

# Response Bias in Research on Religion, Spirituality and Mental Health: A Critical Review of the Literature and Methodological Recommendations

Everton de Oliveira Maraldi<sup>1</sup> 

Published online: 16 May 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** Although a significant body of research supports the psychological benefits of religion and spirituality, more investigations are needed to understand the mechanisms by which they impact mental health. While some studies suggest a causal direct influence, the findings may still be subject to unmeasured factors and confounders. Despite compelling empirical support for the dangers of response bias, this has been a widely neglected topic in mental health research. The aim of this essay is to critically examine the literature addressing the role of response bias in the relationship between religion, spirituality and mental health. A survey of the diverse types of bias in this research area is presented, and methodological and theoretical issues are outlined. The validity and generalizability of the evidence are discussed, as well as the implications for mental health practice. A list of methodological remedies to reduce bias is suggested. The article is then concluded with a summary of the studies reviewed and directions for future research.

**Keywords** Response bias · Religion · Spirituality · Mental health · Social desirability

## Introduction

While religiosity is circumscribed to a more formal commitment to specific religious traditions and practices, spirituality is usually defined in terms of a general and relatively informal adherence to transcendent meanings and beliefs that may or may not arise from involvement with a religious community or belief system (Huguelet and Koenig 2007). But

---

✉ Everton de Oliveira Maraldi  
evertonom@usp.br

<sup>1</sup> Social Psychology of Religion Laboratory, Department of Social and Work Psychology, Institute of Psychology, University of São Paulo (USP), Prof Mello Moraes Avenue, 1721, Cidade Universitária, São Paulo, SP 05508-030, Brazil

despite terminological differences, many are the similarities between the two concepts, and their differentiation is considered complex (Zimbauer and Pargament 2005).

A significant body of research supports the existence of a positive association between religiosity/spirituality (R/S) and a range of mental health indicators, from subjective well-being to enhanced coping capacity (Bonelli and Koenig 2013; Koenig 2009; Pargament and Lomax 2013). There is some controversy, however, as to the generalizability of this relationship, with some studies showing it depends on how spirituality is defined (e.g., Lindeman et al. 2012) and whether contrast groups (such as atheists or agnostics) are included (Galen 2012a, b; Zuckerman 2009; Zuckerman et al. 2016). More investigations are also needed to understand the precise mechanisms by which R/S impacts mental health (Peres et al. 2017). Although some studies suggest a causal direct influence, the findings may still be subject to unmeasured factors and confounders (VanderWeele et al. 2017).

Of all the variables considered, little attention has been paid to response bias. Psychologists have long recognized the role of impression management and demand characteristics on questionnaire responses (Rosenthal 1976). Motivations such as social desirability or the need for consistency in self-representations can influence a wide range of measures, even when unintended (Council 1993). Instead of accurate portrayals of the respondents' mental states, their answers to scale items are significantly influenced by contextual factors, including the expectations and beliefs of the experimenter (Holman et al. 2015).

But despite compelling empirical support for the dangers of response bias (Holman et al. 2015; Reese et al. 2013; Knäuper and Hans-Ulrich 1994), ranging from false conclusions based on spurious correlations between variables and errors in prevalence rates of mental disorders, this has been a widely neglected topic in health research (Rogler et al. 2001; van de Mortel 2008). In a systematic review of nursing and health studies, van de Mortel (2008) verified that, from a total of 14,275 investigations, only 0.2% (31) used a scale to measure social desirability. Of these, 43% found socially desirable responding significantly biased their results. However, only 10% statistically controlled for such effects in the analyses. Similar findings were obtained by Rogler et al. (2001) in their review of the mental health literature. According to the authors, “a very small minority of the articles reviewed mentioned response bias and [...] among those mentioning it, a minority attempted to control for bias effects” (p. 182). These findings are illustrative of how often response biases are ignored in health research, despite their demonstrated influence on research results, with studies on religiosity and spirituality being no exception, as we shall see.

The aim of the present essay is to critically examine the literature addressing the role of response bias in the relationship between religion, spirituality and mental health. A survey of the diverse types of bias in this research area (from social desirability to contextual and experimenter effects) is presented, and methodological and theoretical issues are outlined. The validity and generalizability of the evidence regarding such biases are discussed, as well as the implications for research and mental health practice. A list of methodological remedies to help reduce them and advance this field of investigation is suggested. The article is then concluded with a summary of the studies reviewed and directions for future research.

