

BRIEF REPORT

Comparing Two Books and Establishing Probably Efficacious Treatment for Low Sexual Desire

Alexandra M. Balzer
University of Missouri

Laurie B. Mintz
University of Florida

Using a sample of 45 women, this study compared the effectiveness of a previously studied (Mintz, Balzer, Zhao, & Bush, 2012) bibliotherapy intervention (Mintz, 2009), a similar self-help book (Hall, 2004), and a wait-list control (WLC) group. To examine intervention effectiveness, between and within group standardized effect sizes (interpreted with Cohen's, 1988 benchmarks .20 = small, .50 = medium, .80+ = large) and their confidence limits are used. In comparison to the WLC group, both interventions yielded large between-group posttest effect sizes on a measure of sexual desire. Additionally, large between-group posttest effect sizes were found for sexual satisfaction and lubrication among those reading the Mintz book. When examining within-group pretest to posttest effect sizes, medium to large effects were found for desire, lubrication, and orgasm for both books and for satisfaction and arousal for those reading the Mintz book. When directly comparing the books, all between-group posttest effect sizes were likely obtained by chance. It is concluded that both books are equally effective in terms of the outcome of desire, but whether or not there is differential efficacy in terms of other domains of sexual functioning is equivocal. Tentative evidence is provided for the longer term effectiveness of both books in enhancing desire. Arguing for applying criteria for empirically supported treatments to self-help, results are purported to establish the Mintz book as *probably efficacious* and to comprise a first step in this designation for the Hall book.

Keywords: low sexual desire, self-help, bibliotherapy, evidence-based practice, empirically supported treatments

Based on the need for accessible treatment for a high prevalence concern, Mintz, Balzer, Zhao, and Bush (2012) conducted a randomized clinical trial (RCT) of a bibliotherapy intervention for low sexual desire among women. These researchers reported that

compared to those in a wait-list control group, women who read *A Tired Woman's Guide to Passionate Sex* (Mintz, 2009) made greater gains on measures of sexual desire, sexual arousal, and sexual satisfaction, with all Cohen's *d* between-groups posttest effect sizes in the range considered large by Cohen (1988). Mintz et al. (2012) recommended future studies, including both replication and “. . . comparing the efficacy of this book with another self-help book on the topic” (p. 476).

This article was published Online First March 2, 2015.

Alexandra M. Balzer, Department of Educational, School, and Counseling Psychology, University of Missouri; Laurie B. Mintz, Department of Psychology, University of Florida.

This article is based on a doctoral dissertation conducted at the University of Missouri by Alexandra M. Balzer, directed by Laurie B. Mintz. Laurie B. Mintz is the author of the self-help book under study. Campus Institutional Review Board (IRB) approval was sought, including disclosing that the author of the book was an investigator and addressing conflict of interest issues. As detailed in the IRB application, although the book author stands to benefit from book sales that result from a publication on the efficacy of the book, this conflict is minimal in that book royalties are quite low. Additionally, earlier bibliotherapy efficacy studies have been conducted by the authors of the books under investigation (e.g., Morokoff, & LoPiccolo, 1986). Moreover, book authors holding responsibility for assessing the effectiveness of their products is in line with recent recommendations (McKendree-Smith et al., 2003). Importantly, the book author was not involved in participant recruitment, data collection, or data analysis. We thank Tessa Wimberley and Natalie E. Kelso for assistance with this article.

Correspondence concerning this article should be addressed to Laurie B. Mintz, Department of Psychology, University of Florida, P.O. Box 112250, Gainesville, FL 32611-2250. E-mail: mintzl@ufl.edu

Replication and extension is consistent with recent American Psychological Association (APA) policies and their history. In the 1990s, a Division 12 task force published criteria for empirically supported treatments (ESTs; Chambless et al., 1998), including those deemed “probably efficacious” and those deemed “well-established” (p. 8). These criteria sparked “a decade of enthusiasm and controversy” (APA, 2006, p. 272) during which several groups, including Division 17, weighed in on the issues. Subsequently, a 2005 APA Presidential Task Force endorsed “. . . psychology's fundamental commitment to sophisticated, evidence-based psychological practice” (EPPP), and defined such practice as an integration of “. . . the best available research evidence with clinical expertise in the context of patient characteristics, culture and preferences” (APA, 2006, p. 273). Best research evidence was delineated in accordance with prior APA (2002) criteria stating that RCTs are the most stringent way to evaluate treatment efficacy, that “. . . a single experiment from one setting does not provide sufficient evidence of efficacy . . .” (p. 1055), and that

interventions should be compared to both “doing nothing” (p. 1054) and to “alternative interventions that are known or believed to be effective” (p. 1055). This study replicates a RCT and extends it by comparing two books believed to be similarly effective in increasing sexual desire among women.

