as well as impulsive personality traits.

Several likely correlates of discordant attempt reporting uncovered throughout this dissertation, may inform suicide attempt question modification. For example, if confusion surrounding the construct of attempts potentially influences reporting, questions may benefit from more direct and detailed wording, such as "have you ever tried to kill yourself, in which there was at least some intention to die". Likewise, if attempt severity affects reporting, questions regarding attempts could be followed by a short clarifying statement indicating that "suicide attempts", includes ones that resulted in medical care, as well as ones that did not. Also, if impulsive attempters are less likely to report their attempts, it's plausible that they may be ambivalent about whether there was true intention to die, given that it was done impulsively. If so, questions could be modified to elicit both premeditated, as well as impulsive attempts. Lastly, the finding that respondents were more likely to report their attempt if they were currently depressed may have public health implications. Surveys that include suicide attempt measures may have reason to also include measures of very recent and/or momentary mood. Momentary mood measures could provide researchers with more information on which to base sensitivity analyses, assessing how underreporting may affect their causal associations of interest. This dissertation found that approximately 1% of respondents are likely underreporting attempts, and respondents are almost two times more likely to underreport when they are non-depressed. By reclassifying varying percentages of non-reporters as attempters, based on their mood at the time of the interview, researchers can then model how varying degrees of underreporting may affect their associations of interest.

There may be broader applications of these findings to the field of psychiatric epidemiology as most measures of mental health have well-documented, and varying degrees of unreliability.[30-32] Examination of the degree, but most importantly potential correlates of discordant reporting within the suicide literature may add to the body of evidence examining recanting among other psychiatric disorders and behaviors. For example, male, young, and less educated respondents have been shown to inconsistently report past substance,[33-35] indicating that there may be similar causal mechanisms underlying discordant reporting across behaviors. In particular, respondent current mood as a correlate of reporting may have a broader application. A substantive area of psychiatric epidemiology has been devoted to examining how negative life events impact the likelihood of depression in individuals. It would be prudent for researchers to consider how estimates from retrospective studies may be biased if individuals report events in a mood-congruent manner. That is, if respondents are less likely to report negative life events, such as a suicide attempt, or abuse when they are not depressed, then the underreporting of these events would result in differential misclassification, and a bias away from the null; potentially overestimating the effects between negative life events and depression.

Finally, there may be some indirect implications of these findings. This research, via potentially improved measurement of attempts in the future, may not only aid in the eventual prevention of attempts. It is estimated that 6-10 percent of individuals who attempt suicide, complete within 5 years,[36-38] and 40 percent of individuals who complete, have attempted at least once before.[39] An enhanced understanding of suicide attempts will therefore not only inform prevention and intervention efforts surrounding attempts, but ultimately, completions as well.

Strengths and limitations

This dissertation had a number of limitations, related to external validity, loss to follow-up, and power. The main limitation of Chapter 1, the literature review, was the limited studies available reporting the test-retest reliability of suicidality estimates. Specifically, the wide variations across a small number of studies made it challenging to effectively isolate which characteristics were truly influencing the reliability estimate. Studies were also heterogeneous in regards to reliability statistics used. Since the majority of studies used the Kappa or conditional Kappa statistic, which accounts for chance agreement, the review tended to rely more on Kappa coefficients, and therefore those studies, to examine correlates across studies. While other reliability statistics were strongly considered during correlate assessment, the slight priority given to kappa coefficients may have reduced the generalizability of the findings.

The reliability of attempt reports was only assessed among individuals who passed the MDD gateway questions at Wave 1, thereby excluding 29,340 individuals from our analysis. Unfortunately, due to the survey structure at Wave 2, we are unable to assess how many attempts were not captured due to the gateway questions at Wave 1. However, the NESARC's predecessor survey, the National Longitudinal Alcohol Epidemiologic Survey (NLAES), found that the number of individuals who did not screen into the depression section and reported a suicide attempt was very low (< 0.1% of the sample).[40] Therefore, I do not believe the gateway questions affected our external validity substantially. Unlike at Wave 1, the attempt question at Wave 2 was not preceded by gateway questions, which could also influence our findings. The fact that respondents were not "primed" by the depression questions at Wave 2 may explain the high percentage of Recanters, and the comparatively low percentage of New endorsers. This

lack of priming may have diminished recall or influenced reinterpretation among the respondents. Additionally, this could explain some of the observed differences between Recanters and New endorsers.

While this dataset was rich in respondent sociodemographic and mental health information, many likely important characteristics of suicide attempts were not captured in the NESARC, but could certainly influence reporting reliability. The medical and psychological severity of respondent attempts would have been particularly useful for these analyses. Since information on severity was not available, I used past MDD diagnosis as a proxy for attempt severity; however, the appropriateness of this proxy can be debated, and an actual measure of severity would have been preferred. In addition, there was no way to assess the potential influence of social desirability on the reporting of attempts, and inclusion of such a measure would have been useful in the detection of this possible underlying causal mechanism. The amount of time that had passed between the suicide attempt and interview was only available for New endorsers and Concordant yes individuals; having this information on Recanters as well could have provided a strong test of recall failure as a potential underlying mechanism, and may have helped further tease apart differences between Recanters and New endorsers.

True differences between Recanters and New endorsers were also difficult to conclude given the comparatively smaller sample size of New endorsers. Given that the Recanter group was approximately twice as large as the New endorser group, it is possible that differences in current mood-reporting patterns may have arisen from reduced power within the New endorser group. However, given that mood-congruent findings were detected within the New endorser group

when current mood was measured in alternative ways, we believe that our mood-congruent conclusions were unaffected by this potentially limited power.

Lastly, as with all prospective studies, this study had loss to follow up. Seventeen percent of our sample was lost to follow-up before Wave 2. Individuals lost to follow up were similar to the final sample on most major baseline predictors, but were i) slightly less likely to have a history of depressive or anxiety disorders; ii) less likely to have had more than a high school education; iii) more likely to have had a current depressed mood at Wave 1. Since there is no way of assessing loss to follow-up (LTF) up differential by attempter profile, it is difficult to determine the affect of this LTF. However, if those lost to follow up were also more likely to be discordant reporters, it could perhaps partially explain our null finding between discordant status and history of depressive disorders found in Chapter 2. Further, if Recanters were more likely to be LTF, then our mood-congruent findings found in Chapter 3 would be an underestimate of the true strength of the theory; if New endorsers were more likely to be LTF, our findings would represent an overestimate of the mood-congruent findings. However, there is no reason to believe that Recanters were LTF at a rate different from New endorsers, thereby likely exhibiting minimal influence on our findings.

