healthcare

MDPI

Review

Reporting and Interpreting Effect Sizes in Applied Health-Related Settings: The Case of Spirituality and Substance Abuse

Iván Sánchez-Iglesias [1], Jesús Saiz [2], Antonio J. Molina [2,*] and Tamara L. Goldsby [3]

[1] Department of Psychobiology & Behavioral Sciences Methods, Complutense University of Madrid, 28223 Madrid, Spain; i.sanchez@psi.ucm.es (I.S.-I)
[2] Department of Social, Work and Differential Psychology, Complutense University of Madrid, 28223 Madrid, Spain; jesus.saiz@psi.ucm.es (J.S.)
[3] Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA 92093, USA; tgoldsby@health.ucsd.edu (T.G.)
* Correspondence: antmolin@ucm.es (A.J.M.)

Abstract: Inferential analysis using null hypothesis significance testing (NHST) allows accepting or rejecting a null hypothesis. Nevertheless, rejecting a null hypothesis and concluding there is a statistical effect does not provide a clue as to its practical relevance or magnitude. This process is key to assessing the effect size (ES) of significant results, be it using context (comparing the magnitude of the effect to similar studies or day-to-day effects) or statistical estimators, which also should be sufficiently interpreted. This is especially true in clinical settings, where decision-making affects patients' lives. We carried out a systematic review for the years 2015 to 2020 utilizing Scopus, PubMed, and various ProQuest databases, searching for empirical research articles with inferential results linking spirituality to substance abuse outcomes. Out of nineteen studies selected, eleven (57.9%) reported no ES index, and nine (47.4%) reported no interpretation of the magnitude or relevance of their findings. The results of this review, although limited to the area of substance abuse and spiritual interventions, are a cautionary tale for other research topics. Gauging and interpreting effect sizes contributes to better understanding of the subject under scrutiny in any discipline.

Keywords: Effect size; scientific writing; substance abuse; spiritually-based treatment

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Healthcare* **2022**, *10*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Scientific research in psychology relates to many other disciplines such as epidemiology, biology, and medicine, among others. The research endeavor seeks to gain knowledge of human behavior in all of its aspects, from observable behavior to cognition, through personality traits, beliefs, attitudes, and many other systems and processes related to psychological or physical health. When psychological research addresses issues as prevalent as substance abuse, it becomes a public health issue. As many other scientific disciplines, psychological research also seeks to describe, predict, and explain phenomena. The instruments to do so include a proper and thorough design and adequate data analysis to answer the proposed research question. Although there are certain alternative approaches to data analysis such as Bayesian analyses, the most frequent strategy for inference is null hypothesis significance testing (NHST).

NHST is the key procedure in frequentist inferential statistics, while its use remains a subject of debate and controversy. Many of the criticisms [1-3] may be said to be based on misuse by researchers authoring studies and/or poor understanding on the part of both authors and readers [4-6]. The use of p-values is ubiquitous. Based upon the p-values, categorical, dichotomous judgments may be made regarding the so-called null-hypothesis in terms of accepting or rejecting it. This in turn gives rise to a "significant" vs. "nonsignificant" results determination. Too often, that is the end of the road in a given study and the authors draw conclusions on a substantive and complex issue from that p-value only. Usually, once an effect has been found, no attention is paid to the magnitude of that effect. Authors are just beginning to recommend the calculation and interpretation of

the magnitude of an effect (Effect Size, ES) as part as what they refer to as "the new-statistics movement" [1,7,8]. However, as we shall see, the use of ES has been studied, discussed, and recommended as standard practice for decades now; at least, as far as we know, since 1969 [9].

1.1. Beyond the Null Hypothesis Significance Testing

Reporting ES does not replace the purpose of NHST (i.e., whether an observed effect is statistically significant or not), but supplies additional information regarding the magnitude of a significant observed effect (i.e., "How large an effect do I expect exists in the population?" [10]). The practical relevance of a given significant effect is better assessed by comparing it with reasonable criteria (well-stated effects in similar research settings, or everyday and well-known effects). Authors may call upon previous research in similar (if not the same) settings in order to compare ESs. If there are no contextual benchmarks, arbitrary but published thresholds are available for reference, even when they are not the best option. Labels such as "small", "medium", or "large", may be misleading or simply uninformative as an ES estimator. It would seem that researchers use such labels, more often than not, to interpret ES indices as ubiquitous as Cohen's *d*. However Cohen himself, in 1988, believed that the convention he was proposing would be found to be "reasonable by reasonable people" [10] (p. 13), and warned about the dangers of the use and abuse of arbitrary labels.

Contrasting ES across different populations also assists in gaining knowledge in generalization strategies and identifying potential confounds affecting internal validity in any study. As Shadish et al. stated in 2002, "Demonstrating effect size variation across operations presumed to represent the same cause or effect can enhance external validity by showing that more constructs and causal relationships are involved than was originally envisaged; and in that case, it can eventually increase construct validity by preventing any mislabeling of the cause or effect inherent in the original choice of measures…" [11] (pp. 470-71).

ES may be interpreted using descriptive statistics only (that is, after a result has been deemed as statistically significant and the sample statistics are to be interpreted). When variables are measured in units with intrinsic value (such as height or weight in standardized units) or contextual meaning (such as salaries in dollars), readers may make rapid assessments based on their previous experiences. The American Psychological Association (APA) [12] recommends including measures of effect size in the manuscripts and has been doing so since, at least, 2010. The APA mentions that ES expressed in original units allows for an easier interpretation ("e.g., mean number of questions answered correctly, kilograms per month for a regression slope") [13] (p. 89), but focuses primarily on statistical estimates.

There are entire courses devoted to statistics in social sciences degrees. Descriptive and inferential statistics, psychometry, research methods, and epidemiology are subjects known to be taught in most (if not all) of those degrees and ES indices are included in their syllabi. Additionally, there are numerous published papers that address this topic and offer recommendations regarding ES in several psychological research areas [14-18]. It is not clear whether the recommendations have been fully adopted over time by students, researchers, and reviewers alike. In fact, the misreporting (or the lack of reporting) of ES remains an issue in scientific writing regarding several scientific disciplines and health-related settings. In a systematic review of randomized controlled trials (RCT) published since the year 2000, Martín-Aguilar and Sánchez-Iglesias [19] found that 8 out of 10 statistically significant studies (80%) failed to report ES statistics to assess the magnitude of the effect of pharmacological treatments on depressive symptoms; 1 reported ES statistics but did not interpret it; and only 1 reported and interpreted the ES in its context. In a similar review, Elvira-Flores and Sánchez-Iglesias [20] analyzed 21 experimental studies, 11 (52.4%) of which did not report ES statistics; 5 of which (23.8%) reported statistics but did not interpret them, and the remaining 5 studies (23.8%) reported and interpreted the ES values using the arbitrary thresholds proposed by Cohen [10] but without providing a contextual meaning.

1.2. Inferential Statistics Without Effect Size Estimators and Questionable Research Practices

Failing to report ES indices, or doing so but without discussing them, may be regarded as questionable research practices. Some reasons may include lack of training

in statistical procedures, the rush for publishing imposed by competitive academic environments [21,22], a misunderstanding of the meaning and usefulness of ES, or a willingness to conceal observed poor effects. These practices may be found during statistical analyses, as in the case of ES calculation (or the lack of it) or other inadequate procedures (variable slicing, cherry picking, p-hacking, etc.). However, questionable practices may also occur before or after research [23], such as failing to publish non-significant results [24] or using tendentious causal language [25]. These questionable research practices pose a threat to the credibility of scientific research [26,27].

We assume that, in most cases, the lack of ES reporting is unintentional. One may wonder whether these studies (and their results) may be considered "wrong". We do not think so. However, even if the study was appropriately designed and reliable research methods were utilized, an argument may be made that they are not entirely complete. Readers will not have enough information to make more than educated guesses regarding the substantive relevance of the findings. Assuredly, readers may make their own assumptions regarding effect sizes and their interpretation. Assuming the relevant data (descriptive and inferential statistics) are reported, they may calculate ES indices and then interpret them. However, should it not be the authors who are the first to introduce and discuss the practical relevance of their own findings?