## Social Desirability Bias

The literature specifically investigating the subject of response bias in research on the mental health benefits of religiosity/spirituality is relatively scarce. The studies available were concerned mainly with social desirability, a type of response bias characterized by

socially expected responses, either involving over-reporting of appropriate and positive behaviors or underreporting of negative characteristics, such as psychopathological symptoms. By means of such strategies, participants aim to project a favorable image of themselves, although sometimes their use might be unconscious, as when participants have a distorted positive self-image (Paulhus and Vazire 2007). Response patterns may eventually suggest the occurrence of social desirability (for example, when participants choose only positive responses or consistently avoid answering controversial questions), but there are many scales specifically designed to evaluate respondents' general tendency to give desirable answers; the most widely used is the Marlowe–Crowne Social Desirability Scale or MCSD (Crowne and Marlowe 1960), which was also the most frequently employed measure of desirable responding in the studies reviewed. Another commonly reported measure was the Balanced Inventory of Desirable Responding BIDR (Paulhus 1991).

Ellis and Smith (1991), MacDonald (1997) and Migdal and MacDonald (2013) found that both subjective well-being and spirituality were significantly correlated with different aspects of social desirability from self-deceptive enhancement to impression management (i.e., when individuals attempt to manage or control the impression that others form of them). Given previous research indicating that well-being might not be a good predictor of spirituality (King et al. 2013; Lindeman et al. 2012), and given moderate correlations between socially desirable responses and different well-being measures (including the construct of “spiritual well-being,” Migdal and MacDonald 2013; Ellis and Smith 1991), the above findings were interpreted as suggesting that the psychological benefits of spirituality are indicative of illusory mental health, a form of defensive denial of suffering or distress (MacDonald 1997; Shedler et al. 1993).

Similar findings were also observed for measures of religiosity. Some of the first investigations in this regard were conducted in relation to Eysenck's dimension of psychoticism (Eysenck and Eysenck 1991; Eysenck 1998). Studies found religiosity to be inversely correlated with psychoticism and positively associated with scores on the Eysenck's Lie scale, which can be considered a measure of social desirability (Francis 1992; Francis et al. 1983; Gillings and Joseph 1996; Pearson and Francis 1989). These findings were interpreted as an indication that religious individuals might underreport psychopathological symptoms in order to project a healthy and socially desirable image of themselves, free of symptoms of mental disorders.

Another group of investigations concerned the role of social desirability in intrinsic religiosity. In contradiction with Allport and Ross (1967) account of *extrinsic* religiosity (i.e., religion seen as a means to obtain social and personal benefits) as more vulnerable to self-enhancement and social desirability, some studies identified positive and significant correlations between *intrinsic* religiosity and socially desirable responding, suggesting that descriptions of religion as an end in itself or as having ultimate significance in life might be influenced by social expectations (Batson et al. 1978; Leak and Fish 1989; Trimble 1997). The results remained significant even after controlling for content overlap between scales (Leak and Fish 1989). These findings are important for our discussion of the relationship between religion and mental health because intrinsic religiosity is often considered to have the strongest relationship with well-being and other mental health indicators (Huguelet and Koenig 2007; Trimble 1997).

Not only was social desirability related to general religiosity scales but also to health-related measures of religiousness, such as religious coping (Aguilar-Vafaie and Abiari 2007). Many other studies investigated the association of religiosity with social desirability, not necessarily in relation to mental health. Crandall and Gozali (1969), for example, verified this association in a sample of religious children, indicating that it may

start early in life as a result of upbringing. Socially desirable responding was also found to influence more “objective” behavioral measures, such as church attendance and frequency of religious practice (e.g., Hadaway et al. 1998), which raise concerns about research on the impact of religious involvement on mental health indicators, especially when results are based solely on self-report.

It must be stressed, however, that the effect sizes in these studies were usually weak, and researchers were not always able to replicate the initial findings (Fastame et al. 2017; Gillings and Joseph 1996; Lewis 2000; Plante et al. 2000; Watson et al. 1986). The most comprehensive meta-analysis assessing the relationship between religiosity and social desirability found a positive correlation of only 0.108 ( $p < 0.001$ ) (Sedikides and Gebauer 2010). Moreover, the results are not generalizable to all aspects of religiosity. For example, in Crandall and Gozali (1969) study with religious children, most participants were Catholics submitted to a “rigorous and doctrinaire” (p. 753) training in the context of parochial schools. For this reason, their socialization experiences were considered more demanding than that of most religious children, a finding that could explain their higher social desirability scores. Another important problem with such studies is the fact that they are predominantly composed of Christian samples; it is thus unclear whether the association with social desirability can be reliably extended to members of other religious traditions.