The main strength of this study lies within the novelty of the research questions. This is the first study to report the reliability of adult-reported suicide attempts. All prior studies examining the reliability of attempt reporting were conducted among youth or young adult populations, with limited potential predictors and small, non population-based samples. The NESARC is a rich and detailed dataset, allowing for a greater breadth of potential predictors and its large sample size

provided the power to examine predictors in adjusted models, rather than bivariate analyses alone. Further, this study was the first to assess how a respondent's current mood may affect the reporting of attempts among adults, and did so using competing hypotheses, a powerful scientific technique. Lastly, as a way of assessing the robustness of our mood-congruent findings in Chapter 3, we classified current depressed mood using three alternate measures; findings which not only supported our conclusions, but also strengthened them. We believe the consistent mood-congruent results given varying measures of current mood, lends robustness to our findings.

Future Directions

I believe this research to be only the first, yet requisite step towards the improvement of suicide attempt measures, and consequently, more reliable and valid estimates. I would encourage researchers in the field to identify both existing and new data source opportunities that may be utilized for further exploration of correlates of discordant reporting, and underlying causal mechanisms. Correlates of particular interest, based on both theory and findings from this dissertation, should include the severity of the attempt as well as the presence of prior suicidal ideation.

In addition, the underlying reasons as to why these correlates lead to discordant reporting warrants a more in-depth investigation than this dissertation was able to provide. This dissertation primarily focused on recall failure as a likely candidate for reporting impact. However, it is also critical to investigate the role of reinterpretation, as this explanation is also consistent with our mood-congruent findings. Future research should include qualitative aspects in attempt to tease apart discordant responses resulting from recall failure and reinterpretation.

Social desirability effects were also not able to be assessed in this dissertation work, but likely have some impact on the reporting of attempts. Social desirability effects may be specific to the respondent, and captured through the inclusion of social desirability scales. However, the degree of social desirability effects may also depend on characteristics of the interviewer, which likely change across waves of data collection, and therefore candidates for explaining inconsistent responses over time. For example, in this study, males were more likely to recant attempts than females. If males were to underreport attempts due to effects of social desirability for fear that attempts may be seen as weak or vulnerable, it is plausible that the gender of the interviewer may affect male reporting; males may be less likely to report a past attempt to a male interviewer, compared with a female interviewer. Once again, these types of social desirability influences could be examined with more detailed information on the interviewer, in future studies.

The mood-congruent findings found in Chapter 3 should be explored for extended applicability and robustness. Mood-congruent reporting likely influences the reporting of other suicidal constructs, such as ideation and plans; however this assumption needs to be examined. Further, the robustness of this theory should be explored with other types of moods. For example, because some evidence exists for the association between anxiety disorders and attempts,[41] anxious or impulsive moods may also affect the reporting of attempts in a mood-congruent manner. Similarly, perhaps respondents who are currently using substances, either alcohol, drugs, or smoking, may be more likely to accurately report past substance use, compared with those who are currently abstinent or have low substance use levels.

The absence of mood-congruent findings among New endorsers, as measured with current depressed mood, certainly warrants more focused attention. While mood-congruent patterns were indeed found within New endorsers when depressed mood was measured by "past year MDD and Dysthymia", they were not captured when current depressed mood, our main mood variable, was utilized. There are several plausible explanations for these slightly disparate findings. First, it is possible that New endorsers may have been misclassified. In order to be correctly classified as a New endorser, respondents had to accurately determine if their attempt was "since the last interview" or "before the last interview". Therefore, some New endorsers may have actually attempted soon after the first interview (hence, an incident attempt), but misremembered it as before the last interview (misclassifying them as a New endorser). However, contrary to this supposition, in Chapter 2, we found that New endorsers were younger at their first and last attempts, were more likely to have attempted more than 20 years ago, and have attempted only once, compared with Concordant yes individuals. These findings would therefore greatly challenge the plausibility that many New endorsers were in fact incident attempters. Secondly, it is highly possible that the New endorser group (n=187), which was half the size of the Recanter group (n=377), had comparatively less power to detect significant mood-congruent findings for current mood. Lastly, it is also possible that New endorsers and Recanters are in fact, distinct types of discordant reporters. Individuals who forget or reinterpret-down over time (i.e. explain away attempts over time), which would be the expected direction, may differ from individuals who remember, or reinterpret-up over time. These differences would require a more detailed analysis, perhaps inclusive of qualitative information.

Finally, given that hypothesized correlates of discordant reporting are corroborated in further research, investigators may choose to adapt suicide attempt questions in attempt to combat underreporting, and potentially capture a wider, more representative range of attempts and attempters. Ideally, suicide attempt questions should capture the full range of medical severity, intent to die, impulsivity, and recency of attempts. A randomized experiment examining different question phrasing may help illuminate any differences in reporting they may elicit. Of course, given the rarity of suicide attempts, this would be most feasibly done within a high-risk sample.

Conclusion

My hope is that this dissertation work will serve as just the beginning of a longer line of research investigating the adequacy of our suicide attempt measures and estimates. By highlighting the ways in which our estimates may be flawed, my aim was to motivate, rather than dishearten those who use attempt estimates in their research work. Continual enhancement of attempt measures will allow for better estimation of predictors and prevalence, with the ultimate goal of reducing future suicide attempts.

References

1. Astruc B, Torres S, Jollant F, et al. A history of major depressive disorder influences intent to die in violent suicide attempters. *J Clin Psychiatry.* May 2004;65(5):690-695.

2. Brown GK, Henriques GR, Sosdjan D, Beck AT. Suicide intent and accurate expectations of lethality: predictors of medical lethality of suicide attempts. *J Consult Clin Psychol.* Dec 2004;72(6):1170-1174.

3. Dhossche DM, Meloukheia AM, Chakravorty S. The association of suicide attempts and comorbid depression and substance abuse in psychiatric consultation patients. *Gen Hosp Psychiatry.* Jul-Aug 2000;22(4):281-288.

4. Hasley JP, Ghosh B, Huggins J, Bell MR, Adler LE, Shroyer AL. A review of "suicidal intent" within the existing suicide literature. *Suicide Life Threat Behav.* Oct 2008;38(5):576-591.

5. Kumar CT, Mohan R, Ranjith G, Chandrasekaran R. Characteristics of high intent suicide attempters admitted to a general hospital. *J Affect Disord.* Mar 2006;91(1):77-81.

6. Suominen K, Isometsa E, Henriksson M, Ostamo A, Lonnqvist J. Hopelessness, impulsiveness and intent among suicide attempters with major depression, alcohol dependence, or both. *Acta Psychiatr Scand.* Aug 1997;96(2):142-149.

7. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. *J Health Soc Behav.* Sep 1987;28(3):232-246.

8. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T. The accuracy of prevalence estimations for suicide attempts. how reliably do adolescents and young adults report their suicide attempts? *Arch Suicide Res.* 2006;10(3):253-263.