1.3. The Role of Effect Size in Spirituality, Religion, and Substance Abuse Studies

Psychological and social factors play a part in health problems. Religiousness is considered a key variable in health improvement [28-30], and. researchers have studied spirituality and religiosity as relevant variables with regard to public health [31-33].

King and Koenig defined religion as "an organized system of beliefs, practices, rituals, and symbols designed a) to facilitate closeness to the sacred or transcendent (God, higher power, or ultimate truth/reality) and b) to foster an understanding of one's relationship and responsibility to others in living together in a community" [34] (p. 2). From Western to Eastern beliefs, varied conceptions exist regarding the definition of a transcendent greater power. At the same time, the term spirituality refers to a broader concept that encompasses everything from a term that may be used to identify deeply religious people [32], to a characteristic of individuals who are only superficially religious, religion seekers, a well-being concept or even secular individuals [35] (Koenig, 2008).

We focus the present study on substance abuse disorders. Recovery from other health and behavioral issues, such as gambling disorders (among other disorders), has been studied in regard to spiritual beliefs and practices [36,37]. Although they are occasionally classified as addictions, these issues are not directly related to substance usage and will not be addressed in the present study. The DSM-5 [38] recognizes substance use disorders as a pattern of troublesome symptoms resulting from substance use. The DSM-5 lists eleven criteria including: Using more drugs than one should; failing to reduce consumption; spending a great deal of time on substance-related activities; cravings for the substance; struggling with daily tasks or giving up other activities as a result; using the drug despite problems (psychological, relationship-related, or physically endangering); and developing withdrawal and tolerance symptoms. The list covers a broad spectrum of substances, including alcohol, tobacco, caffeine, cannabis, hallucinogens, opioids, anxiolytics, cocaine, stimulants in general, and even other unidentified substances.

Addiction intervention programs are fundamentally divided into *harm reduction* and *recovery-based* programs. Harm reduction programs aim to minimize the main negative consequences of drug addiction, especially the consequences of associated infections and criminal behaviors related to substance use [39], while recovery is a concept used to contextualize a process of treatment and subsequent social reintegration [40]. *Recovery* is occasionally used interchangeably with *rehabilitation*, however, there are differences. The goal of *rehabilitation* is to assist individuals with a handicap or difficulty (such as a physical problem, addiction and/or psychiatric problem) to reintegrate the individual into the community [41]. However, *recovery* involves the development of personal autonomy, the performance of socially valuable roles, maintaining significant socio-affective relationships and existence with the symptoms that allow socio-community integration and a relatively satisfactory life [42]. *Recovery* implies not only

reducing or eliminating drug use, as could even be achieved by spontaneous remission [43], (referred to in some cases as "natural recovery" [44]), but becoming an active member of society [40]. The term *recovery capital* refers to the connections between personal and social networks, competencies, reciprocal norms, and the capacity for trust and bonding generated between the individual in treatment and their reference groups [45-47], in the framework of a "science of recovery" [48].

Harm reduction programs, recovery (or therapeutic) community programs, and psychosocial integration programs are currently included in treatment networks [49]. Therapeutic communities have traditionally been associated with recovery programs. Therefore, these programs not only focus on the presence or absence of substances. Today, therapeutic communities feature empowerment, peer support, active participation, and social support [50].

An individual approach (with cognitive behavioral therapy) that incorporates relapse prevention is typically the basis of health system treatment interventions. It may be challenging to incorporate other service types such as psychological support, self-help groups, peer support / social support groups, other supporting programs, or programs oriented to minorities into treatment networks.

Spirituality is another aspect that is frequently excluded from treatment, but has received increased attention with regard to its role in the maintenance of recovery from alcoholism. Spirituality has been defined as that which gives meaning and purpose in life [51] as well as a sense of personal identity and transcendence that motivates individuals beyond the practicalities of daily living [52].

Recovery interventions such as the Twelve-Step programs of Alcoholics Anonymous, advocate acceptance of a "higher power," promote spiritual awakening, and use prayer and meditation as instruments for recovery and healing [53]. In this context, spirituality has been linked to betterment in certain health outcomes, including state anxiety in alcohol recovery [54] and relapse avoidance [55].

1.4. Objective

Using several databases, the present authors carried out a systematic review to obtain a non-biased sample of studies with inferential outcomes that linked spirituality or religion to substance abuse. The selected studies were analyzed to determine the number of studies that utilized ES estimators and/or interpreted the magnitude of their findings.

2. Methods

The systematic review procedure utilized in the present study was the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines by Page et al. in 2021 [56].

2.1. Eligibility Criteria

In order to be included in the systematic review, the studies needed to be published between 2015 and 2020, in Spanish or English, in a peer-reviewed scientific journal. The studies could use any methodology (experimental or not).

The target population was people who had a problem of substance abuse (any substance). For observational studies, the substance-related problem could have appeared at any time prior to the measurement of the variables.

Also, the studies could be observational (assessing the relationship between spirituality and outcomes related to substance abuse) or include treatments, programs, or interventions based on spirituality or religion.

The studes had to present at least one significant outcome measure assessing the relationship between a relevant variable and a change in the abuse behavior, relapse prevention, or a theoretically related variable.

We excluded studies without significant outcomes, solely qualitative methodologies, or case reports.

2.2. Information Sources

The present authors carried out a systematic literature search, searching for relevant studies. The following ProQuest databases were utilized: PsycINFO, the Applied Social Sciences Index & Abstracts [ASSIA], Sociological Abstracts, and Sociology Database (the latter three are included in the Sociology Collection), PubMed, and Scopus, for the period 2015 to 2020.

2.3. Search Strategy

The same search terms were entered in each selected database, in English and Spanish, using the following Boolean expression: "(addiction OR "substance abuse") AND (spirituality OR spiritual) AND (relapse OR treatment)", adapting the syntax to the specific rules of each database engine. The search was restricted by title, abstract, and keywords. The present authors also restricted the search to peer-reviewed papers published in scientific journals, excluding theses and dissertations, chapters, books, and gray literature items. The publication date was also restricted in the database, allowing registers from 2015 to 2020, both inclusive. The specific sequences of terms used for the ProQuest databases can be found in Appendix A.

2.4. Selection Process

In order to identify and remove duplicate records, we entered the data from the previous stage into a single Excel spreadsheet. To determine whether a record was suitable for retrieval and reading, two reviewers independently evaluated each record's title and abstract. The final judgment was made with the assistance of a third researcher when appropriate. Disagreements among the reviewers were settled by consensus.

2.5. Data Collection Process

The present authors attempted to retrieve all eligible records. Two reviewers independently read these reports to determine final inclusion and data extraction.