In the meta-analysis conducted by Sedikides and Gebauer (2010), the association of religiosity with social desirability showed to be moderated by culture; countries displaying higher levels of religious involvement also evidenced a stronger positive relationship between the two variables. Such pattern was also found to depend on the facet of religiosity that was under investigation: whereas intrinsic religiosity was positively related to social desirability, religion-as-quest (i.e., religiosity seems as an existential search for truth independent of an involvement with organized religion) was negatively correlated. Stavrova et al. (2013) found that religious individuals scored higher in measures of happiness and satisfaction with life when compared with non-religious individuals, but this seemed to partially depend on how much social recognition religion had in each country investigated.

The evidence regarding an association between religiosity and social desirability is sometimes difficult to interpret. Trimble (1997) suggested that socially desirable responses somehow capture true religiosity. More specifically, it might be that religious individuals are genuinely more sensitive to moral values and social norms, and this is reflected in their questionnaire responses. This could eventually explain the finding that intrinsic religiosity was found to be positively related to social desirability. In fact, from this perspective, religiosity and social desirability would represent two sides of the same coin, since religion would be essentially a form of establishing social cohesiveness and support. Such account is consistent with a long Durkheimian tradition in religious studies (Durkheim 1912), with the view of social desirability as a substantive trait or general personality factor (e.g., van der Linden et al. 2016), as well as with theories describing religion in terms of prosociality (e.g., Norenzayan et al. 2016). It might even be imagined that it is precisely the social desirability of religion which explains its health benefits, as can be inferred from Gelfand et al. (1965) finding that religiosity and social desirability were the most important predictors of pain tolerance and placebo responsiveness.

After a thorough review of the evidence, Galen (2012a) suggested, however, that the religious prosociality might actually consist in a desire to appear prosocial, instead of genuine prosociality, a hypothesis that finds some support in studies showing how religious practices and beliefs are influenced by reputational concerns, stereotype biases, and ingroup favoritism. The debate continues (Galen 2012b, 2016).

## Going Beyond Social Desirability

Response bias is not restricted, however, to a general tendency toward social desirability; it encompasses many other forms of response sets. One example is what Council (1993) has called “contextual effects.” Contextual effects are usually observable when different measures are administered in the same assessment condition. Council (1993) verified, for example, that when measures of paranormal belief and psychopathological symptoms were administered together, on a single occasion, no correlation could be identified between them. But when the same measures were answered on separate occasions, without participants being able to infer the purposes of the study, positive and statistically significant correlations emerged. Contextual effects might be influenced by an individual’s tendency to give socially desirable responses, but are relatively independent of it. In the case of Council’s study, the presentation of both scales in the same context elicited the occurrence of transient defensive responding; participants underreported psychopathological symptoms as a way to defend their belief systems and self-image from a negative evaluation.

The concept of social desirability emphasizes the role of the respondent in the distortion of results. However, the experimenter’s role is equally important. Scientific research is not an inherently neutral activity. Scientists are vulnerable to a series of cognitive and social biases that influence their methods and results (Klein et al. 2012; Podsakoff et al. 2003; Rosenthal and Rubin 1978). According to Holman et al. (2015, p. 1), “these biases are strongest when researchers expect a particular result, are measuring subjective variables, and have an incentive to produce data that confirm predictions.”

Since nineteenth-century speculations on a “personal equation” underlying subtle differences between astronomers’ calculations, many psychologists—including pioneers such as Wilhelm Wundt (1832–1920), William James (1842–1910) and Carl Gustav Jung (1875–1961)—have dedicated their attention to the unintended role played by the experimenter in his/her observations (Shamdasani 2003). Rosenthal (1976, 1994) is known to have developed an entire research program devoted to the investigation of “unconscious experimenter bias,” “observer bias” and “interpersonal expectancy effects.”

Rosenthal and collaborators could demonstrate that expectancy effects have a significant impact on different psychological measures, not only with methodological implications but with social consequences as well, as exemplified by their study of teachers’ expectations regarding students’ achievements in the classroom (Rosenthal and Jacobson 1992). Despite criticisms as to the generalization and validity of the claimed effect (e.g., Barber and Silver 1968), the problem of confirmation bias still attracts the attention of psychologists and other scientists alike (Holman et al. 2015; Klein et al. 2012).

Despite considerable evidence of its importance “blind data recording is often neglected, and its use appears to vary between scientific disciplines” (Holman et al. 2015, p. 2). In a survey of the prevalence of blind assessment in different research areas, Sheldrake (1999) found that only 24.2% of studies in medical sciences and 4.9% of investigations in psychology and animal behavior used blind or double-blind methodologies. A slightly better picture was described years later by Watt and Nagtegaal (2004) in a replication study which found 36.8% of publications in medical sciences and 14.5% of studies in psychology and animal behavior reporting blind methods. It is known that observer bias tends to inflate effect sizes and even foster spurious results. In their meta-analyses of studies in life sciences, Holman et al. (2015) found that lack of blindness increased effect sizes up to 27% exaggerating the benefits of clinical interventions by 22%. The authors concluded that observer bias is both “common and strong” (p. 9).