9. Simon GE, VonKorff M. Recall of psychiatric history in cross-sectional surveys: implications for epidemiologic research. *Epidemiol Rev.* 1995;17(1):221-227.

10. Blaney PH. Affect and memory: a review. *Psychol Bull.* Mar 1986;99(2):229-246.

11. Bower GH. Mood and memory. *Am Psychol.* Feb 1981;36(2):129-148.

12. Airaksinen E, Larsson M, Lundberg I, Forsell Y. Cognitive functions in depressive disorders: evidence from a population-based study. *Psychol Med.* Jan 2004;34(1):83-91.

13. Calev A. Affect and memory in depression: evidence of better delayed recall of positive than negative affect words. *Psychopathology.* 1996;29(2):71-76.

14. Ellwart T, Rinck M, Becker ES. Selective memory and memory deficits in depressed inpatients. *Depress Anxiety.* 2003;17(4):197-206.

15. Ilsley JE, Moffoot AP, O'Carroll RE. An analysis of memory dysfunction in major depression. *J Affect Disord.* Oct 9 1995;35(1-2):1-9.

16. Kindermann SS, Brown GG. Depression and memory in the elderly: a meta-analysis. *J Clin Exp Neuropsychol.* Oct 1997;19(5):625-642.

17. Murray LA, Whitehouse WG, Alloy LB. Mood congruence and depressive deficits in memory: a forced-recall analysis. *Memory.* Mar 1999;7(2):175-196.

18. Raphael KG, Cloitre M. Does mood-congruence or causal search govern recall bias? A test of life event recall. *J Clin Epidemiol.* May 1994;47(5):555-564.

19. Dutta S, Kanungo RN. *Affect and Memory: A reformulation*. London: Pergamon; 1975.

20. Hunt R. The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review.* 1995;2(1):105-112.

21. Matlin M, Stang D. *The Pollyanna Principle: Selectivity in language, memory, and thought*. Cambridge, MA.: :Schenkman; 1978.

22. Tversky A, Kahneman D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology.* 1973;5:207-232.

23. Borges G, Angst J, Nock MK, Ruscio AM, Walters EE, Kessler RC. A risk index for 12-month suicide attempts in the National Comorbidity Survey Replication (NCS-R). *Psychol Med.* Dec 2006;36(12):1747-1757.

24. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry.* Jul 1999;56(7):617-626.

25. Kuo WH, Gallo JJ, Tien AY. Incidence of suicide ideation and attempts in adults: the 13-year follow-up of a community sample in Baltimore, Maryland. *Psychol Med.* Oct 2001;31(7):1181-1191.

26. Baca-Garcia E, Diaz-Sastre C, Basurte E, et al. A prospective study of the paradoxical relationship between impulsivity and lethality of suicide attempts. *J Clin Psychiatry.* Jul 2001;62(7):560-564.

27. Baca-Garcia E, Diaz-Sastre C, Garcia Resa E, et al. Suicide attempts and impulsivity. *Eur Arch Psychiatry Clin Neurosci.* Apr 2005;255(2):152-156.

28. Brent DA. Correlates of the medical lethality of suicide attempts in children and adolescents. *J Am Acad Child Adolesc Psychiatry.* Jan 1987;26(1):87-91.

29. Groholt B, Ekeberg O, Haldorsen T. Adolescents hospitalised with deliberate self-harm: the significance of an intention to die. *Eur Child Adolesc Psychiatry.* Dec 2000;9(4):244-254.

30. Andreasen NC, Grove WM, Shapiro RW, Keller MB, Hirschfeld RM, McDonald-Scott P. Reliability of lifetime diagnosis. A multicenter collaborative perspective. *Arch Gen Psychiatry.* Apr 1981;38(4):400-405.

31. Bromet EJ, Dunn LO, Connell MM, Dew MA, Schulberg HC. Long-term reliability of diagnosing lifetime major depression in a community sample. *Arch Gen Psychiatry.* May 1986;43(5):435-440.

32. Mazure C, Gershon ES. Blindness and reliability in lifetime psychiatric diagnosis. *Arch Gen Psychiatry*. May 1979;36(5):521-525.

33. Fendrich M, Rosenbaum DP. Recanting of substance use reports in a longitudinal prevention study. *Drug Alcohol Depend.* Jun 5 2003;70(3):241-253.

34. Percy A, McAlister S, Higgins K, McCrystal P, Thornton M. Response consistency in young adolescents' drug use self-reports: a recanting rate analysis. *Addiction.* Feb 2005;100(2):189-196.

35. Rehm J, Irving H, Ye Y, Kerr WC, Bond J, Greenfield TK. Are lifetime abstainers the best control group in alcohol epidemiology? On the stability and validity of reported lifetime abstention. *Am J Epidemiol.* Oct 15 2008;168(8):866-871.

36. Cooper J, Kapur N, Webb R, et al. Suicide after deliberate self-harm: a 4-year cohort study. *Am J Psychiatry*. Feb 2005;162(2):297-303.

37. Johnsson Fridell E, Ojehagen A, Traskman-Bendz L. A 5-year follow-up study of suicide attempts. *Acta Psychiatr Scand.* Mar 1996;93(3):151-157.

38. Nordstrom P, Samuelsson M, Asberg M. Survival analysis of suicide risk after attempted suicide. *Acta Psychiatr Scand.* May 1995;91(5):336-340.

39. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. *Psychol Med.* Apr 2003;33(3):395-405.

40. Baca-Garcia E, Perez-Rodriguez MM, Keyes KM, et al. Suicidal ideation and suicide attempts in the United States: 1991-1992 and 2001-2002. *Mol Psychiatry.* Sep 9 2008.

41. Hawgood J, De Leo D. Anxiety disorders and suicidal behaviour: an update. *Curr Opin Psychiatry*. Jan 2008;21(1):51-64.

Appendices

Appendix 2A: MDD gateway question frequencies

MDD gateway questions		Total %	Concordant Yes %	Concordant No %	Recanter %	New Endorser %
In your entire life, have you ever had a time when you felt sad, blue, or down most of the time for at least 2 weeks?	**No**	7.1	3.9	7.4	5.7	4.0
	Yes	92.9	96.1	92.6	94.3	96.0
In your entire life, have you ever had a time, lasting at least 2 weeks, when you didn't care about the things that you usually cared about, or when you didn't enjoy the things you usually enjoyed?	**No**	22.2	4.7	23.7	7.9	16.5
	Yes	77.8	95.3	76.3	92.1	83.5

Appendix 2B: Composite measure of Lifetime Suicidal Ideation at Wave 2. If the respondent endorsed any lifetime ideation item at Wave 1, or any "since last interview" ideation item at Wave 2, they were classified as having Lifetime Suicidal Ideation at Wave 2.