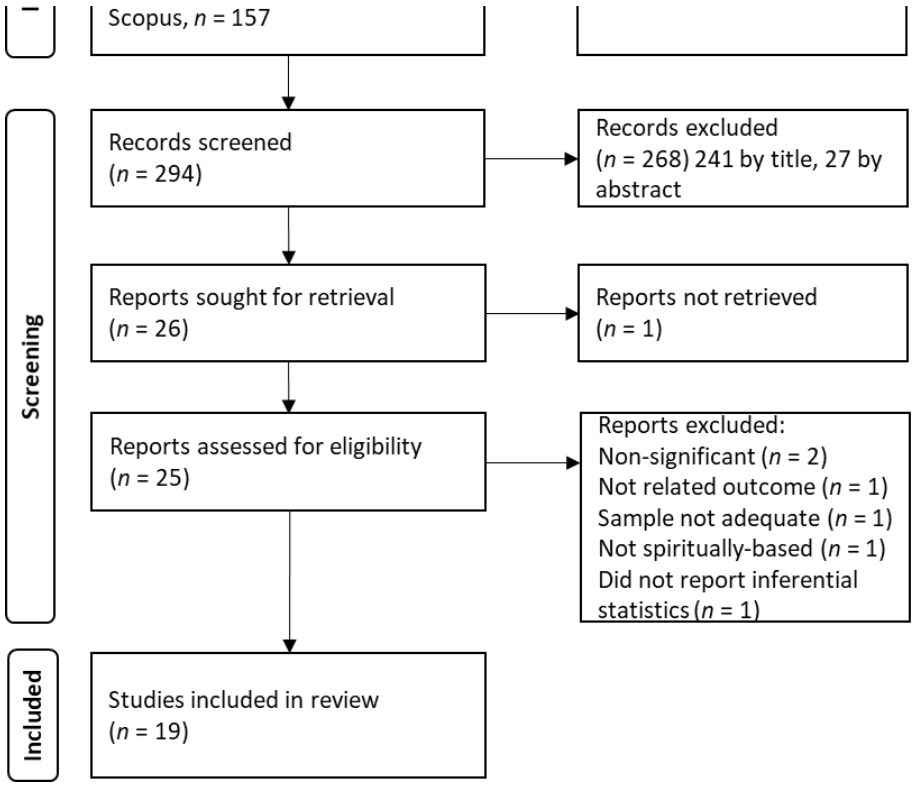
2.6. Data Items / Assessment of Effect Size Estimators and their Interpretation

Each reviewer, on their own, searched for and extracted the methodology, statistical analysis techniques, ES estimators, and ES interpretations for each selected study. The ES estimators (contextual or statistical) were sought in the results section of each study. The reviewers also looked for effect size interpretations of the significant findings, in both the results and discussion section of each report. The studies were classified according to their methodology, main data analysis techniques, ES indices reported (explicitly as ES estimators or not), and the interpretation of the magnitude of the significant effects observed (again, explicitly reported as such or not). Disagreements were settled by consensus and with the aid of a third researcher, as in the previous step.

3. Results

3.1. Study Selection

We identified a total of 477 studies, and 294 non-duplicate records were screened. We excluded 268 records (241 by title and 27 by abstract); 26 were sought for retrieval and evaluated for eligibility. Of those, seven articles were excluded for the following reasons: The outcomes of two studies were non-significant, so ES was not necessary [57,58]; the outcome of one study was not related to change in substance abuse or improvement of relapse prevention [59]; the sample was not comprised of participants with a problem of substance abuse [60]; in another study, the intervention was not spiritually-based [61]; one did not report inferential statistics [62]; and one study could not be retrieved for full text [63]. Finally, 19 studies were included in the review. The flowchart of the search and selection of studies is displayed in Figure 1.



265

266 **Figure 1.** Flow Diagram of the Search and Selection of Studies.

267 3.2. Study Characteristics

268 Table 1 shows the key characteristics of the selected studies. In summary, we
269 identified the following designs in the 19 studies selected: cross-sectional, six studies
270 (31.6%); longitudinal, five studies (26.3%); pre-experimental (one-group pretest-posttest
271 design), three studies (15.8%) and one three static, non-equivalent groups design (5.3%).
272 The remaining four studies used experimental or quasi-experimental design (21.1%). 12
273 studies (63.2%) included some kind of spiritually-related intervention.

274 The following is a summary of the selected articles.

275 Abdollahi and Talib [64], in a cross-sectional study, examined the associations of
276 several variables using a moderation test and structural equations modeling. The
277 authors argued that spirituality and hardiness played a protective role against suicidal
278 ideation (an abuse related outcome) in a population with substance abuse referred to
279 addiction treatment centers. They used the percentage of variance accounted for as an ES
280 index with no benchmark or contextual comparison, stating that "hardiness and
281 spirituality explain 46.0% of the variance in suicidal ideation. These findings indicate
282 that hardiness and spirituality are valuable predictors of suicidal ideation (p. 17)".

283 Andó et al. [54] used path analysis to study the mediation effect of spirituality
284 between anxiety and depressive symptoms and alcohol recovery, in a three-static, non-
285 equivalent groups (three distinct alcohol treatment settings) design. They concluded that
286 there is a beneficial effect of spirituality, decreasing state anxiety, when attending long
287 term Twelve step-based interventions, such as the ones provided by Alcoholics
288 Anonymous. This study did not quantify the ES of these long-term interventions.

289 Beckstead et al. [65] used a one-group pretest-posttest design to assess the change of
290 young patients in a substance use treatment center, after the incorporation of Dialectical
291 Behavior Therapy, a spiritually-related treatment. They used a paired T-test to assess

change, and Cohen's d and its arbitrary benchmarks [66] to estimate the ES, stating that "The effect size of treatment, using Cohen's d was 1.315, a large effect by Cohen's standards" (p. 86). They also used arbitrary benchmarks to assess the ES of the percentage of change (clinically significant and reliable change on the YOQ-SR, a questionnaire designed to assess perceived functioning and distress). The authors reported that "...the clinical significance of change was substantial within individuals over time, with 96% of the youth either recovering or improving at the time of discharge (according to Jacobson & Truax, 1991 criteria)" [65] (p. 86).

Crutchfield and Güss [67] designed a cross-sectional study examining a link between successful long-term substance abuse recovery and goal-oriented, educational or vocational achievements. Their data analyses included T-test, Pearson's correlation, and hierarchical linear regression. There was neither explicit ES estimator for correlation values nor contextual interpretation for R^2 in the regression models. However, they used η2 as an ES estimator, using expressions such as "The magnitude of the differences in the means […] was large (eta squared = .12)" (p. 10.). Moreover, they reported descriptive statistics and used them to express ES as a ratio in a meaningful metric, stating "…This equates to roughly 10 years clean for those who said yes versus 5 years to those who said no" (p. 10.).

Dickerson et al. [68], in a cross-sectional study with adults seeking substance use treatment examined the relationship between several measures: demographic, mental health, physical health, cognitive functioning, cultural identity and spiritual involvement, and substance use-related variables. The authors found that higher frequency in traditional, spiritual traditional practice correlated with lower depression and with lower generalized anxiety disorder scores. They used correlation analyses and reported p-values without r-values.

Kelly and Eddie [69] used a cross-sectional, representative sample of adults who had had a problem with alcohol or drugs. They examined differences in spiritual and religious identification across groups, and whether those differences related to alcohol and other drug abuse recovery. Through chi square analyses and post hoc tests, they found that spirituality (but not religiousness) related to recovery, but with some notable differences by ethnicity and gender. No ES estimators were calculated or discussed.

Kerlin [70] found, in a one-group pretest-posttest design, a statistically significant decrease in self-reported health symptoms and therapeutic improvement as a result of a spiritually integrated treatment program for substance use disorder. She conducted multiple paired-samples t-tests in a sample of 30 women. However, the author did not report ES estimators, so the magnitude of the change could not be quantified.

Lashley [71] used a time series design to assess the impact of stay in a faith-based, addiction recovery program for homeless residents. She used paired t-tests to assess change and ANOVAs to compare differences based on demographic variables. The author found improvements in self-esteem, depressive symptomatology and physical activity levels at follow-up periods after admission. No ES estimator was calculated, although some descriptive differences between groups were highlighted when reported in units with a contextual meaning. For example, the author stated "On average, men reporting other religious affiliations having 54 fewer days in the program than men affiliated with the Christian religion ($p < .05$)." and "On average, men who had not used recovery resources in the past stayed nearly 67 days longer than men who had utilized past recovery resources…".

Lee et al. [72] performed a longitudinal study with youths diagnosed with substance dependency (alcohol and other drugs) in residential treatment with Twelve-step programs. They argued that this treatment played a role in promoting change. However, change was not quantified; no ES estimators were reported, although the authors used many statistical tools: Fisher's exact test, Kruskal-Wallis chi-squared test, Proportional hazard regression, Binomial logistical regression, and Random effects regression.