The research concerning experimenter effects has identified many different situations in which an observer may involuntarily bias scientific data, ranging from simple psychological measures such as IQ test scores to studies of athletic and scholastic performance (Rosenthal 1994), but an investigation of the occurrence and impact of such effects in the study of the mental health benefits of religion/spirituality has yet to be undertaken. The history of psychology and psychiatry is filled with narratives both in favor and in disfavor of religion, with authors such as Ellis (1980) and Freud (1927) emphasizing its negative consequences, while Pfister (1928) and others have pointed out its positive and adaptive features. These are not simply empirical matters. As Bergin (1983, p. 171) has observed, “values and ideology influence theoretical axioms. Conceptions of personality and psychopathology have subjective as well as empirical bases, as do rationales for intervention and goals of outcome.”

It is thus reasonable to conjecture that those who are in favor of a positive link between religious/spiritual practices and health (be they researchers or respondents) would not wish to see their convictions proved wrong. The same could be expected of those who are incredulous of such a link. Researchers from both sides of the debate usually disseminate their ideas to the lay public through books, internet videos and participation in television programs or documentaries. In this regard, their studies may receive media coverage beyond academics. Their students and assistants probably know their hypotheses, and this might be true even for some of those that collaborate as volunteers in their studies. This entire context may help to disseminate certain expectations, thereby extraneously affecting studies’ results. Such expectations could be significantly fostered by “demand characteristics,” a type of response bias derived from participants being actively engaged in an experiment or survey to the point of adopting behaviors or responses that they suppose are demanded of them (Orne 1962).

But there is also a series of other methodological shortcomings beyond experimenter bias and social desirability that may significantly affect studies’ results, from measurement to cultural biases, and to which we must be equally aware. Measures of widespread use are characteristically biased toward certain types of beliefs—e.g., Christian—thus covering a very limited range of supernatural and religious worldviews, which bears important implications for cross-cultural research (King and Crowther 2004). Self-report measures are also known to be vulnerable to a series of methodological biases, including question-order bias (i.e., items that appear first may sometimes influence responses to further questions), and variations due to mood state during questionnaire completion (i.e., transient emotions or impactful events may influence participants’ responses, for example, to well-being measures). These biases may interact in complex ways. For example, poorly designed measures in terms of cultural sensitivity could elicit negative emotions as a result of respondents’ discontentment with the way their beliefs and worldviews are presented. In this scenario, participants’ responses could involve some degree of resistance and distrust, perhaps reflected in underreporting tendencies, missing responses and non-response bias.

## Ways to Reduce Response Bias

Since response biases are not systematically controlled nor usually taken into consideration in mental health research (Rogler et al. 2001), we are not entirely able to evaluate their influence on our methods and results. The first step to reducing them is to acknowledge their possibility. Specific investigations should be conducted to estimate their prevalence

and impact in studies investigating the relationship between religiosity/spirituality and mental health. A host of different methods might be employed for that purpose. One example would be to conduct a meta-analysis of the literature to see how much of these investigations addressed the occurrence of response biases and whether any difference in effect sizes could be attributed to them. Another possibility would be to carry out experimental studies to evaluate the role of different types of bias in participants' responses, comparing results with and without controls for experimenter effects, question-order bias and contextual effects.

Some of the methodological remedies that can be employed to reduce bias include:

1. amelioration of available measures or the development of new ones. This may include the creation of items specifically related to cultural or local beliefs and practices, and exclusion or rewording of socially desirable items;
2. use of self-administered questionnaires, preferably with controls for question-order bias (i.e., question randomization);
3. use of double-blind protocols, whenever possible, as an essential measure to avoid observer or experimenter bias;
4. and eventually deception (for experimental purposes). Deception may take many forms, from experiments in which the purposes are almost entirely hidden from participants to more nuanced variations involving the presentation of the most necessary information and omission of other aspects, such as the original title of a scale or the project's title. The use of deception is acceptable provided compliance with ethical recommendations and standards. These include offering debriefing to participants after the experiment, providing psychological support in case of negative emotional reactions and ensuring the right of participants to withdraw from the study at any time they wish to do so (Cozby and Bates 2014).