Having too little sexual desire is the most common sexual complaint identified in epidemiological surveys of women (Basson & Brotto, 2009). Twenty-five to 52% of women are distressed by low desire at some point in their lives (West et al., 2008), with the vast majority experiencing low desire “despite good health, good will, and even good relationships” (Leiblum, 2010, p. xii). However, only four studies could be located which offer empirical support for interventions for women with low desire, three being group face-to-face treatment (Brotto, Basson, & Luria, 2008; Hurlbert, 1993; Trudel et al., 2001) and one (Mintz et al., 2012) being the bibliotherapy intervention under investigation here. That there is only one study on the effectiveness of bibliotherapy for low sexual desire is surprising considering that women with this concern commonly seek out self-help resources (Rosen et al., 2009) and that bibliotherapy occupies a “stable position” in alleviating sexual concerns (van Lankveld, 2009, p. 143). Indeed, when reviewing studies which directly compared self-help and face-to-face treatment for sexual concerns, van Lankveld (2009) concluded that the “face-to-face treatment is certainly not always superior and that at least for certain types of sexual dysfunctions, the self-help format is to be recommended because of its better cost effectiveness ratio” (p. 150).

Whether this is true for women’s low desire has not been directly tested. However, computing effect sizes across independent studies is informative. Using data provided in published articles, the unbiased Cohen’s *d* pretest to posttest effect size (Hedges’ g_{av}) for the Female Sexual Function Index Desire Subscale (Rosen et al., 2000) was .69 (considered medium) for the 26 women undergoing face-to-face treatment in Brotto et al.’s (2008) study and 1.45 (considered large) for the 19 women reading a self-help book in Mintz et al.’s (2012) bibliotherapy study. These effects bolster the notion that bibliotherapy may be the “treatment of first choice in a stepped care model” (van Lankveld, 2009, p. 150).

Stepped care models promote seeking the least costly, least intensive treatment presumed to be effective first (Norcross, 2006). This model is long-recognized as particularly applicable to sexual concerns, with Annon (1974) first contending that 80% to 90% of sexual problems could be resolved without intensive therapy. Conducting research based on the stepped-care model is particularly pertinent in today’s era where psychological concerns are increasingly treated outside the therapist’s office (Teachman, 2014). Yet, the field’s definitions of EBPP and ESTs have not caught up to this reality. The two aforementioned APA policies (APA, 2002, 2006) implicitly and explicitly, respectively, restrict “. . . its consideration to evidence-based practice as it relates to health services provided by psychologists” (APA, 2006, p. 273). One of the authors of the Division 17 Principles of Empirically Supported Interventions in Counseling Psychology (Wampold, Lichtenberg, & Waehler, 2002) questions this restriction, noting that it is inconsistent with the aforementioned Counseling Psychology Principles (C. Waehler, personal communication, September 29, 2014). Along similar lines, the foremost authority on self-help for sexual dysfunction advocates applying EST criteria to biblio-

therapy, noting that two successful RCTs showing a book is more effective than a wait-list control group would establish that book as a probably efficacious treatment (J. van Lankveld, personal communication, January 24, 2014). It is hoped this study will spur the field to apply EST criteria to self-help.

Hypotheses of this study are that, among women self-reporting low sexual desire, reading the Mintz (2009) book and a similar self-help book (Hall, 2004) will be more effective than a wait-list control group in enhancing desire and other aspects of sexual functioning. Reading both books is hypothesized to result in positive change over time, with neither more effective than the other. This study could thus help establish a probably efficacious bibliotherapy treatment for women with low sexual desire and provide a first step toward this designation for a second book.

Method

Participants

Fifty-five women were randomly assigned to the Mintz Intervention (MI), Hall Intervention (HI), or a wait-list control (WLC) group. Only those completing pretest and posttest measures were included in the final sample, which included 45 participants (MI: $n = 13$; HI: $n = 18$; WLC: $n = 14$), as five participants did not complete the pretest (MI: $n = 2$; HI: $n = 1$; WLC: $n = 2$) and five participants (MI: $n = 4$; WLC: $n = 1$) did not complete posttest measures. All participants identified as heterosexual and married. Marriage length ranged from 1 to 34 years ($M = 15.13$, $SD = 9.68$). Ages ranged from 29 to 57 ($M = 42.64$, $SD = 8.66$). All participants were residing in the United States and most ($n = 41$; 91.1%) identified as White, one as African American, two as Biracial/Multiracial, and one as Asian/Pacific Islander. Most ($n = 32$; 71.1%) identified as Christian; other affiliations were nonreligious ($n = 6$), agnostic ($n = 2$), or “Other Religion” ($n = 5$). About 4.4% reported having a high school degree, 4.4% some college, 11.1% an associate’s degree, 28.9% a bachelor’s degree, 17.8% some graduate or professional training, 17.8% a master’s degree, 11.1% a doctoral degree, 2.2% an advanced professional degree, and 2.2% reported “Other Education.” Most (97.8%) were employed, with 91.1% working full-time. Household incomes were: \$25,000 to \$50,000 (6.7%); \$50,000 to \$75,000 (33.3%); \$75,000 to \$100,000 (35.6%); \$100,000 or more per year (24.4%). About 63% had children living at home.