Mallik et al. [73] designed a quasi-experimental study, in which they compared the effect of spiritually-based meditation with relaxation and with standard treatment, on substance abstinence, psychological distress, and psychological dysfunction. They concluded that the spiritually-based approach might add further support to substance use disorder patients. They used several statistical tests: ANOVA, chi-square test, logistic regression, ANCOVA, and moderation analysis. For the logistic regression, they

used the odds ratio as an ES estimator, and interpreted it in terms of likelihood ratio, e.g.: "...participants in the Meditation condition were 22 times more likely to maintain abstinence than participants in the Relaxation condition and 15 times more likely to maintain abstinence than participants in the TAU condition" (p. 61).

Medlock et al. [74] examined adult patients requiring medical detoxification for severe substance use disorders in another cross-sectional study. The researchers used bivariate analyses via Pearson´s correlation, and multivariate linear regression models. In the former, no explicit ES measures (such as R^2) were reported. They concluded that positive religious coping was negatively associated with days of substance use and positively associated with use of mutual help. Furthermore, they associated religious coping with "…very modestly, yet statistically significantly lower craving" (p. 747), providing a clue regarding that specific ES. In the linear regression models, the change in R^2 when including new variables to the model was mentioned, but not interpreted.

Montes and Tonigan [75] administered measures longitudinally to a sample from community-based Alcoholics Anonymous (AA) and outpatient treatment programs. They examined spiritual and religious (S/R) practices as a mediator of the relationship between AA attendance and reductions in drinking behavior: They found this mediation effect (via mediation and moderated-mediation models) and concluded that some S/R practices should be fostered in order to positively change the drinking behavior. No ES index was calculated, however, some of the findings were placed into context, stating "...the magnitude of the prognostic effect of gains in S/R practices on later increases in alcohol abstinence observed in the current study fell within the range explicated in a reported by Tonigan (2015, in Magill et al., 2015)" (p. 8).

Ranes et al. [76], designed a longitudinal study with participants recruited from a Twleve Step-based residential program. Researchers asked participants to complete multiple instruments at baseline, end of treatment, and three follow-up measures. They utilized repeated measures ANCOVA to assess changes in level of spirituality over time, while controlling for the effects of several variables. They also used multiple linear regression to evaluate predictive models, using R^2 as ES estimator, but without further interpretation. The authors reported data plots and provided their opinion regarding the magnitude of the observed increment. For instance, they stated "Data plots also demonstrated that spirituality increased throughout the duration of the study for all participants, with a large increase between baseline and the end of treatment" and "Participants with low baseline religiousness […] experienced a fairly large increase in spirituality during the first month following treatment" (p. 11).

Ransome et al. [77] studied religious involvement and race differences in opioid use disorder risk and found that religious involvement may be important for prevention and treatment practices. They utilized bivariate logistic regression to estimate the lifetime risk of opioid use disorder and data plots for visual interpretation of certain results; no explicit ES estimators were calculated nor interpreted.

Shorey et al. [78] considered mindfulness-based interventions promising as an effective intervention for improving substance use disorder and associated depressive symptoms. Using correlation and hierarchical regression analyses in a cross-sectional study, the researchers found that dispositional mindfulness and spirituality were negatively associated with depressive symptoms. They reported R^2 and change in R^2 without further interpretation.

Temme and Kopak [79] recruited participants from an inpatient residential therapeutic community. In an experimental design, they randomized the sample into an intervention group and a treatment as usual group. Using path analysis, the authors tested the model of relationships between mindfulness, spirituality (as a mediator), and warning signs of relapse. They did not report or interpret any ES estimator.

Tianingrum et al. [80] designed a one-group pretest-posttest study and concluded that a Narcotics Anonymous style intervention and rehabilitation may improve relapse prevention among prisoners with substance abuse problems. The authors used ANOVA and correlation analyses; however, they did not use ES indices, nor did they interpret the magnitude of their findings.

Yaghubi et al. [81] randomly assigned a sample of patients into two groups to evaluate the efficacy of a religious-spiritual group therapy on the spiritual well-being and the quality of life in methadone-treated patients, versus a no-treatment group. The

authors found a significant increase in spiritual well-being for the experimental group using ANOVA, but they did not quantify the magnitude of the effect.

Yeterian et al. [82] studied religiosity and spirituality as predictors of cannabis use and heavy drinking, recruiting a sample of adolescents in outpatient treatment. The researchers randomly assigned the sample to a Twelve-Step Facilitation treatment group or to a Motivational/Cognitive-Behavioral Therapy group. Data were analyzed via correlation, hierarchical multiple linear regression, and logistic regression. ES were reported for linear regression (change in R2), but not interpreted. For logistic regression, the authors interpreted ES using the odd ratio, in terms of increase or decrease of likelihood of a behavior; for instance "For each 1-point decrease on the STS at baseline, individuals were 3.34 times more likely to report HDD [heavy drinking day] at follow-up (i.e., 1/OR = 1/.299 = 3.34)." (p. 6).

Altogether, in the 19 studies selected, the authors reported results from 42 main statistical techniques. The analyses were quite varied. They included descriptive statistics and graphical plots, chi-square and Fisher's exact tests, independent and paired t-tests, ANOVAs and Kruskal-Wallis tests, ANCOVA, Pearson's correlations, regression models (linear, hierarchical, logistic, random effects, and hazard regression), path analysis, moderation tests, and structural equation models. From these 42 techniques, 29 (the 69.0%) did not include any ES index. The 13 ES estimators reported were: R^2 (and / or change in R^2), six times; odds ratio, two times; ratio expressed in a contextual frame, one time; *r*-value, one time; eta squared (η^2), one time; Cohen's *d*, one time; percentage of change in a test scoring, one time. Out of 19 studies, 11 (the 57.9%) did not report any ES index at all.

ES interpretations were found in 12 occasions. Three indices were interpreted using arbitrary benchmarks (for Cohen's *d*, η^2, and Jacobson & Truax criteria [66]). Two ES indices (both from a single study) were interpreted as mean differences in a natural context (days). Two ES indices (both odds ratio) were interpreted as likelihood ratio. Another ES was interpreted as a ratio of years between two distinct groups. In one study, the authors did not report ES indices, but they put the results into context, comparing them to a similar study by other researchers. One ES estimator, expressed as "percentage of variance accounted for" was interpreted arbitrarily, without benchmarks or contextual framing. In the last two studies in which the magnitude of an effect was addressed, the authors did not report ES indices and they used subjective judgments or opinions. Out of 19 studies, 9 (47.4%) did not report any interpretation of the magnitude or relevance of their findings at all.

Table 1. Design, Main Statistical Analyses and Effect Size (ES) Interpretations in the Studies Selected.

Citation	Methodology	Statistical analysis	ES	ES interpretation
Abdollahi and Talib (2015)	Cross-sectional	Structural Model	Variance accounted for (%)	Arbitrary, no benchmark or context.
		Moderation Test via SEM	No	-
Andó et al. (2016)	Three static, non-equivalent groups design *	Path analysis	No	-
Beckstead et al. (2015)	Pre-experimental (one-group pretest-posttest design) *	T-test	Cohen's *d*	Arbitrary benchmarks
		Descriptive statistics	% of clinically significant change	Arbitrary benchmarks
Crutchfield and Güss (2018)	Cross-sectional	T-test	η^2	Arbitrary benchmarks
		Descriptive statistics	Ratio	Natural context
		Pearson's correlation	*r*	-