Among the available methodological safeguards, we can also mention the use of measures which showed to be more robust against socially desirable responses in comparison with self-report and introspection. For that purpose, a handful of tasks assessing implicit cognition were developed (see Gawronski and Paine 2010 for a review). The most widely used is the Implicit Association Test (IAT), which is based on speeded responses to word association tasks (Greenwald et al. 1998). The IAT was shown to be less prone to faking in comparison with self-report measures (Steffens 2004) and has been successfully employed in a series of investigations in the psychology of religion and spirituality (see Jong 2013 for a review).

Galen (2012a, b) suggested the use of actual behavioral observations instead of self-report, whenever possible or feasible. Podsakoff et al. (2003) recommend using different sources of data to reduce response bias, for example, by contrasting questionnaire responses to archival data and hospital records.

Once data collection is concluded, there are statistical remedies to which the researcher may resort to exert control over response bias during data analysis—providing the potential confounder was previously assessed (see Podsakoff et al. 2003 for a review of such methods). The most widely used is the partial correlation procedure in which the confounder (e.g., social desirability, mood state) is entered as a covariate, and its effects are partialled out from the analyses. The coefficients for zero-order correlations can then be compared to those of partial correlations to estimate the impact of the confounder on the results (reduced/increased coefficients or nonsignificant results).

To the extent that religiosity and social desirability share common variance, Trimble (1997, p. 983) recommended not to use partial correlation to control for the effects of social



desirability since this could “potentially eliminate part of the religious construct of interest.” However, this decision may depend on theoretical preferences, as discussed earlier, and is not necessarily supported by existing research. Byrd et al. (2007), for example, verified that intrinsic religiosity significantly predicted subjective well-being even after controlling for social desirability. The issue of whether desirable responding and religiosity are theoretically indistinct is still debatable, and partialling out the effects of the first is a decision that will vary according to the purposes of the study and definitions of religiosity. In any case, providing information on the strength of the relationship between religiosity and mental health variables before and after controlling for the effects of social desirability could be useful to precise the role of desirable responding in the obtained results.

## Conclusion

The correlations between religiosity and social desirability showed to differ according to the facet of religiosity under investigation (an example is the inverse correlation between religion-as-quest and social desirability). Future investigations will thus benefit from a more detailed assessment of particular aspects of religious and spiritual involvement, including the identification of subgroups of religious or spiritual individuals potentially more prone to biased responding. Even though religiosity and spirituality are profoundly interconnected, some studies have found differences in the way each one impacts mental health (e.g., Granqvist et al. 2014; King et al. 2013). Saucier and Skrzypinska (2006) verified that impression management was positively related to traditional religiousness while spirituality was not, a finding that apparently contradicts some of the studies reviewed above. Therefore, more evidence is needed to ascertain whether spirituality significantly differs from religiosity in its association with social desirability and other types of response bias.

The net impact of desirable responding on the relationship between religiosity and mental health seems to be relatively weak and can be interpreted in ways that do not necessarily involve distorted or biased responding. However, as Koenig (2011) has remarked, “when studying a behavior such as R/S that is important to nearly 200 million people in the USA alone, even small correlations with health may translate into substantial public health importance” (p. 265). In this sense, the correlations between R/S and desirable responding are not much different in strength from those associating R/S with mental health (e.g., Smith et al. 2003). Moreover, as shown by the studies reviewed, response biases are not limited to social desirability, and a series of other methodological shortcomings must be taken into consideration by researchers in their investigation of the mental health implications of religiosity and spirituality. It is hoped that the present paper will help increase awareness to this often neglected but fundamental research topic.

When systematically ignored, the effects of response bias might become cumulative. Although meta-analyses can be successfully employed to provide information regarding published experiments, some meta-analyses may lack studies with specific controls for response bias (Savović et al. 2012). In this regard, van Elk et al. (2015) defend the relevance of registered publications. In a reanalysis of data indicating a prosocial effect of religious priming, the authors verified that, depending on the statistical analyses employed, the results could be ascribed to publication bias. Therefore, the use of meta-analysis alone might be insufficient to establish an effect. The conduction of a large-scale preregistered



replication project would be essential to advance the field of R/S and scientifically establish the mental health benefits of religious and spiritual involvement, especially if controls for the most common methodological biases in this research area are appropriately acknowledged, both in the data collection and in the analysis process.

The impact of response bias on diagnosis and mental health practice is another issue of concern and an important area for future investigations since its manifestations are potentially more dangerous in the therapeutic setting when compared to a controlled research context. This is even more problematic given the role of client and therapist religious values in clinical judgments (Houts and Graham 1986), and the influence of cultural stereotypes on the relation between religion and mental health (Lewis 2001). Special attention should be given to clients' evaluations of improvement and other information concerning their perception of the therapeutic context, including the relationship with the therapist. For example, Reese et al. (2013) found evidence that client's ratings of the therapeutic alliance were significantly influenced by demand characteristics and social desirability. Many are the factors that may influence client disclosure, including therapist's disclosure of religious beliefs (Chesner and Baumeister 1985). Religiosity (be that of the client or the therapist) is thus an important variable to be considered in the evaluation of clients' ratings. Given the complex interrelationship between religiosity, self-perception, social desirability and mental health, response biases, if accurately identified, may shed light on other aspects of an individual's psychological profile beyond faking or impression management.