Suicidal Ideation Questions		Total %	Concordant Yes %	Concordant No %	Recanter %	New Endorser %
Wave 1 – Lifetime measures						
Did you ever think about committing suicide?	No	73.6	3.8	80.0	6.6	42.3
	Yes	26.4	96.2	20.0	93.4	57.7
Did you ever feel like you wanted to die?	No	68.0	3.6	73.9	7.4	35.8
	Yes	32.0	96.4	26.1	92.6	64.2
Did you ever think a lot about your own death?	No	74.9	26.3	79.1	30.1	63.3
	Yes	25.1	73.7	20.9	69.9	36.7
Any Wave 1 Ideation	No	57.5	0.8	62.6	4.6	29.5
	Yes	42.5	99.2	37.4	95.4	70.5
Wave 2- Since last interview measures						
Did you ever think about committing suicide?	No	92.7	64.2	94.8	87.6	62.0
	Yes	7.3	35.8	5.2	12.4	38.0
Did you ever feel like you wanted to die?	No	89.2	61.8	91.5	84.1	44.6
	Yes	10.8	38.2	8.5	15.9	55.4
Did you ever think a lot about your own death?	No	90.7	68.6	92.3	87.5	68.5
	Yes	9.3	31.4	7.7	12.5	31.5
Any Ideation Since Last Interview	No	84.9	53.5	87.4	80.1	37.6
	Yes	15.1	46.5	12.6	19.9	62.4
Either Wave 1 or Wave 2 ideation	No	52.8	0.8	57.7	4.6	17.2
	Yes	47.2	99.2	42.3	95.4	82.8

Appendix 2C: Prevalence of Wave 2 lifetime mental health disorders among samples of interest

		Total %	Concordant Yes %	Concordant No %	Recanter %	New Endorser %
Depressive disorder	Yes	60.3	91.5	57.3	86.9	87.5
MDD	Yes	58.4	89.2	55.5	83.3	87.2
Dysthymia	Yes	14.5	43.3	12.3	28.7	30.0
Anxiety disorder	Yes	46.7	75.3	44.2	60.5	81.9
GAD	Yes	17.5	39.3	15.7	23.5	48.7
Panic disorder	Yes	4.4	16.6	3.3	9.3	18.4
Social Phobia	Yes	14.2	33.6	12.6	19.3	40.0
Agoraphobia	Yes	0.7	2.9	0.6	0.6	1.1
Specific Phobia	Yes	25.1	45.9	23.3	35.0	47.3
PTSD	Yes	17.8	42.2	15.7	25.5	54.2
Substance use disorder	Yes	45.4	70.6	43.5	58.5	58.8
Alcohol abuse	Yes	19.7	19.9	19.9	16.6	14.0
dependence	Yes	22.7	41.9	21.0	36.4	37.7
Heroin abuse	Yes	0.2	0.4	0.2	0.1	0.0
dependence	Yes	0.3	1.1	0.2	0.0	0.5
Inhalant abuse	Yes	0.6	1.3	0.5	0.8	2.1
dependence	Yes	0.1	0.7	0.0	0.2	0.0
Cocaine abuse	Yes	2.4	5.8	2.7	6.3	5.1
dependence	Yes	3.0	9.6	1.9	6.6	6.0
Hallucinogen abuse	Yes	2.7	8.2	2.3	5.0	5.2
dependence	Yes	0.7	2.4	0.6	0.8	0.5
Cannabis abuse	Yes	10.9	18.9	10.1	17.1	20.8
dependence	Yes	3.4	11.6	2.9	5.0	7.4
Amphetamine abuse	Yes	2.6	6.0	2.4	3.8	6.4
dependence	Yes	1.4	4.5	1.1	3.0	4.1
Opioid abuse	Yes	1.2	8.5	2.1	4.6	4.0
dependence	Yes	2.5	5.9	0.8	5.2	2.3
Tranquilizer abuse	Yes	1.7	4.7	1.5	3.3	3.0
dependence	Yes	0.8	4.7	0.5	3.5	2.0
Sedative abuse	Yes	1.9	6.4	1.6	2.6	5.1
dependence	Yes	0.8	3.9	0.6	1.8	0.5
Other drug abuse	Yes	0.20	0.1	0.2	0.8	0.0
dependence	Yes	0.10	1.3	0.0	0.8	0.0
Personality disorder	Yes	21.7	65.8	18.5	34.2	58.8
Borderline	Yes	11.1	50.1	8.3	17.6	48.5
Antisocial	Yes	7.2	25.7	6.0	15.8	11.7
Schizotypal	Yes	7.2	26.7	5.8	9.7	25.5
Narcisstic	Yes	8.4	18.5	7.5	12.5	21.1

Appendix 2D: Weighted percentage frequencies of Wave 2 lifetime mental health disorders crosstabs

		Dep %		Anx %		Sub %		Pers %		Idea %	
		Yes	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
Dep %	**Yes**										
	No										
Anx %	**Yes**	13.8	14.3								
	No	9.9	62.0								
Sub %	**Yes**	11.7	26.2	13.1	24.8						
	No	11.9	50.2	15.0	47.1						
Pers %	**Yes**	6.9	6.8	7.6	6.1	8.6	5.1				
	No	16.8	69.5	20.5	65.8	29.3	57.0				
Idea %	**Yes**	38.8	9.3	28.6	19.5	25.5	22.6	15.7	32.5		
	No	21.9	30.0	18.4	33.5	20.2	31.6	6.4	45.4		

Appendix 2E: Odds of mental health comorbity

	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
	Dep	**Anx**	**Sub**	**Pers**	**Idea**
Dep	1				
Anx	6.02 (5.63, 6.44)	1			
Sub	1.90 (1.77, 2.03)	1.67 (1.57, 1.77)	1		
Pers	4.17 (3.84, 4.52)	4.05 (3.74, 4.40)	3.25 (3.01, 3.51)	1	
Idea	5.69 (5.10, 6.35)	2.67 (2.44, 2.92)	1.76 (1.61, 1.93)	3.42 (3.05, 3.84)	1

Appendix 2F: Reliability across levels of predictor variables

		Kappa
W2 Lifetime depressive disorder	Yes	0.52
	No	0.39
MDD	Yes	0.52
	No	0.41
Dysthymia	Yes	0.59
	No	0.46
W2 Lifetime anxiety disorder	Yes	0.53
	No	0.46
GAD	Yes	0.55