		Hierarchical linear regression	R^2	-
Dickerson et al. (2021)	Cross-sectional	Correlation (w/o *r* value, only p-value)	No	-
Kelly and Eddie (2020)	Cross-sectional	Chi-square analyses, *post hoc* tests	No	-
Kerlin (2017)	Pre-experimental (one-group pretest-posttest design) *	Paired and independent T-tests	No	-
Lashley (2018)	Longitudinal *	Paired T-tests	No	Difference in mean (days)
		Correlation (w/o *r* value, only figures)	No	-
		ANOVA	No	Difference in mean (days)
Lee et al. (2017)	Longitudinal *	Fisher's exact test	No	-
		Kruskal-Wallis chi-squared test	No	-
		Proportional hazard regression	No	-
		Binomial logistical regression	No	-
		Random effects regression	No	-
Mallik et al. (2019)	Quasi-experimental *	ANOVA	No	-
		Chi-square test	No	-
		Logistic regression	Odds ratio	Likelihood ratio
		ANCOVA	No	-
		Moderation analysis	No	-
Medlock et al. (2017)	Cross-sectional	Correlation	No	Subjective judgment
		Multivariable linear regression	ΔR^2	-
Montes and Tonigan (2017)	Longitudinal *	Mediation and moderated-mediation	No	Context (similar studies)
Ranes et al. (2016)	Longitudinal *	ANCOVA	No	-
		Multiple linear regression	R^2	-
		Data plots	No	Subjective judgment
Ransome et al. (2019)	Longitudinal	Logistic regression	No	-
		Data plots	No	-
Shorey et al. (2015)	Cross-sectional	Correlation	No	-
		Hierarchical linear regression	R^2 and ΔR^2	-
Temme and Kopak (2016)	Experimental *	Path analysis	No	-
Tianingrum et al. (2018)	Pre-experimental (one-group pretest-posttest design) *	ANOVA	No	-
		Correlation	No	-
Yaghubi et al. (2019)	Experimental *	ANOVA	No	-
Yeterian et al. (2018)	Experimental *	Correlation	No	-

	Hierarchical linear regression	ΔR^2	-
	Logistic regression	Odds ratio	Likelihood ratio

Note. SEM: Structural equation modeling.

* This design comprised a spiritually-based intervention.

448
449
450
451
452
453

4. Discussion

454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498

This paper addresses the importance of properly gauging the magnitude of effects inferred via Null hypothesis significance testing (NHST). The present authors chose an unbiased sample of articles, thanks to a systematic review, to illustrate the need for a better reporting of ESs in the applied field of substance abuse disorder interventions. Using the dichotomous decision method of NHST, we can test whether empirical data conform to a null model (as suggested by Fisher in 1925 [83]) or to an alternative one (as proposed in 1928 by Neyman and Pearson [84]). Both proposals were combined into the ubiquitous NHST. However, NHST was never intended for inferring clinical significance from statistical significance. Since then, multiple effect size (ES) indices have been proposed and widely used. In addition, authors are encouraged to interpret ES into a contextual framework. Despite recommendations for estimating and interpreting ES, as in other research fields [19,20], the authors of the present paper found that studies of spiritual treatments in substance abuse patients rarely report any statistical index or any other type of estimator. Interpretations of the estimators are also infrequent and when they do occur, they are mostly arbitrary thresholds, using the "small", "medium", and "large" labels. Contextual references are very rare. As reported by other authors [85], we also found instances of interpretations left to the subjective judgment of the author. In the present study, approximately half of the selected studies did not report any ES index and roughly the same number did not interpret the magnitude or relevance of their findings either.

The present authors believe it would be useful to revisit and validate the relevance of recovery-based programs [86]. It is important to develop theoretical models and useful interventions based on scientific evidence, with data gathered in applied studies with people who have problems with addictive behaviors [87]. However, the practical relevance of a given intervention cannot be assessed, even if it yielded significant differences with a control group or a treatment-as-usual group, without gauging the magnitude of the differences (i.e., estimating and interpreting the ES). It is through the effect size that we will know whether, in the applied context of the research, the intervention is worthwhile. Small ESs may be relevant if they can be obtained with short, simple, inexpensive intervention programs, or a combination of the above. Large ESs may determine that even the most expensive and complex intervention programs will be implemented. These are decisions that need to be made by hospital, institutional or government managers. It is up to us, the researchers, to calculate and provide clear and accurate indications of the ES. It is we, from academia and applied research, who have the duty to report adequately on this fundamental aspect. Statistical indicators should not be a straitjacket for interpreting effect sizes, using strict thresholds and benchmarks with arbitrary meanings. Even after calculating indices such as Cohen's d or r^2, researchers need to interpret them in the framework of the actual context of the research. For instance, the standardized differences obtained from intervention and control groups can be compared between similar studies A and B. Is one of the interventions relatively better than the other? This comparison would be even better if, instead of the d indices of each study, we compared the scores on an interpretable metric. For example, mean difference (between experimental and control group) of days elapsed without relapse in study A compared to that in study B. As we have found in this review, ratios

[67], odd ratios and likelihood ratios [73,82] are also statistics susceptible to straightforward contextual interpretation.

All this is especially true in clinical settings, where decision-making affects patients' lives. The choice of an effective intervention is of paramount importance in substance abuse programs. The literature presents promising data on the inclusion of spirituality in recovery-oriented programs when it comes to treatment, relapse prevention, and social integration (particularly those emphasizing social support and recovery capital, participatory activity, and a biopsychosocial perspective) (49,50). Any comprehensive network for treating addictive behaviors should contain programs based on previously verified data, and the ES is essential to gauge their usefulness.

4.1. Limitations

This study is not without limitations. Firstly, the present study assessed the methodological rigor when reporting and interpreting ES for a very specific setting, spiritual-based interventions and programs, and their effect on recovery in substance abuse. Therefore, the search terms used were limited; the search may have been conducted with a more comprehensive search equation. Secondly, our search led the present authors to a relatively small number of studies (n = 19) that we considered suitable according to the current study's proposed inclusion and exclusion criteria. To acquire a larger sample of publications, an option might have been to conduct the search using various criteria, such as more databases, a wider range of publication dates, synonymous search phrases, etc. Moreover, one option might have been to search for articles on the impact of spiritual therapies on various aspects of health to gain access to a broader sample. Nevertheless, the primary goal was to discuss the importance of having a measure of the magnitude of the effects found in spiritual treatments of substance abuse. A non-biased selection of articles was obtained through the systematic review process. This non-biased sample allows us to assess how well the ES is addressed by the publishing authors. In addition, we do not suspect that the manner in which our research topic was approached by researchers differs from that of other studies. However, given the limitations of the size of our sample of studies, caution should be used when generalizing. Future studies could utilize a new applied research question to address this objective.

4.2. Conclusions

In this paper, a systematic review was conducted on a very specific health-related issue to highlight this argument in an applied setting of interest. The present research revealed that approximately half of the studies do not report effect size indicators. In addition, approximately half of the studies do not interpret effect size in any way. There is a promising body of research demonstrating the usefulness of spiritual therapies in health conditions, including substance abuse relapse. However, there is a need for improved methodological rigor when reporting and interpreting effect sizes. It is not only desirable to calculate and report statistical indicators, but also to place them in the context of the research. It could be argued that research in spiritual or religious interventions in substance abuse is not representative of general scientific research. However, the authors writing on this specific topic do not necessarily report their findings differently from other researchers. Thus, the present authors argue that the results of the current review stand as a cautionary tale, a warning for researchers in any area of applied research.