**Funding** This work was funded by São Paulo Research Foundation, FAPESP (Grant Number #2015/05255-2)

#### **Compliance with Ethical Standards**

**Conflict of interest** The author declares that he has no conflict of interest.

## **References**

- Aguilar-Vafaie, M. E., & Abiari, M. (2007). Coping Response Inventory: Assessing coping among Iranian college students and introductory development of an adapted Iranian Coping Response Inventory (CRI). *Mental Health, Religion & Culture, 10*, 489–513.
- Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology, 5*, 432–443.
- Barber, T. X., & Silver, M. J. (1968). Fact, fiction, and the experimenter bias effect. *Psychological Bulletin Monograph Supplement, 70*, 1–29.
- Batson, C. D., Naife, S. J., & Pate, S. (1978). Social desirability, religious orientation and racial prejudice. *Journal for the Scientific Study of Religion, 17*, 31–41.
- Bergin, A. E. (1983). Religiosity and mental health: A critical re-evaluation and meta-analysis. *Professional Psychology: Research and Practice, 14*, 170–184.
- Bonelli, R. M., & Koenig, H. G. (2013). Mental disorders, religion and spirituality 1990 to 2010: A systematic evidence-based review. *Journal of Religion and Health, 52*(2), 657–673.
- Byrd, K. R., Hageman, A., & Belle Isle, D. (2007). Intrinsic motivation and subjective well-being: The unique contribution of intrinsic religious motivation. *International Journal for the Psychology of Religion, 17*, 141–156.
- Chesner, S. P., & Baumeister, R. F. (1985). Effect of therapist's disclosure of religious beliefs on the intimacy of client self-disclosure. *Journal of Social and Clinical Psychology, 3*, 97–105.
- Council, J. R. (1993). Context effects in personality research. *Current Directions in Psychological Science, 2*(2), 31–34.
- Cozby, P. C., & Bates, S. C. (2014). *Methods in behavioral research* (12th ed.). London: McGraw-Hill Education.
- Crandall, V. C., & Gozali, J. (1969). The social desirability of responses of children of four religious-cultural groups. *Child Development, 40*(3), 751–762.

- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354.
- Durkheim, E. (1912). *The elementary forms of religious life* (translated and with an introduction by Karen E. Fields). New York: The Free Press.
- Ellis, A. (1980). Psychotherapy and atheistic values: A response to A. E. Bergin's, "Psychotherapy and religious values.". *Journal of Consulting and Clinical Psychology, 48*, 635–639.
- Ellis, J. B., & Smith, P. C. (1991). Spiritual well-being, social desirability and reasons for living: Is there a connection? *The International Journal of Social Psychiatry, 37*(1), 57–63.
- Eysenck, M. W. (1998). Personality and the psychology of religion. *Mental Health, Religion and Culture, 1*, 11–19.
- Eysenck, H. J., & Eysenck, S. B. G. (1991). *Manual of the Eysenck personality scales*. London: Hodder & Stoughton.
- Fastame, M. C., Hitchcott, P. K., & Penna, M. P. (2017). Does social desirability influence psychological well-being, perceived physical health and religiosity of Italian elders? A developmental approach. *Aging and Mental Health, 21*(4), 348–353.
- Francis, L. J. (1992). Is psychoticism really a dimension of personality fundamental to religiosity? *Personality and Individual Differences, 13*, 645–652.
- Francis, L. J., Pearson, P. R., & Kay, W. K. (1983). Are religious children bigger liars? *Psychological Reports, 52*, 551–554.
- Freud, S. (1927). The future of an illusion. In *The standard edition of the complete psychological works of Sigmund Freud* (trans & ed. Strachey, Vol. 21). London: Hogarth Press.
- Galen, L. W. (2012a). Does religious belief promote prosociality? A critical examination. *Psychological Reports, 138*(5), 876–906.
- Galen, L. W. (2012b). The complex and elusive nature of religious prosociality: Reply to Myers (2012) and Saroglou (2012). *Psychological Bulletin, 138*, 918–923.
- Galen, L. W. (2016). Commentary on Norenzayan et al.: The cultural evolution of prosocial religions. Big Gods: Extended prosociality or group binding? *Behavioral and Brain Sciences, 39*, 29–30.
- Gawronski, B., & Paine, B. K. (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: The Guilford Press.
- Gelfand, D. M., Gelfand, S., & Rardin, M. W. (1965). Some personality factors associated with placebo responsiveness. *Psychological Reports, 17*, 555–562.
- Gillings, V., & Joseph, S. (1996). Religiosity and social desirability: Impression management and self-deceptive positivity. *Personality and Individual Differences, 21*(6), 1047–1050.
- Granqvist, P., Broberg, A. G., & Hagekull, B. (2014). Attachment, religiousness, and distress among the religious and spiritual: Links between religious syncretism and compensation. *Mental Health, Religion & Culture, 17*(7), 726–740.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480.
- Hadaway, C. K., Marler, P. L., & Chaves, M. (1998). Overreporting church attendance in America: Evidence that demands the same verdict. *American Sociological Review, 63*, 122–130.
- Holman, L., Head, M. L., Lanfear, R., & Jennions, M. D. (2015). Evidence of experimental bias in the life sciences: Why we need blind data recording. *PLoS Biology, 13*(7), e1002190.
- Houts, A. C., & Graham, K. (1986). Can religion make you crazy? Impact of client and therapist religious values on clinical judgments. *Journal of Consulting and Clinical Psychology, 54*(2), 267–271.
- Huguelet, P., & Koenig, H. G. (2007). *Religion and spirituality in psychiatry*. New York: Cambridge University Press.
- Jong, J. (2013). Implicit measures in the experimental psychology of religion. In G. W. Dawes (Eds.), *A new science of religion* (pp. 65–78). London: Routledge.
- King, J. E., & Crowther, M. R. (2004). The measurement of religiosity and spirituality: Examples and issues from psychology. *Journal of Organizational Change Management, 17*(1), 83–101.
- King, M., Marston, L., McManus, S., Brugha, T., Meltzer, H., & Bebbington, P. (2013). Religion, spirituality and mental health: results from a national study of English households. *The British Journal of Psychiatry, 202*, 68–73.
- Klein, O., Doyen, S., Leys, C., Gama, P. A. M. S., Miller, S., & Questienne, L., et al. (2012). Low hopes, high expectations: Expectancy effects and the replicability of behavioral experiments. *Perspectives on Psychological Science, 7*(6), 572–584.
- Knäuper, B., & Hans-Ulrich, W. (1994). Diagnosing major depression in the elderly: Evidence for response bias in standardized diagnostic interviews? *Journal of Psychiatry Research, 28*(2), 147–164.

- Koenig, H. G. (2009). Research on religion, spirituality, and mental health: A review. *Canadian Journal of Psychiatry, 54*(5), 283–291.
- Koenig, H. G. (2011). *Spirituality and health research: Methods, measurement, statistics, and resources*. Pennsylvania: Templeton Press.
- Leak, G. K., & Fish, S. (1989). Religious orientation, impression management, and self-deception: Toward a clarification of the link between religiosity and social desirability. *Journal for the Scientific Study of Religion, 28*(3), 355–359.
- Lewis, C. A. (2000). The religious-psychoticism relationship and the two factors of social desirability: A response to Michael W. Eysenck. *Mental Health, Religion and Culture, 3*(1), 39–45.
- Lewis, C. A. (2001). Cultural stereotype of the effects of religion on mental health. *British Journal of Medical Psychology, 74*, 359–367.
- Lindeman, M., Blomqvist, S., & Takada, M. (2012). Distinguishing spirituality from other constructs: Not a matter of well-being but of belief in supernatural spirits. *Journal of Nervous and Mental Disease, 200*(2), 167–173.
- MacDonald, D. A. (1997). *The development of a comprehensive factor analytically derived measure of spirituality and its relationship to psychological functioning* (doctoral dissertation). Windsor, Ontario, Canada: University of Windsor.
- Migdal, L., & MacDonald, D. A. (2013). Clarifying the relation between spirituality and well-being. *The Journal of Nervous and Mental Disease, 201*(4), 274–280.
- Norenzayan, A., Shariff, A. F., Gervais, W., Willard, A. K., McNamara, R. A., Slingerland, E., et al. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences, 39*(1), 1–65.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*(11), 776–783.
- Pargament, K., & Lomax, J. (2013). Understanding and addressing religion among people with mental illness. *World Psychiatry, 12*(1), 26–32.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego: Academic Press.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). New York: Guilford.
- Pearson, P. R., & Francis, L. J. (1989). The dual nature of the Eysenckian lie scales: Are religious adolescents more truthful? *Personality and Individual Differences, 10*, 1041–1048.
- Peres, M. F. P., Kamei, H. H., Tobo, P. R., & Lucchetti, G. (2017). Mechanisms behind religiosity and spirituality's effect on mental health, quality of life and well-being. *Journal of Religion and Health*. <https://doi.org/10.1007/s10943-017-0400-6>.
- Pfister, O. (1928). The illusion of a future: A friendly disagreement with Prof. Sigmund Freud (edited, with an Introduction note, by Paul Rozen). *International Journal of Psycho-Analysis, 74*(557–579), 1993.
- Plante, T. G., Yancey, S., Sherman, A., & Guertin, M. (2000). The association between strength of religious faith and psychological functioning. *Pastoral Psychology, 48*, 405–412.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.
- Reese, R. J., Gillaspay, J. A., Owen, J. J., Flora, K. L., Cunningham, L. C., Archie, D., et al. (2013). The influence of demand characteristics and social desirability on clients' ratings of the therapeutic alliance. *Journal of Clinical Psychology, 69*(7), 696–709.
- Rogler, L., Mroczek, D., Fellows, M. J., & Loftus, S. T. M. A. (2001). The neglect of response bias in mental health research. *The Journal of Nervous and Mental Disease, 189*(3), 182–187.
- Rosenthal, R. (1976). *Experimenter effects in behavioral research*. New York: Irvington.
- Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3*(6), 176–179.
- Rosenthal, R., & Jacobson, L. (1992). *Pygmalion in the classroom*. New York: Irvington.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences, 1*(3), 377–386.
- Saucier, G., & Skrzypinska, K. (2006). Spiritual but not religious? Evidence for two independent dispositions. *Journal of Personality, 74*, 1257–1292.
- Savovic, J., et al. (2012). Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: Combined analysis of meta-epidemiological studies. *Annals of Internal Medicine, 157*, 429–438.
- Sedikides, C., & Gebauer, J. E. (2010). Religiosity as self-enhancement: A meta-analysis of the relation between socially desirable responding and religiosity. *Personality and Social Psychology Review, 14*(1), 17–36.

