Non-response to sad mood induction: implications for emotion research

Jonathan Rottenberg, Maria Kovacs & Ilya Yaroslavsky

To cite this article: Jonathan Rottenberg, Maria Kovacs & Ilya Yaroslavsky (2017): Non-response to sad mood induction: implications for emotion research, Cognition and Emotion, DOI: 10.1080/02699931.2017.1321527

To link to this article: https://doi.org/10.1080/02699931.2017.1321527

Published online: 03 May 2017.
Non-response to sad mood induction: implications for emotion research

Jonathan Rottenberg, Maria Kovacs and Ilya Yaroslavsky

ABSTRACT
Experimental induction of sad mood states is a mainstay of laboratory research on affect and cognition, mood regulation, and mood disorders. Typically, the success of such mood manipulations is reported as a statistically significant pre- to post-induction change in the self-rated intensity of the target affect. The present commentary was motivated by an unexpected finding in one of our studies concerning the response rate to a well-validated sad mood induction. Using the customary statistical approach, we found a significant mean increase in self-rated sadness intensity with a moderate effect size, verifying the “success” of the mood induction. However, that “success” masked that, between one-fifth and about one-third of our samples (adolescents who had histories of childhood-onset major depressive disorder and healthy controls) reported absolutely no sadness in response to the mood induction procedure. We consider implications of our experience for emotion research by (1) commenting upon the typically overlooked phenomenon of nonresponse, (2) suggesting changes in reporting practices regarding mood induction success, and (3) outlining future directions to help scientists determine why some subjects do not respond to experimental mood induction.

Introduction: a surprising finding
Experimental induction of sad mood states is a mainstay of laboratory research on the interface of affect and cognition, mood regulation, and mood disorders (Martin, 1990). In our recent study of how adolescents with a history of major depressive disorder (and normal controls) regulate sad mood (Kovacs et al., 2015), we encountered an unexpected result with our sad mood induction, a 163-second clip from The Champ, which depicts a boy’s immediate reactions to the death of a loved one. The Champ clip has been a mainstay of sadness induction in the field of affective science and has been extensively used with paediatric and adult samples (e.g. Gross & Levenson, 1995; Rottenberg, Gross, Wilhelm, Najmi, & Gotlib, 2002). Furthermore, prior to selecting the Champ clip, we did extensive pilot studies that compared it with alternative film clips (Kovacs et al., 2015, Online Appendix). It was absolutely essential for our study that subjects report sadness after the induction procedure, because our research question concerned how youth take advantage of specific opportunities to repair sadness (i.e. we could not determine whether a youth succeeded at mood repair unless there was sad mood to repair).

When we used the customary approach to determine the success of the mood induction procedure, all appeared well. A repeated measures ANOVA on self-ratings of sadness revealed that, after watching the Champ clip, sadness increased in the overall sample with a moderate effect size. However, another result lurked: a simple frequency count revealed that one-fifth of the emotionally healthy, control adolescents, and somewhat more than one-third of the adolescents with childhood-onset depression histories, reported absolutely no sadness in response to our induction procedure (Kovacs et al., 2015).
Searching for whether our rates of nonresponse were unusual, we found that very few articles actually reported response rates to mood inductions in their samples. And, in the literature as a whole, we found strikingly little discussion of mood induction failure. Given the critical importance of successful sad mood induction across a number of research areas, these seemed to be important gaps in the literature that made us pause and reflect. In the present article, we first consider the history of mood induction success and the implications of nonresponse. Second, we discuss the need to improve scientific reporting conventions for mood inductions and propose alternatives. Finally, we call for future research to better understand why some participants will not respond to even the most robust mood manipulation procedures.