Author Contributions: Conceptualization, I.S.-I. and J.S; methodology, I.S.-I. and A.J.M..; validation, J.S.; formal analysis, I.S.-I. and A.J.M.; investigation, A.J.M.; writing—original draft preparation, I.S.-I. and A.J.M.; writing—review and editing, J.S. and T.G.; visualization, T.G.; supervision, I.S.-I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

552 Search Strings for the Systematic Review (ProQuest Databases)

553 (NOFT(addiction) OR NOFT("substance abuse")) AND (NOFT(spirituality) OR NOFT(spiritual))
554 AND (NOFT(relapse) OR NOFT(treatment)) AND stype.exact("Scholarly Journals") AND (stype-
555 .exact("Scholarly Journals" NOT ("Dissertations & Theses" OR "Books" OR "Reports")) AND
556 pd(20150101-20201231))

557

558 **References**

559 1. Calin-Jageman, R.J.; Cumming, G. The New Statistics for Better Science: Ask How Much, How Uncertain, and What Else
560 Is Known. *The American Statistician* **2019**, *73*, 271–280. DOI: 10.1080/00031305.2018.1518266
561 2. Gill, J. The insignificance of null hypothesis significance testing. *Political Research Quarterly* **1999**, 52, 647-674. DOI:
562 10.1177/106591299905200309
563 3. Schneider, J.W. Null hypothesis significance tests: A mix-up of two different theories: The basis for widespread confusion
564 and numerous misinterpretations. *Scientometrics* **2015**, 102, 411-432. DOI: 10.1007/s11192-014-1251-5
565 4. García-Pérez, M.A. Thou shalt not bear false witness against null hypothesis significance testing. *Educational and Psycho-*
566 *logical Measurement* **2017**, *77*, 631-662. DOI: 10.1177/0013164416668232
567 5. Mayo, D.G.; Hand, D. Statistical significance and its critics: practicing damaging science, or damaging scientific practice?.
568 *Synthese* **2022**, *200*, 1-33. DOI: 10.1007/s11229-022-03692-0
569 6. Nickerson, R.S. Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*
570 **2000**, 5, 241-301. DOI: 10.1037/1082-989x.5.2.241
571 7. Cumming, G. *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge: New York, NY,
572 USA, 2012.
573 8. Cumming, G. The new statistics: Why and how. Psychological Science **2014**, 25, 7–29. DOI: 10.1177/0956797613504966
574 9. Cohen, J. *Statistical power analysis for the behavioral sciences*, 1st ed. Academic Press: New York, NY, USA, 1969
575 10. Cohen, J. *Statistical power analysis for the behavioral sciences,* 2nd ed. Academic Press: New York, NY, USA, 1988
576 11. Shadish, W.R.; Cook, T.D.; Campbell, D.T. *Experimental and quasi-experimental designs for generalized causal inference*.
577 Houghton Mifflin: Boston, MA, USA, 2002.
578 12. American Psychological Association. *Publication manual of the American Psychological Association* 7th ed. American Psycho-
579 logical Association: Washington, DC, USA, 2019.
580 13. American Psychological Association. *Publication manual of the American Psychological Association* 6th ed. American Psycho-
581 logical Association: Washington, DC, USA, 2010.
582 14. Eich, E. Business not as usual. *Psychological Science* **2014**, *25*, 3–6. DOI: 10.1177/0956797613512465
583 15. Ferguson, C.J. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*
584 **2009**, *40*, 532–538. DOI: 10.1037/a0015808
585 16. Vacha-Haase, T.; Nilsson, J.E.; Reetz, D.R.; Lance, T.S.; Thompson, B. Reporting Practices and APA Editorial Policies Re-
586 garding Statistical Significance and Effect Size. *Theory & Psychology* **2000**, 10, 413–425. DOI: 10.1177/0959354300103006
587 17. Vacha-Haase, T.; Thompson, B. How to estimate and interpret various effect sizes. *Journal of counseling psychology* **2004**, *51*,
588 473. DOI: 10.1037/0022-0167.51.4.473
589 18. Volker, M.A. Reporting effect size estimates in school psychology research. *Psychology in the Schools* **2006**, *43*, 653-672.
590 DOI: 10.1002/pits.20176653
591 19. Martín-Aguilar, C.; Sánchez-Iglesias, I. Interpretación del Efecto de Tratamientos Farmacológicos con Antidepresivos: ¿La
592 Diferencia Estadística Implica Relevancia Clínica? [Interpreting the effect size of pharmacological treatments with antide-
593 pressants. Does statistical difference imply clinical relevance?]. In Proceedings of the V Congreso Nacional de Psicología.
594 Online, (9-11 July 2021)
595 20. Elvira-Flores, E.; Sánchez-Iglesias, I. Rigor metodológico en la interpretación del efecto del ejercicio físico sobre síntomas
596 depresivos: la importancia del tamaño del efecto [Methodological rigor in the interpretation of the effect size in physical
597 exercise on depressive syntoms]. In Proceedings of the V Congreso Nacional de Psicología. Online, (9-11 July 2021)
598 21. Anderson, M.S.; Ronning, E.A.; De Vries, R.; Martinson, B.C. The Perverse Effects of Competition on Scientists' Work and
599 Relationships. *Science and Engineering Ethics* **2007**, *13*, 437–461 . DOI: 10.1007/s11948-007-9042-5
600 22. Fanelli, D. Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLoS ONE* **2010**,
601 *5*, e10271. DOI: 10.1371/journal.pone.0010271
602 23. Picho, K.; Artino, A.R. 7 Deadly Sins in Educational Research. *Journal of Graduate Medical Education* **2016**, *8*, 483–487. DOI:
603 10.4300/jgme-d-16-00332.1
604 24. Wicherts, J.M.; Veldkamp, C.L.; Augusteijn, H.E.; Bakker, M.; Van Aert, R.; Van Assen, M.A. Degrees of freedom in plan-
605 ning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in psychology* **2016**,
606 *7*, 1832. DOI: 10.3389/fpsyg.2016.01832
607 25. Sánchez-Iglesias, I.; González-Castaño, M.; Molina, A.J. Use of Causal Language in Studies on the Relationship between
608 Spiritually-based Treatments and Substance Abuse and Relapse Prevention. *Religions* **2021**, *12*, 1075. DOI: 10.3390/
609 rel12121075
610 26. Banks, G.C.; Rogelberg, S.G.; Woznyj, H.M.; Landis, R.S.; Rupp, D.E. Editorial: Evidence on Questionable Research Prac-
611 tices: The Good, the Bad, and the Ugly. *Journal of Business and Psychology* **2016**, *31*, 323–338. DOI: 10.1007/s10869-016-9456-
612 *7*