- Shamdasani, S. (2003). *Jung and the making of modern psychology: The dream of a science*. Cambridge: Cambridge University Press.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48(11), 1117–1131.
- Sheldrake, R. (1999). How widely is blind assessment used in scientific research? *Alternative Therapies in Health and Medicine*, 5(3), 88–91.
- Smith, T. B., McCullough, M. E., & Poll, J. (2003). Religiousness and depression: Evidence for a main effect and the moderating influence of stressful life events. *Psychological Bulletin*, 129, 614–636.
- Stavrova, O., Fecthenhauer, D., & Schlösser, T. (2013). Why are religious people happy? The effect of the social norm of religiosity across cultures. *Social Science Research*, 42, 90–105.
- Steffens, M. C. (2004). Is the implicit association test immune to faking? *Experimental Psychology*, 5(13), 165–179.
- Trimble, D. E. (1997). The religious orientation scale: Review and meta-analysis of social desirability effects. *Educational and Psychological Measurement*, 57(6), 970–986.
- van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, 25(4), 40–48.
- van der Linden, D., Dunkel, C. S., & Petrides, K. V. (2016). The general factor of personality (GFP) as social effectiveness: Review of the literature. *Personality and Individual Differences*, 101, 98–105.
- van Elk, M., Matzke, D., Gronau, Q. F., Guan, M., Vandekerckhove, J., & Wagenmakers, E.-J. (2015). Meta-analyses are no substitute for registered replications: A skeptical perspective on religious priming. *Frontiers in Psychology*, 6, 1–7.
- VanderWeele, T. J., Balboni, T. A., & Koh, H. K. (2017). Health and spirituality. *JAMA*, 318(6), 519–520.
- Watson, P. J., Morris, R. J., Foster, J. E., & Hood, R. W. (1986). Religiosity and social desirability. *Journal for the Scientific Study of Religion*, 25(2), 215–232.
- Watt, C., & Nagtegaal, M. (2004). Reporting of blind methods: An interdisciplinary survey. *Journal of the Society for Psychical Research*, 68, 105–114.
- Zimbauer, B. J., & Pargament, K. I. (2005). Religiousness and spirituality. In R. F. Paloutzian (Eds.), *Handbook of the psychology of religion and spirituality*. New York: The Guilford Press.
- Zuckerman, P. (2009). Atheism, secularity and well-being: How the findings of social science counter negative stereotypes and assumptions. *Sociology Compass*, 10, 949–971.
- Zuckerman, P., Galen, L. W., & Pasquale, F. L. (2016). *The nonreligious: Understanding secular people and societies*. Oxford: Oxford University Press.