	No	0.49
Panic Disorder	Yes	0.53
	No	0.51
Social Phobia	Yes	0.56
	No	0.49
Agoraphobia	Yes	0.78
	No	0.51
Specific Phobia	Yes	0.55
	No	0.48
PTSD	Yes	0.51
	No	0.50
W2 Lifetime personality disorder	Yes	0.58
	No	0.40
Borderline Personality	Yes	0.56
	No	0.44
Antisocial Personality	Yes	0.63
	No	0.48
Schizotypal Personality	Yes	0.60
	No	0.48
Narcisstic Personality	Yes	0.56
	No	0.51
W2 Lifetime Substance use disorder	Yes	0.54
	No	0.49
Alcohol Use Disorder	None	0.48
	abuse	0.58
	dependence	0.51
Heroin Use Disorder	None	0.52
	abuse	0.54
	dependence	0.63
Inhalant Use Disorder	None	0.52
	abuse	0.40
	dependence	0.60
Cocaine Use Disorder	None	0.51
	abuse	0.46
	dependence	0.53
Hallucinogen Use Disorder	None	0.51
	abuse	0.60
	dependence	0.63
Cannabis Use Disorder	None	0.50
	abuse	0.50
	dependence	0.64
Amphetamine Use Disorder	None	0.51
	abuse	0.53
	dependence	0.49
Opioid Use Disorder	None	0.51
	abuse	0.55
	dependence	0.38
Tranquilizer Use Disorder	None	0.51
	abuse	0.46
	dependence	0.47

Sedative Use Disorder	None	0.51
	abuse	0.54
	dependence	0.73
Other drug disorder	None	0.52
	abuse	0.16
	dependence	0.37

Appendix 2G: Multinomial logistic regression models examining bivariate odds of being a) Discordant vs. a Concordant No individual; b) Discordant vs. a Concordant Yes individual; c) New endorser vs. a Recanter; d) Recanter vs. Concordant Yes; and e) New endorser vs. Concordant Yes

	OR (95% C.I.) Discordant vs. Concordant No	**OR (95% C.I.) Discordant vs. Concordant Yes**	**OR (95% C.I.) New endorser vs. Recanter**	**OR (95% C.I.) Recanter vs. Concordant Yes**	**OR (95% C.I.) New endorser vs. Concordant Yes**
Mental Health (W2 lifetime)					
Depressive disorder	*5.02 (3.68, 6.86)	0.63 (0.39, 1.01)	1.05 (0.58, 1.91)	0.62 (0.37, 1.03)	0.65 (0.35, 1.22)
MDD	*4.38 (3.25, 5.89)	0.66 (0.43, 1.03)	1.37 (0.75, 2.48)	*0.60 (0.37, 0.97)	0.82 (0.45, 1.51)
Dysthymia	*2.94 (2.33, 3.71)	*0.54 (0.39, 0.75)	1.06 (0.65, 1.74)	*0.53 (0.37, 0.76)	*0.56 (0.36, 0.89)
Anxiety disorder	*2.63 (2.12, 3.25)	*0.68 (0.49, 0.94)	*2.96 (1.78, 4.93)	*0.50 (0.36, 0.70)	1.48 (0.90, 2.46)
GAD	*2.50 (1.96, 3.18)	*0.72 (0.53, 0.97)	*3.09 (1.96, 4.86)	*0.47 (0.33, 0.68)	1.47 (0.99, 2.17)
Panic disorder	*4.07 (2.83, 5.86)	0.71 (0.45, 1.11)	*2.19, 1.01, 4.73)	*0.52 (0.29, 0.91)	1.13 (0.59, 2.16)
Social Phobia	*2.45 (1.85, 3.25)	*0.70 (0.49, 0.99)	*2.79 (1.77, 4.41)	*0.47 (0.31, 0.71)	1.32 (0.85, 2.04)
Agoraphobia	1.28 (0.35, 4.62)	0.26 (0.06, 1.10)	1.75 (0.14, 21.43)	0.21 (0.04, 1.13)	0.36 (0.04, 3.11)
Specific Phobia	*2.11 (1.69, 2.62)	0.75 (0.56, 1.01)	*1.67 (1.05, 2.66)	*0.63 (0.45, 0.89)	1.06 (0.70, 1.59)
PTSD	*2.89 (2.26, 3.69)	0.73 (0.53, 1.01)	*3.45 (2.11, 5.64)	*0.47 (0.32, 0.68)	*1.62 (1.01, 2.58)
Substance use disorder	*2.47 (1.98, 3.09)	*0.62 (0.45, 0.85)	1.10 (0.65, 1.86)	*0.60 (0.43, 0.84)	0.66 (0.40, 1.08)
Alcohol use none	---	---	---	---	---
abuse	0.98 (0.74, 1.30)	*0.63 (0.41, 0.97)	0.82 (0.44, 1.53)	0.67 (0.42, 1.08)	0.55 (0.31, 1.00)
dependence	*2.18 (1.70, 2.79)	0.70 (0.49, 1.02)	1.01 (0.58, 1.77)	0.70 (0.47, 1.05)	0.71 (0.41, 1.22)
Heroin use none	---	---	---	---	---
abuse	0.00 (0.00, 0.00)¥	0.00 (0.00, 0.00)¥	1.26 (0.97, 1.65)	0.00 (0.00, 0.00)¥	0.00 (0.00, 0.00)¥
dependence	0.95 (0.19, 4.77)	*0.20 (0.04, 0.98)	5.07 (0.30, 85.66)	*0.09 (0.01, 0.53)	0.44 (0.05, 3.59)
Inhalant use none	---	---	---	---	---
abuse	2.49 (0.96, 6.41)	0.96 (0.28, 3.32)	2.69 (0.43, 16.74)	0.62 (0.12, 3.27)	1.67 (0.40, 7.02)
dependence	3.77 (0.40, 35.18)	0.24 (0.04, 1.69)	0.00 (0.00, 0.00)¥	0.36 (0.05, 2.51)	0.00 (0.00, 0.00)¥
Cocaine use none	---	---	---	---	---
abuse	*2.42 (1.44, 4.05)	0.99 (0.50, 1.96)	0.79 (0.28, 2.25)	1.06 (0.49, 2.32)	0.84 (0.33, 2.15)
dependence	*3.74 (2.44, 5.73)	0.65 (0.37, 1.12)	0.89 (0.37, 2.17)	0.67 (0.37, 1.22)	0.60 (0.26, 1.40)