613 27. Xie, Y.; Wang, K.; Kong, Y. Prevalence of research misconduct and questionable research practices: a systematic review
614 and meta-analysis. *Science and Engineering Ethics* **2021**, *27*, 1-28. DOI: 10.1007/s11948-021-00314-9
615 28. Bergin, A.E. *Values and religious issues in psychotherapy and mental health. American Psychologist* **1991**, *46, 394–403.* DOI:
616 10.1037/0003-066x.46.4.394
617 29. Koenig, H.G.; Ford, S.M.; George, L.K.; Blazer, D.G.; Meador, K.G. *Religion and anxiety disorder: An examination and compar-*
618 *ison of associations in young, middle-aged, and elderly adults. Journal of Anxiety Disorders* **1993**, *7, 321–342.* DOI: 10.1016/0887-
619 6185(93)90028-j
620 30. Steffen, P.R.; Hinderliter, A.L.; Blumenthal, J.A.; Sherwood, A. *Religious Coping, Ethnicity, and Ambulatory Blood Pressure.*
621 *Psychosomatic Medicine* **2001**, *63, 523–530.* DOI: 10.1097/00006842-200107000-00002
622 31. Contrada, R.J.; Goyal, T.M.; Cather, C.; Rafalson, L.; Idler, E.L.; Krause, T.J. *Psychosocial Factors in Outcomes of Heart*
623 *Surgery: The Impact of Religious Involvement and Depressive Symptoms. Health Psychology* **2004**, *23, 227–238.* DOI:
624 10.1037/0278-6133.23.3.227
625 32. Koenig, H.G.; King, D.E.; Carson, V.B. *Handbook of religion and health,* 2nd ed. New York: Oxford University Press. 2012.
626 33. Saiz, J.; Pung, M.A.; Wilson, K.L.; Pruitt, C.; Rutledge, T.; Redwine, L.; … Mills, P.J. *Is Belonging to a Religious Organization*
627 *Enough? Differences in Religious Affiliation Versus Self-ratings of Spirituality on Behavioral and Psychological Variables in Indi-*
628 *viduals with Heart Failure. Healthcare* **2020**, *8, 129.* DOI: 10.3390/healthcare8020129
629 34. King, M.B.; Koenig, H.G. Conceptualising spirituality for medical research and health service provision. *BMC Health Ser-*
630 *vices Research* **2009**, *9*, 1-7. DOI: 10.1186/1472-6963-9-116
631 35. Koenig, H.G. Concerns About Measuring "Spirituality" in Research. *The Journal of Nervous and Mental Disease* **2008**, *196*,
632 349–355. DOI: 10.1097/nmd.0b013e31816ff796
633 36. Gavriel-Fried, B.; Moretta, T.; Potenza, M.N. Modeling intrinsic spirituality in gambling disorder. *Addiction Research &*
634 *Theory* **2019**, *28*, 1–7. DOI: 10.1080/16066359.2019.1622002
635 37. Gutierrez, I.A.; Chapman, H.; Grubbs, J.B.; Grant, J. Religious and spiritual struggles among military veterans in a resi-
636 dential gambling treatment programme. *Mental Health, Religion & Culture* **2020**, *23*, 187–203. DOI:
637 10.1080/13674676.2020.1764513
638 38. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, 5th ed.; American Psychiatric Pub-
639 lishing: Arlington, VA, USA, 2013. DOI: 10.1176/appi.books.9780890425596
640 39. Laespada, T.; Iraurgi, I. *Reducción de daños: lo aprendido de la heroína* [Harm reduction: what we have learned from heroin].
641 Ed. Universidad de Deusto: Deusto, Bilbao, España, 2009.
642 40. Yates, R. *Tackling addiction: Pathways to recovery*. Jessica Kingsley Pub.: London, UK, 2010.
643 41. Klingemann, H.; Sobell, L.C. *Natural recovery from alcohol and drug problems: A methodological review of the literature from*
644 *1999 through 2006*. Springer: New York, NY, USA, 2007.
645 42. MacGregor, S. Addiction recovery: A movement for social change and personal growth in the UK, by David Best,
646 Brighton. *Drugs: Education, Prevention and Policy* **2012**, *19*, 351–352. DOI: 10.3109/09687637.2012.69259
647 43. Carballo, J.L.; Fernández-Hermida, J.R.; Secades-Villa, R.; García-Rodríguez, O. Determinantes de la recuperación de los
648 problemas de alcohol en sujetos tratados y no tratados en una muestra española [Determinants of recovery from alcohol
649 problems in treated and untreated subjects in a Spanish sample]. *Adicciones* **2008**, *20*, 49-58.
650 44. Moos, R.H.; Finney, J.W. Commentary on Lopez-Quintero et al.(2011): Remission and relapse–the Yin-Yang of addictive
651 disorders. *Addiction* **2011**, *106*, 670-671. DOI: 10.1111/j.1360-0443.2010.00003284.x
652 45. Granfield, R.; Cloud, W. Social context and "natural recovery": The role of social capital in the resolution of drug-associ-
653 ated problems. *Substance Use & Misuse* **2001**, *36*, 1543–70. DOI: 10.1081/JA-100106963
654 46. Groshkova, T.; Best, D. The Evolution of a UK Evidence Base for Substance Misuse Recovery. *Journal of Groups in Addic-*
655 *tion & Recovery* **2011**, *6*, 20–37. DOI: 10.1080/1556035x.2011.571135
656 47. Putnam, R.D. *Bowling alone: The collapse and revival of American community*. Simon & Schuster: New York, NY, USA, 2001.
657 48. Groshkova, T.; Best, D.; White, W. The Assessment of Recovery Capital: Properties and psychometrics of a measure of ad-
658 diction recovery strengths. *Drug and Alcohol Review* **2012**, *32*, 187–194. DOI: 10.1111/j.1465-3362.2012.00489.x
659 49. Best, D.; Bliuc, A.-M.; Iqbal, M.; Upton, K.; Hodgkins, S. Mapping social identity change in online networks of addiction
660 recovery. *Addiction Research & Theory* **2017**, *26*, 163–173. DOI: 10.1080/16066359.2017.1347258
661 50. Best, D. *Addiction recovery: a movement for social change and personal growth in the UK*. Pavilion Publishing: Brighton, UK,
662 2012.
663 51. Puchalski, C.M. The spiritual dimension: The healing force for body and mind. In *Caregiving book series*; Puchalski, C.M.,
664 Ed.; Rosalyn Carter Institute for Human Development, Georgia Southwestern State University: Americus, GA, USA,
665 2003; pp. 174–195.
666 52. Galanter, M.; Dermatis, H.; Talbot, N.; McMahon, C.; Alexander, M.J. Introducing spirituality into psychiatric care.
667 *Journal of Religion and Health* **2011**, *50*, 81–91. DOI: 10.1007/s10943-009-9282-6
668 53. Alcoholics Anonymous. *Alcoholics Anonymous, the Big Book*, 4th ed., Alcoholics Anonymous World Services: New York,
669 NY, USA, 2001.
670 54. Andó, B.; Álmos, P.Z.; Németh, V.L.; Kovács, I.; Fehér-Csókás, A.; Demeter, I.; Rózsa, S.; Urbán, R.; Kurgyis, E.; Szikszay,
671 P.; Janka, Z.; Demetrovics, Z; Must, A. Spirituality mediates state anxiety but not trait anxiety and depression in alcohol
672 recovery. *Journal of Substance Use* **2016**, *21*, 344-348. DOI: 10.3109/14659891.2015.1021869
673 55. Magura, S.; Cleland, C.M.; Tonigan, J.S. Evaluating Alcoholics Anonymous's Effect on Drinking in Project MATCH Using
674 Cross-Lagged Regression Panel Analysis. *Journal of Studies on Alcohol and Drugs* **2013**, *74*, 378–385. DOI: 10.15288/
675 jsad.2013.74.378
676 56. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; … Moher, D. The PRISMA 2020
677 statement: an updated guideline for reporting systematic reviews. *BMJ* **2021**, *71*, 1-9. DOI: 10.1136/bmj.n71