**Historical background: how to define mood induction success?**

How do we judge the success of a sad mood manipulation in the laboratory? Historically, most definitions of a sad mood induction success consider to what extent a sample on average reports a statistically significant change in the targeted affect. This change may be within-subjects, such as a pre- to post-induction intensity ratings of a target affect, or a between-subjects metric post-induction across-group differences in ratings of the target affect, or less commonly, a predetermined percentage point change in affect intensity from pre- to post-induction (Martin, 1990). By such metrics, induction techniques for various affects have been judged to succeed from 50% to 75% of the time, and the general conclusion in the field has been that most experimental mood induction techniques are “effective” (Gerrards-Hesse, Spies, & Hesse, 1994; Martin, 1990; Westermann, Spies, Stahl, & Hesse, 1996). Film clips or stories have often been seen as among the strongest induction approaches, based on narrative (Rottenberg, Ray, & Gross, 2007) and meta-analytic reviews (Westermann et al., 1996; but see Zhang, Hui, & Barrett, 2014). Effect sizes for sad mood induction via films/stories have been reported to range from medium to large, partly depending on the instructions provided to subjects (Westermann et al., 1996). Using individual change scores, Gerrards-Hesse et al. (1994) concluded that sadness induction was effective for about 70% of the cases, but the estimates varied wildly from 30% to 93%.

Despite decades of research using mood induction, and major reviews focusing specifically on the issue of its efficacy (i.e. Lench, Flores, & Bench, 2011), the field has moved away from reporting success rates or the actual distribution of scores. Indeed, since reviews conducted in the 1990s (Gerrards-Hesse et al., 1994; Martin, 1990; Westermann et al., 1996), there has been scant notice of the phenomenon of nonresponse. We thus had difficulty judging whether our rates of nonresponse to mood induction were all that unusual. Furthermore, the reports that we did locate used varying definitions of nonresponse. For instance, in two studies, 18% (Singer & Dobson, 2009) and 14% (Singer & Dobson, 2007) of remitted depressed adults exhibited a poor response to negative mood induction, based on an a priori defined percentage change in affect rating. In another study, response was defined as a rating of at least 3 on a 1–9 point intensity scale: 22% of the depressed but none of the remitted depressed adults were deemed to be nonresponders to a sad film (Werner-Seidler & Moulds, 2012). Furthermore, in a sample of normal adolescents, in vivo peer rejection (used for mood induction) elicited a response, defined as “marked deterioration in state mood” or a reliable change (see Christensen & Mendoza, 1986) only in 38.1% of the youths (Reijntjes, Stegge, Terwogt, Kamphuis, & Telch, 2006). In contrast, using multiple sadness induction techniques, and defining a response as mouse clicks in a predetermined negative region of an “affect grid”, response rates of close to 95% also have been reported (Larcom & Isaacowitz, 2009).

Because lack of a uniform definition of a “response” makes it difficult to compare nonresponse rates and related findings across studies, we suggest two standardised options. First, nonresponders can be defined as those participants not reporting any of the target affect subsequent to the mood manipulation: this is the approach we used. This definition of nonresponse has two advantages: it is clear-cut, and it acknowledges that some subjects may already have the required mood, irrespective of the induction. However, some contexts will require a different definition. In particular, studies that hinge upon demonstrating the effects of mood manipulation (rather than just the presence of the required mood) should define nonresponders as those participants who do not report any increase in the target affect subsequent to the manipulation. This particular definition would exclude those cases who report none of the target affect and those who report either unchanged or
decreased levels of the target affect from baseline to the manipulation. Some investigators have proposed using a certain percent-change in affect self-rating as the criterion of response to a mood manipulation. However, unless the magnitude of change is germane to a particular study, there is no clear standard for what percent of change should define a responder.

**Why should we be concerned about nonresponse?**

We have no reason to doubt that our experience, in which one out of five participants failed to report any reaction to a mood induction, is atypical. What are the implications for the field? Why should we be concerned? One concern is that current practices may provide a misleading picture of our experimental induction methods: By reporting only mean level change, many scientific reports leave the impression of robust sad mood manipulation: A manipulation is judged to have worked when it may have worked only for some participants. Therefore, current reporting practices may camouflage methodologically weak procedures.