Hallucinogen	none	---	---	---	---	---
	abuse	*2.21 (1.30, 3.77)	0.58 (0.30, 1.15)	1.04 (0.35, 3.13)	0.57 (0.27, 1.23)	0.60 (0.22, 1.65)
	dependence	1.16 (0.43, 3.11)	*0.27 (0.09, 0.84)	0.61 (0.07, 5.77)	0.31 (0.09, 1.06)	0.19 (0.02, 1.53)
Cannabis use	none	---	---	---	---	---
	abuse	*2.09 (1.56, 2.79)	0.89 (0.58, 1.37)	1.32 (0.66, 2.63)	0.81 (0.50, 1.30)	1.07 (0.56, 2.04)
	dependence	*2.28 (1.43, 3.65)	*0.45 (0.25, 0.82)	1.61 (0.66, 3.96)	*0.38 (0.19, 0.74)	0.61 (0.27, 1.40)
Amphetamine	none	---	---	---	---	---
	abuse	*2.05 (1.20, 3.52)	0.75 (0.38, 1.48)	1.75 (0.61, 5.03)	0.60 (0.27, 1.32)	1.05 (0.41, 2.70)
	dependence	*3.06 (1.68, 5.58)	0.71 (0.34, 1.50)	1.43 (0.45, 4.60)	0.63 (0.25, 1.59)	0.90 (0.35, 2.31)
Opioid use	none	---	---	---	---	---
	abuse	*2.19 (1.24, 3.86)	*0.45 (0.24, 0.95)	0.84 (0.29, 2.45)	0.51 (0.23, 1.11)	0.43 (0.16, 1.10)
	dependence	*5.49 (2.88, 10.5)	0.68 (0.34, 1.34)	0.42 (0.14, 1.27)	0.84 (0.38, 1.86)	*0.36 (0.16, 0.82)
Tranquilizer use	none	---	---	---	---	---
	abuse	*2.25 (1.22, 4.15)	0.66 (0.30, 1.45)	0.88 (0.26, 2.94)	0.68 (0.28, 1.67)	0.60 (0.20, 1.85)
	dependence	*6.63 (2.93, 15.0)	0.62 (0.27, 1.41)	0.56 (0.15, 2.08)	0.73 (0.29, 1.81)	0.41 (0.12, 1.37)
Sedative use	none	---	---	---	---	---
	abuse	*2.19 (1.07, 4.47)	0.50 (0.24, 1.05)	2.01 (0.46, 8.86)	*0.38 (0.17, 0.86)	0.76 (0.22, 2.64)
	dependence	*2.42 (1.06, 5.55)	*0.33 (0.14, 0.82)	0.26 (0.03, 2.24)	0.44 (0.17, 1.14)	*0.12 (0.02, 0.85)
Other drug use	none	---	---	---	---	---
	abuse	2.98 (0.68, 13.13)	4.75 (0.62, 36.34)	0.00 (0.00, 0.00)¥	7.11 (0.92, 54.72)	0.00 (0.00, 0.00)¥
	dependence	*24.0 (3.3, 176.4)	0.41 (0.05, 3.39)	0.00 (0.00, 0.00)¥	0.62 (0.08, 5.11)	0.00 (0.00, 0.00)¥
Personality disorder		*3.23 (2.59, 4.03)	*0.38 (0.28, 0.52)	*2.75 (1.69, 4.45)	*0.27 (0.19, 0.39)	0.74 (0.48, 1.15)
Borderline		*4.23 (3.29, 5.43)	*0.38 (0.28, 0.52)	*4.42 (2.76, 7.08)	*0.21 (0.14, 0.32)	0.94 (0.62, 1.42)
Antisocial		*2.67 (1.89, 3.76)	*0.49 (0.33, 0.73)	0.71 (0.30, 1.65)	*0.54 (0.34, 0.86)	*0.38 (0.18, 0.81)
Schizotypal		*2.85 (2.04, 3.97)	*0.48 (0.32, 0.72)	*3.18 (1.76, 5.74)	*0.30 (0.17, 0.51)	0.94 (0.58, 1.53)
Narcissistic		*2.21 (1.60, 3.05)	0.80 (0.54, 1.17)	*1.88 (1.04, 3.39)	0.63 (0.39, 1.00)	1.18 (0.71, 1.94)

* $p < 0.05$

¥ Unable to be calculated to extremely small cell size

Appendix 2H: Lost to follow up analysis, comparing final sample to those lost to follow-up on Wave 1 characteristics

		Lost to follow-up (n =2361) (Weighted %)	Final sample (n= 11090) (Weighted %)	X^2 p-value
Age	18-29	11.0	15.4	0.64
	30-64	70.4	69.2	
	65 +	18.6	15.4	
Sex	Female	65.8	62.4	0.60
	Male	34.2	37.6	
Education	≤ HS	26.0	12.6	0.03
	> HS	74.0	87.4	
Income	0-34,999	12.3	13.0	0.99
	35-69	20.2	20.2	
	70 +	67.5	66.8	
Marital Status	Never married	25.1	16.8	0.31
	Married	49.5	58.9	
	Divorced	25.5	24.3	
Lifetime Suicidal Ideation	No	42.1	42.7	0.72
	Yes	57.9	57.3	
Lifetime Depressive disorder	No	82.0	88.5	0.00
	Yes	18.0	11.5	
Lifetime Anxiety disorder	No	70.0	64.8	0.00
	Yes	30.0	35.2	
Lifetime substance use disorder	No	62.5	58.0	0.005
	Yes	37.6	42.0	

Note: Personality disorders of interest only assessed at Wave 2.

Appendix 3A: Measure of **"*Depressed mood*",** based on one question from the 12-Item Short-Form Survey Instrument (SF-12)[63]; How much of the time during the past 4 weeks have you felt downhearted and depressed?

Respondents reporting "depressed mood" at Wave 1

Response options	Total sample (%)	Concordant Yes (%)	Concordant No (%)	Recanters (%)	New endorsers (%)
None of the time	32.9	14.4	34.2	26.1	19.6
A little of the time	34.6	31.1	35.3	26.5	23.1
Some of the time	21.0	26.8	20.6	21.7	28.9
Most of the time	8.4	17.1	7.5	18.5	16.2
All of the time	3.1	10.6	2.4	7.2	12.2

Respondents reporting "depressed mood" at Wave 2

Response options	Total sample (%)	Concordant Yes (%)	Concordant No (%)	Recanters (%)	New endorsers (%)
None of the time	37.76	19.98	39.26	31.07	15.63
A little of the time	35.21	30.31	35.73	28.37	33.13
Some of the time	17.26	25.24	16.58	24.53	18.88
Most of the time	7.22	14.84	6.41	11.97	22.80
All of the time	2.55	9.63	2.03	4.06	9.56

Appendix 3B: Unadjusted odds ratios for stratum of "depressed mood" variable

	OR (95% C.I.) Recanters vs. Concordant Yes
Wave 1 - Depressed mood	
None of the time	----
Little of the time	0.47 (0.28, 0.77) *
Some of the time	0.44 (0.26, 0.76) *
Most of the time	0.59 (0.34, 1.03)
All of the time	0.37 (0.19, 0.74) *
Wave 2 - Depressed mood	
None of the time	---
Little of the time	0.60 (0.38, 0.96) *
Some of the time	0.62 (0.38, 1.04)
Most of the time	0.52 (0.29, 0.93) *
All of the time	0.27 (0.13, 0.56) *
	OR (95% C.I.) New endorsers vs. Concordant Yes
Wave 1 - Depressed mood	
None of the time	---
Little of the time	0.54 (0.29, 1.04)
Some of the time	0.79 (0.39, 1.60)
Most of the time	0.69 (0.34, 1.40)
All of the time	0.85 (0.36, 1.99)
Wave 2 - Depressed mood	
None of the time	---
Little of the time	1.40 (0.74, 2.64)
Some of the time	0.96 (0.52, 1.75)
Most of the time	1.96 (1.04, 3.72) *
All of the time	1.27 (0.61, 2.64)

* < 0.05

Appendix 3C: ***"Wave 1:Wave 2 depressed mood profiles"***, in the format of D_{xy}; where x= current depression status at Wave 1
(1= depressed, 0= not depressed), and y= depression status at Wave 2 (1= depressed, 0= not depressed).