678 57. Webster, D. The effects of spirituality on drug use. *Journal of Human Behavior in the Social Environment* **2015**, *25*, 322-332. DOI: 10.1080/10911359.2014.969126
679
680 58. Yeterian, J.D.; Bursik, K.; Kelly, J.F. Religiosity as a predictor of adolescents' substance use disorder treatment outcomes. *Substance abuse* **2015**, *36*: 453-461. DOI: 10.1080/08897077.2014.960550
681
682 59. Luna, N.; Horton, G.; Newman, D.; Malloy, T. An empirical study of attachment dimensions and mood disorders in inpa-
683 tient substance abuse clients: The mediating role of spirituality. *Addiction Research & Theory* **2016**, *24*. 248-260. DOI:
684 10.3109/16066359.2015.1119267
685 60. McClintock, C.H.; Worhunsky, P.D.; Xu, J.; Balodis, I.M.; Sinha, R.; Miller, L.; Potenza, M.N. Spiritual experiences are re-
686 lated to engagement of a ventral frontotemporal functional brain network: Implications for prevention and treatment of
687 behavioral and substance addictions. *Journal of Behavioral Addictions* **2019**, *8*, 678–691. DOI: 10.1556/2006.8.2019.71
688 61. Nurulhuda, M.H.; Haneem, N.; Khairi, C.M.; Norwati, D.; Aniza, A.A. Spiritual influence towards relapse in opioid ad-
689 dicts in therapy. *IIUM Medical Journal Malaysia* **2018**, 17, 71-74. DOI: 10.31436/imjm.v17i1.1032
690 62. Saari, C.Z.; Basirah, S.; Muhsin, S.; Syukri, M.; Abidin, Z.; Mohammad, S.; Syed, H.; Rahman, A.; Ahmad, S.S.; Ab Raman,
691 Z.B.; Manawi, M.; Akib, M.M.M.; Hamjah, S.H.; Farooque, M.H.; Rashid, M.A. Critical review of Sufi healing therapy in
692 drug addiction treatment. *Journal of Critical Reviews* **2020**, *7*, 1155-1160. DOI: 10.31838/jcr.07.05.220
693 63. Ng, S.M.; Rentala, S.; Chan, C.L.; Nayak, R.B. Nurse-Led Body–Mind–Spirit Based Relapse Prevention Intervention for
694 People With Diagnosis of Alcohol Use Disorder at a Mental Health Care Setting, India: A Pilot Study. *Journal of Addictions*
695 *Nursing* **2020**, 31, 276-286. DOI: 10.1097/JAN.0000000000000368
696 64. Abdollahi, A.; Abu Talib, M. Hardiness, spirituality, and suicidal ideation among individuals with substance abuse: The
697 moderating role of gender and marital status. *Journal of Dual Diagnosis* **2015**, *11*, 12-21. DOI: 10.1080/15504263.2014.988558
698 65. Beckstead, D.J.; Lambert, M.J.; DuBose, A.P.; Linehan, M. Dialectical behavior therapy with American Indian/Alaska Na-
699 tive adolescents diagnosed with substance use disorders: Combining an evidence based treatment with cultural, tradi-
700 tional, and spiritual beliefs. *Addictive behaviors* **2015**, *51*, 84-87. DOI: 10.1016/j.addbeh.2015.07.018
701 66. Jacobson, N.S.; Truax, P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy re-
702 search. *Journal of Consulting and Clinical Psychology* **1991**, *59*, 12–19. DOI: 10.1037/10109-042
703 67. Crutchfield Jr, D.A.; Güss, C.D. Achievement linked to recovery from addiction: Discussing education, vocation, and non-
704 addict identity. *Alcoholism Treatment Quarterly* **2019**, *37*, 359-376. DOI: 10.1080/07347324.2018.1544058t
705 68. Dickerson, D.L.; D'Amico, E.J.; Klein, D.J.; Johnson, C.L.; Hale, B.; Ye, F. Mental health, physical health, and cultural char-
706 acteristics among american indians/alaska natives seeking substance use treatment in an urban setting: A descriptive
707 study. *Community mental health journal* **2021**, *57*, 937-947. DOI: 10.1007/s10597-020-00688-3t
708 69. Kelly, J.F.; Eddie, D. The role of spirituality and religiousness in aiding recovery from alcohol and other drug problems:
709 An investigation in a national US sample. *Psychology of religion and spirituality* **2020**, *12*, 116–123. DOI: 10.1037/rel0000295
710 70. Kerlin, A.M. Therapeutic change in a Christian SUD program: Mental health, attachment, and attachment to God. *Alco-*
711 *holism Treatment Quarterly* **2017**, *35*, 395-411. DOI: 10.1080/07347324.2017.1355218
712 71. Lashley, M. The impact of length of stay on recovery measures in faith-based addiction treatment. *Public Health Nursing*
713 **2018**, *35*, 396-403. DOI: 10.1111/phn.12401
714 72. Lee, M.T.; Pagano, M.E.; Johnson, B.R.; Post, S.G.; Leibowitz, G.S.; Dudash, M. From defiance to reliance: Spiritual virtue
715 as a pathway towards desistence, humility, and recovery among juvenile offenders. *Spirituality in Clinical Practice* **2017**, *4*,
716 161-175. DOI: 10.1037/scp0000144
717 73. Mallik, D.; Bowen, S.; Yang, Y.; Perkins, R.; Sandoz, E.K. Raja yoga meditation and medication-assisted treatment for re-
718 lapse prevention: A pilot study. *Journal of Substance Abuse Treatment* **2019**, *96*. 58-64. DOI:10.1016/j.jsat.2018.10.012
719 74. Medlock, M.M.; Rosmarin, D.H.; Connery, H.S.; Griffin, M.L.; Weiss, R.D.; Karakula, S.L.; McHugh, R.K. Religious coping
720 in patients with severe substance use disorders receiving acute inpatient detoxification. *The American Journal on Addictions*
721 **2017**, *26*, 744-750. DOI: 10.1111/ajad.12606
722 75. Montes, K.S.; Tonigan, J.S. Does Age Moderate the Effect of Spirituality/Religiousness in Accounting for Alcoholics
723 Anonymous Benefit? *Alcoholism Treatment Quarterly* **2017**, *35*, 96–112. DOI: 10.1080/07347324.2017.1288487
724 76. Ranes, B.; Johnson, R.; Nelson, L.; Slaymaker, V. The Role of Spirituality in Treatment Outcomes Following a Residential
725 12-Step Program. *Alcoholism Treatment Quarterly* **2016**, *35*, 16–33. DOI: 10.1080/07347324.2016.1257275
726 77. Ransome, Y.; Haeny, A.M.; McDowell, Y.E.; Jordan, A. Religious involvement and racial disparities in opioid use disorder
727 between 2004–2005 and 2012–2013: Results from the National Epidemiologic Survey on Alcohol and Related Conditions.
728 *Drug and Alcohol Dependence* **2019**, *205*, 107615. DOI: 10.1016/j.drugalcdep.2019.107615
729 78. Shorey, R.C.; Gawrysiak, M.J.; Anderson, S.; Stuart, G.L. Dispositional Mindfulness, Spirituality, and Substance Use in
730 Predicting Depressive Symptoms in a Treatment-Seeking Sample. *Journal of Clinical Psychology* **2014**, *71*, 334–345. DOI:
731 10.1002/jclp.22139
732 79. Temme, L.J.; Kopak, A.M. Maximizing recovery through the promotion of mindfulness and spirituality. Journal of Reli-
733 gion & Spirituality in Social Work: *Social Thought* **2016**, *35*, 41-56. DOI: 10.1080/15426432.2015.1067591
734 80. Tianingrum, N.A.; Feriani, P.; Susanti, E.W.; Purdani, K.S.; Winarti, Y.; Safrudin, B. The Effect of Narcotics Anonymous
735 Meeting toward Relapse Prevention among Prisoners. *Indian Journal of Public Health Research & Development* **2019**, *10*, 634-
736 38. DOI: 10.5958/0976-5506.2019.00579.5
737 81. Yaghubi, M.; Abdekhoda, M.; Khani, S. Effectiveness of Religious-Spiritual Group Therapy on Spiritual Health and Qual-
738 ity of Life in Methadone-treated Patients: A Randomized Clinical Trial. *Addiction & Health* **2019**, *11*, 156-64. DOI:
739 10.22122/ahj.v11i3.238
740 82. Yeterian, J.D.; Bursik, K.; Kelly, J.F. "God put weed here for us to smoke": A mixed-methods study of religion and spiritu-
741 ality among adolescents with cannabis use disorders. *Substance abuse* **2018**, *39*, 484-492. DOI:
742 10.1080/08897077.2018.1449168
743 83. Fisher, R.A. *Statistical methods for research workers*. Oliver and Boyd: Edinburgh, UK, 1932.

744 84. Neyman, J.; Pearson, E.S. On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference: Part
745 I. Biometrika **1928**, 20A, 175–240.
746 85. Heene, M.; Ferguson, C.J. The need for Bayesian hypothesis testing in psychological science. In *Psychological science under*
747 *scrutiny: Recent challenges and proposed solutions*; Lilienfeld, S.O., Waldman, I., Eds.; Wiley: New York, NY, USA, 2017; pp.
748 34-52.
749 86. Molina, A.; Saiz, J.; Gil, F.; Cuenca, M.L.; Goldsby, T. Psychosocial Intervention in European Addictive Behaviour Recov-
750 ery Programmes: A Qualitative Study. *Healthcare* **2020**, *8*, 268. DOI: 10.3390/healthcare8030268
751 87. Bumbarger, B.K.; Campbell, E.M. A State Agency–University Partnership for Translational Research and the Dissemina-
752 tion of Evidence-Based Prevention and Intervention. *Administration and Policy in Mental Health and Mental Health Services*
753 *Research* **2011**, *39*, 268–277. DOI: 10.1007/s10488-011-0372-x
754