A related concern is that different samples are likely to vary considerably in rates of nonresponse. But since nonresponse is rarely reported, this source of between-study heterogeneity will remain hidden from investigators. Given the centrality of sad mood manipulation to studies that examine the correlates of sadness, the presumed heterogeneity in response rates represents error variance, and ultimately makes it less likely that findings are replicable across samples. This issue is particularly acute in light of recent attention on problems in replicating results in the field of psychology (Open Science Collaboration, 2015).

Third, the hidden group of nonresponders may lead to misinterpretations of research findings. To give one example, Sheppard-Sawyer, McNally, and Fischer (2000) induced sad mood with a film in a sample of restrained and unrestrained eaters and then used the standard manipulation check to verify mood induction success. These authors reported that restrained eaters showed greater increases in popcorn consumption than unrestrained eaters after the sad film, a finding that has implications for how eating disorders develop. Alternatively, if unrestrained eaters were simply less likely to respond to mood induction, affective reactivity rather than restraint may be the explanatory variable (e.g. people who do not react to the induction will not alter their popcorn consumption). Similarly, Williams, Barnhofer, Crane, and Beck (2005) reported that formerly depressed subjects with a history of suicide attempts displayed a greater deterioration of problem-solving following a Velten mood manipulation than did formerly depressed subjects with no history of suicide attempts or healthy controls. Nonresponse to mood manipulation was not reported. Thus, it is conceivable that the groups that exhibited less deterioration of problem-solving were more likely to be nonresponders to the mood induction. If Williams et al.’s (2005) findings were not due to differential behaviour responses to the same mood, but instead represent different responses to mood induction, the treatment implications of the results would differ from those proposed by the authors.

Importantly, it may be inappropriate to include in some study samples those subjects who report no sad mood. For example, if an investigator is examining the effects of induced sad mood on another construct (whether it be attention, memory, or amygdala activation), any viable interpretation of the results is predicated on the presence of sad mood. Logically, then, participants who do not report experiencing that mood do not belong in the study. In our study cited earlier, it would not have been possible to make inferences about subjects’ skills at repairing sad mood if the subjects did not first experience some degree of sadness.

**What should be done? Implications for the field of emotion research**

Although this article has specifically focused on sad mood induction, and the largely hidden nature of nonresponse, the implications apply broadly to affect induction, and highlight the need for change in emotion science practices. First, we suggest changes in contemporary reporting practices. The prevailing convention regarding how mood manipulation success is reported – the mean change across a sample – is insufficient. We recommend that investigators adopt our definitions of nonresponse and report its rates. Descriptors of the score distributions also should be reported to facilitate across-study comparisons. When possible, the reports should include hit rates; the hit rate is the proportion of the sample that reports the targeted affect to a greater degree than some non-targeted affect (Gross & Levenson, 1995).
Such information corrects the overreliance on the mean response for validity checks; it also provides a critical basis for comparing findings across studies, which is particularly pressing when sadness induction continues to be a popular experimental manipulation (Rottenberg et al., 2007; Schaefer, Nils, Sanchez, & Phillips, 2010). Once it becomes routine to report nonresponse rates, it will be possible to use meta-analysis to determine which types of mood inductions minimise them (e.g., films, music) and how nonresponse is influenced by methodological factors (e.g., duration of the induction).

Second, there should be efforts to reduce nonresponse rates by introducing procedural changes to mood induction. For example, some investigators have altered the duration of mood induction so that the experiment continues only when subjects report achieving a specific mood severity criterion (Liotti et al., 2000). Although this may reduce nonresponse rates, different participants are likely to need different “doses” of mood inductions to reach the same threshold (which could be problematic for certain research questions). Moreover, requiring a certain level of affect intensity to be present before the experiment can proceed carries with it an increased risk of experimental demand, perhaps to unacceptable levels.

However, even subtle changes to instructions may reduce nonresponse rates. For example, for film viewing, some studies have altered the typical “watch” instructions by telling subjects to put themselves in the scene or the role of the protagonist (e.g., Werner-Seidler & Moulds, 2012). Nonresponse also may be minimised by combining mood induction procedures, such as listening to depressogenic music while engaging in sad self-statements (Larcom & Isaacowitz, 2009). Finally, investigators may consider different kinds of stimuli entirely, such as personally tailored, sad stimuli (e.g., recall of one’s worst life moments), as a strategy to combat nonresponse (e.g., Rottenberg, Gross, & Gotlib, 2005).