D_{xy}	Total sample (%)	Concordant Yes (%)	Concordant No (%)	Recanter (%)	New Endorser (%)
D_{00}	82.5	60.5	84.6	68.6	52.8
D_{10}	7.7	15.1	7.0	15.4	14.8
D_{01}	6.0	11.9	5.5	5.7	18.8
D_{11}	3.8	12.5	2.9	10.3	13.6
Total	100	100	100	100	100

Appendix 3D: Assessing potential covariates. Bivariate associations between covariates of interest, and main exposure (current depressed mood) and outcome (Discordant reporters) of interest. All covariates are Wave 2 measures.

		Depressed mood (Wave 1) OR (95% CI)	Depressed mood (Wave 2) OR (95% CI)	Recanters vs. Concordant Yes OR (95% CI)	New endorsers vs. Concordant Yes OR (95% CI)
Age	18-29	---	---	---	---
	30-64	1.25 (1.07, 1.49)*	1.28 (1.08, 1.53)*	0.58 (0.38, 0.89)*	1.43 (0.79, 2.60)
	65 +	1.27 (1.05, 1.55)*	1.41 (1.17, 1.71)*	1.29 (0.67, 2.48)	0.86 (0.27, 2.73)
Sex	Female	---	---	---	---
	Male	0.57 (0.51, 0.64)*	0.67 (0.60, 0.76)*	1.33 (0.92, 1.92)	0.87 (0.54, 1.38)
Education	≤ HS	---	---	---	---
	> HS	0.38 (0.33, 0.43)*	0.39 (0.35, 0.44)*	0.72 (0.45, 1.13)	0.95 (0.52, 1.73)
Lifetime Suicidal Ideation	No	---	---	---	---
	Yes	2.28 (2.00, 2.62)*	3.98 (3.40, 4.66)*	0.16 (0.05, 0.52)*	0.04 (0.01, 0.11)*

* < 0.05

Note: Sex and education were not significantly associated with Recanting and New endorsing in the bivariate analyses. However, in Chapter 2 they became significant in the final adjusted analyses, and were found to be significant correlates in prior studies (see Chapter 1), therefore were included as potential covariates in Chapter 3 analyses.

Appendix 3E: Assessing degree of association between Lifetime Ideation at Wave 2, and Current Depressed Mood at Waves 1 and 2.

Lifetime Ideation Wave 2

Depressed mood Wave 1	**Freq % Row % Col %**	**Yes**	**No**
	Yes	7.8 67.4 16.4	3.8 32.6 7.1
	No	39.5 44.6 83.6	49.0 55.4 92.9

Lifetime Ideation Wave 2

Depressed Mood Wave 2	**Freq % Row % Col %**	**Yes**	**No**
	Yes	7.2 73.8 15.3	2.6 26.2 4.9
	No	40.0 44.4 84.7	50.2 55.6 95.1

Appendix 3F: Composite measure of Lifetime Suicidal Ideation at Wave 2. If the respondent endorsed any lifetime ideation item at Wave 1, or any "since last interview" ideation item at Wave 2, they were classified as having Lifetime Suicidal Ideation at Wave 2.

Suicidal Ideation Questions		**Total %**	**Concordant Yes %**	**Concordant No %**	**Recanter %**	**New Endorser %**
Wave 1 – Lifetime measures						
Did you ever think about committing suicide?	No	73.6	3.8	80.0	6.6	42.3
	Yes	26.4	96.2	20.0	93.4	57.7
Did you ever feel like you wanted to die?	No	68.0	3.6	73.9	7.4	35.8
	Yes	32.0	96.4	26.1	92.6	64.2
Did you ever think a lot	No	74.9	26.3	79.1	30.1	63.3

about your own death?	Yes	25.1	73.7	20.9	69.9	36.7
Any Wave 1Ideation	No	57.5	0.8	62.6	4.6	29.5
	Yes	42.5	99.2	37.4	95.4	70.5
Wave 2- Since last interview measures						
Did you ever think about committing suicide?	No	92.7	64.2	94.8	87.6	62.0
	Yes	7.3	35.8	5.2	12.4	38.0
Did you ever feel like you wanted to die?	No	89.2	61.8	91.5	84.1	44.6
	Yes	10.8	38.2	8.5	15.9	55.4
Did you ever think a lot about your own death?	No	90.7	68.6	92.3	87.5	68.5
	Yes	9.3	31.4	7.7	12.5	31.5
Any Ideation Since Last Interview	No	84.9	53.5	87.4	80.1	37.6
	Yes	15.1	46.5	12.6	19.9	62.4
Either Wave 1 or Wave 2 ideation	No	52.8	0.8	57.7	4.6	17.2
	Yes	47.2	99.2	42.3	95.4	82.8

Appendix 3G: Wave 1: Wave 2 depressed mood profile analysis, with reference groups containing only concordant attempt reporters and concordant mood reporters.

Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) Recanters** (vs. CY + CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10} vs. (D_{00} + D_{11})	2.30 (1.66, 3.19)*	2.13 (1.48, 3.07)*	1.60 (1.11, 2.30)*
D_{01} vs. (D_{00} + D_{11})	1.08 (0.65, 1.80)	1.03 (0.61, 1.75)	0.68 (0.40, 1.15)
D_{10} vs. **D_{01}**	2.13 (1.20, 3.78)*	2.11 (1.18, 3.76)*	2.44 (1.36, 4.40)*
Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) New Endorsers** (vs. CY + CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10} vs. (D_{00} + D_{11})	2.64 (1.52, 4.61)*	2.49 (1.47, 4.24)*	1.99 (1.14, 3.47)*
D_{01} vs. (D_{00} + D_{11})	4.23 (2.63, 6.82)*	4.04 (2.55, 6.41)*	2.95 (1.85, 4.69)*
D_{01} vs. **D_{10}**	1.60 (0.82, 3.14)	1.58 (0.81, 3.11)	1.48 (0.76, 2.89)

* < 0.05

Appendix 3H: Binomial logistic regression models examining odds of being a) Recanters vs. Concordant Yes; and b) New Endorsers vs. Concordant Yes, given respondent depression status at *i)* Wave 1 and *ii)* Wave 2. Current depressed mood assessed based on MDD and/or Dysthymia in the past year.

	OR (95% C.I.) Recanters vs. Concordant Yes	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
Wave 1			
Past year MDD	0.85 (0.60, 1.20)	0.82 (0.57, 1.18)	0.84 (0.59, 1.22)
Past year Dysthymia	0.44 (0.28, 0.71)*	0.45 (0.28, 0.72)*	0.47 (0.30, 0.75)*
Past year MDD or Dysthymia	0.80 (0.56, 1.14)	0.77 (0.54, 1.12)	0.80 (0.55, 1.15)

Wave 2			
Past year MDD	0.45 (0.31, 0.67)*	0.46 (0.31, 0.68)*	0.47 (0.32, 0.70)*
Past year Dysthymia	0.33 (0.15, 0.72)*	0.32 (0.15, 0.71)*	0.34 (0.15, 0.75)*
Past year MDD or Dysthymia	0.46 (0.32, 0.68)*	0.47 (0.31, 0.69)*	0.48 (0.32, 0.72)*
	OR (95% C.I.) New endorsers vs. Concordant Yes	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
Wave 1			
Past year MDD	0.89 (0.58, 1.36)	0.88 (0.56, 1.39)	1.02 (0.65, 1.60)
Past year Dysthymia	0.50 (0.29, 0.85)*	0.48 (0.28, 0.83)*	0.55 (0.32, 0.95)*
Past year MDD or Dysthymia	0.77 (0.50, 1.20)	0.77 (0.49, 1.22)	0.89 (0.56, 1.40)
Wave 2			
Past year MDD	2.58 (1.60, 4.17)*	2.60 (1.61, 4.18)*	3.33 (2.11, 5.26)*
Past year Dysthymia	1.15 (0.61, 2.16)	1.12 (0.59, 2.15)	1.26 (0.67, 2.39)
Past year MDD or Dysthymia	2.55 (1.58, 4.13)*	2.56 (1.59, 4.15)*	3.29 (2.08, 5.20)*

AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation
* < 0.05

Appendix 3I: Discordant responders (Recanters, New endorsers) odds of Wave 1:Wave 2 depressed mood profiles (as measured by Past year MDD)

Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) Recanters** (vs. NE, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	1.99 (1.46, 2.72)*	1.76 (1.28, 2.41)*	1.42 (1.03, 1.97)*
D_{01}†	1.03 (0.68, 1.56)	1.02 (0.67, 1.55)	0.66 (0.44, 1.01)
D_{10} vs. **D_{01}**	1.66 (1.02, 2.71)*	1.58 (0.97, 2.55)	1.90 (1.18, 3.06)*
Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) New Endorsers** (vs. Rec, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	0.66 (0.41, 1.08)	0.60 (0.36, 1.00)	0.52 (0.31, 0.86)*
D_{01}†	4.96 (3.22, 7.65)*	4.65 (3.01, 7.17)*	3.54 (2.31, 5.43)*
D_{01} vs. **D_{10}**	5.24 (2.93, 9.37)*	5.41 (2.94, 9.97)*	4.91 (2.71, 8.88)*

† (reference = all other mood profiles)
AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation
* < 0.05

Appendix 3J: Discordant responders (Recanters, New endorsers) odds of Wave 1:Wave 2 depressed mood profiles (as measured by Past year Dysthmia)

Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) Recanters** (vs. NE, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	2.46 (1.65, 3.69)*	2.42 (1.61, 3.63)*	1.49 (0.99, 2.25)
D_{01}†	1.30 (0.57, 2.96)	1.14 (0.49, 2.64)	0.81 (0.35, 1.87)
D_{10} vs. **D_{01}**	1.77 (0.71, 4.45)	1.94 (0.73, 5.12)	1.74 (0.68, 4.48)
Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) New Endorsers** (vs. Rec, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	2.32 (1.30, 4.14)*	2.14 (1.20, 3.80)*	1.48 (0.84, 2.63)
D_{01}†	4.43 (2.40, 8,17)*	3.98 (2.09, 7.58)*	3.00 (1.56, 5.75)*
D_{01} vs. **D_{10}**	1.93 (0.90, 4.16)	1.98 (0.88, 4.42)	2.13 (0.95, 4.78)

† (reference = all other mood profiles)
AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation
* < 0.05

Appendix 3K: Discordant responders (Recanters, New endorsers) odds of Wave 1:Wave 2 depressed mood profiles (as measured by Past year MDD or Past year Dysthymia)

Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) Recanters** (vs. NE, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	2.01 (1.47, 2.75)*	1.79 (1.31, 2.44)*	1.45 (1.05, 2.00)*
D_{01}†	1.05 (0.69, 1.60)	1.03 (0.67, 1.57)	0.68 (0.45, 1.05)
D_{10} vs. **D_{01}**	1.64 (1.00, 2.67)	1.54 (0.95, 2.50)	1.83 (1.13, 2.96)*
Wave 1: Wave 2 Depressed Mood Profiles	**OR (95% C.I.) New Endorsers** (vs. Rec, CY, CN)	**AOR[1] (95% C.I.)**	**AOR[2] (95% C.I.)**
D_{10}†	0.62 (0.37, 1.05)	0.57 (0.34, 0.97)	0.49 (0.29, 0.83)*
D_{01}†	4.99 (3.24, 7.67)*	4.49 (3.06, 7.21)*	3.61 (2.36, 5.52)*
D_{01} vs. **D_{10}**	5.55 (3.03, 10.16)*	5.74 (3.08, 10.70)*	5.22 (2.84, 9.61)*

† (reference = all other mood profiles)
AOR[1] = Adjusted for age, sex, education
AOR[2] = Adjusted for Wave 2 age, sex, education, and lifetime suicidal ideation

Appendix 3L: Lost to follow-up analysis

		Lost to follow-up (n =2361) (Weighted %)	Final sample (n= 11090) (Weighted %)	X^2 p-value
Age	18-29	11.0	15.4	0.64
	30-64	70.4	69.2	
	65 +	18.6	15.4	
Sex	Female	65.8	62.4	0.60
	Male	34.2	37.6	
Education	≤ HS	26.0	12.6	0.03
	> HS	74.0	87.4	
Wave 1 lifetime suicidal ideation	No	57.9	57.5	0.72
	Yes	42.1	42.6	
Wave 1 current depressed mood	No	82.0	88.5	0.000
	Yes	18.0	11.5	