Given that nonresponders are likely to persist, even with the most robust procedures, sound statistical analyses are the final type of countermeasure. We believe that the simplest remedy, to delete nonresponders to mood induction from the analyses, is often the best (e.g., Werner-Seidler & Moulds, 2012). However, if authors want to retain subjects who fail to respond to mood induction, we recommend that: (a) the decision be justified, (b) nonresponse rates be reported, and (c) an additional categorical (yes, no) variable be included in statistical models to account for the variance associated with nonresponse.

Finally, it may be time to think more broadly about how we validate mood inducing stimuli. Historically, validation studies and experiments typically have prioritised the subjective self-rated component of affective experience as the criterion of successful mood induction (e.g. Zhang et al., 2014). However, eliciting a mood also can be expected to induce changes in cognition, physiology, and/or behaviour (Lench et al., 2011). Thus, a more stringent validation of mood induction success might require converging evidence from several response systems (i.e. behavioural change and changes in subjective experience) in a given subject. Indeed, bolstering the need for multi-method verification, low correspondences are often observed between affective response domains (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). Thus, mood inductions that are “successful” via the criterion of subjective experience may fail to systemically alter mood-related physiology, cognition, and/or behaviour.

**Towards an understanding of nonresponse**

In this comment, we have approached nonresponse to sadness induction as a methodological problem in research. But it is equally important to understand why nonresponse occurs in the first place, which then raises both methodological and substantive issues. That is, nonresponse to sadness is not only “nuisance variance” but may reveal important individual differences in affect processing (Brenner, 2000). Indeed, some subjects will not evidence a response to any mood induction procedure, no matter how robust it is. Future work must be directed to gain a better (systematic) understanding of how best to characterise these nonresponders.

Several factors may contribute to nonresponse. For instance, if the sad stimulus depicts loss and grief in a protagonist, then difficulties with empathy and/or perspective taking (poor socioemotional functioning) may explain nonresponse among some participants. Apparent nonresponse in other participants may reveal a lack of insight into or monitoring of internal states, as is seen in conditions such as alexithymia (Taylor, Bagby, & Parker, 1991), or among subjects high in trait avoidance of negative experiences (e.g., experiential avoidance, Kashdan, Barrios, Forst, & Steger, 2006). Nonresponse to mood induction may also reflect differences in reporting biases, as when...
participants may experience but are unwilling to report negative affect because of reasons such as embarrassment. Finally, nonresponse to sad mood induction may reflect individual differences in central nervous system processes, such as hypoactivity of subcortical structures (Kiehl et al., 2001).

To provide an illustration of the kinds of analyses that might be useful, we return to our preliminary efforts to characterise nonresponders in our sample of adolescents. Our analyses (reported in supplementary material) controlled for key variables and examined the potential roles of clinical and physiological variables. A clinically useful finding was that youths with depression histories and comorbid conduct/oppositional disorders were more likely to be nonresponders to sadness induction than those without such comorbidities, possibly because the sad stimulus film required the ability to empathise. In the physiological domain, we found preliminary evidence that atypical profiles of respiratory sinus arrhythmia (RSA), an index of the functioning of the autonomic nervous system, were associated with nonresponse to mood induction (Yaroslavsky, Rotenberg, & Kovacs, 2013). Because RSA has been implicated in attention deployment (Suess, Porges, & Plude, 1994), nonresponse to sadness induction could mirror impaired attention to stimuli with sad affective content.

Finally, we note that our entire sample, including our nonresponders, consisted of adolescents. Along with Reijntjes et al. (2006), the high rates of nonresponse in our study raise the question as to whether some (or all) adolescent age groups may be particularly “resistant” to sad mood induction. Such a possibility, and reasons for it, also warrants further study, and underscores the importance of incorporating a developmental perspective in future work on the phenomenon of nonresponse to sadness induction.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This study has been supported by NIMH [grant number MH084938].

References