Measures

The Hurlbert Index of Sexual Desire (HISD; Apt & Hurlbert, 1992) was the primary measure of sexual desire. This 25-item inventory yields total scores ranging from zero to 100, with higher scores indicating higher levels of sexual desire. Individual items are rated on a 5-point Likert-type scale, ranging from 0 (*all of the time*) to 4 (*never*). Beck (1995) reports good internal consistency ($\alpha = .86$), test-retest reliability ($r = .86$), and concurrent, construct, and discriminant validity. Mintz et al. (2012) reported internal consistencies of $\alpha = .93$ at pretest and $\alpha = .94$ at posttest, whereas in this study they were $\alpha = .91$ at pretest and $\alpha = .94$ at posttest.

The Female Sexual Function Index (FSFI; Rosen et al., 2000) was a second measure of desire, as well as five related aspects of

sexual functioning (see list below). This 19-item measure is the most widely used inventory of sexual functioning, with strong validity across multiple studies (Meyer-Bahlburg & Dolezal, 2007) and domain scores supported via factor analysis (Rosen et al., 2000). Scores are calculated with an algorithm. Score ranges are: Desire 1.2–6; Arousal 0–6; Lubrication 0–6; Orgasm 0–6; Satisfaction .8–6; and Pain 0–6. Higher scores represent higher functioning. Rosen et al. (2009) reports good 2- to 4-week test-retest reliability ($r = .79 - .86$) and internal consistency ($\alpha = .89 - .96$) across domains. Internal consistency (α) in this study for pretest and posttest, respectively, was Desire (.84, .84), Arousal (.96, .79), Lubrication (.95, .92) Orgasm (.95, .86), Satisfaction (.66, .83), and Pain (.97, .82).

Procedure

After receiving IRB approval, advertisements were placed in a weekly e-mail distributed at a public university. These ads stated: “Seeking heterosexual married women who feel satisfied with their marriages but who are bothered by a low sex drive” and explained the study involved reading one of two self-help books and answering personal questions on sexual functioning. For the first two ads the tagline was “Seeking Women with Diminished Sexual Desire for Intervention Study” and for the third it was: “Lacking Libido? Reclaim your Sexual Desire.”

Taking into account the 30% attrition rate observed in other studies of bibliotherapy (e.g., Mintz et al., 2012; van Lankveld et al., 2006), a priori power analyses revealed that 54 participants would be needed in order to obtain a power of 0.81 with an estimated population effect size of 1.00 (Cohen’s d) at an alpha of $p < .05$. The first 55 participants who met inclusion criteria (heterosexual, happily married, suffering from low sexual desire) were randomly assigned to the MI, HI, or WLC group, and e-mailed a link to the informed consent and pretest measures. Intervention group participants were then mailed their respective books and instructed to read it in 6 weeks, while WLC group participants were sent a letter stating they would receive the book in 6 weeks after completing additional measures. Six weeks after completing pretest measures, participants were e-mailed links to the posttest measures and asked if they had done anything outside of the study to address their low sexual desire. Intervention participants were asked about their favorite part of the book (manipulation check). No participants were excluded based on seeking outside help or not passing the manipulation check. To examine maintenance of gains, intervention group participants were also sent a link to the measures 6 weeks after completing the posttest. Twelve of the 13 participants in the MI and 13 of the 18 participants in the HI completed the follow-up measures. After study completion, participants were offered counseling referrals. Across all surveys, participants not responding within 5 days were e-mailed up to two reminders, spaced 5 days apart. The only incentive provided was a book. The WLC group was given the Mintz (2009) book, since prior research found it effective.

Interventions

A Tired Woman’s Guide to Passionate Sex (Mintz, 2009) is a 237-page self-help book for heterosexual, married women with low sexual desire and otherwise satisfying relationships. It con-

tains three foundational chapters detailing the author’s personal story, information about the causes of women’s low desire, and information on the physical and emotional benefits of sex. Five chapters comprise a psycho-educational and cognitive-behavioral treatment approach, including information and exercises on thought processes, communication strategies, time management, and sex therapy techniques (e.g., sensate focus, scheduled sexual encounters).

Reclaiming Your Sexual Self (Hall, 2004) is a 228-page self-help book for women with low sexual desire who have otherwise satisfying romantic relationships. It contains 10 chapters covering topics such as the author’s personal story, education about causes of low desire, partner communication strategies, and exercises for enhancing sexual desire (e.g., sensate focus, sensual meditation). The book is written from a systemic-feminist perspective and is recommended in the highly acclaimed and award winning book by Joannides (2012).

Results

Preliminary Analyses

Analyses of variance (ANOVAs) revealed no significant differences in demographic or dependent variables between the MI, HI, and WLC groups at baseline or between participants recruited using different advertisement taglines. Also in terms of demographic and dependent variables, independent samples t tests yielded no significant differences between those in the intervention groups who did and did not complete the follow-up measures, and one difference between those who completed the posttest and those who did not. The latter had poorer orgasm functioning at pretest ($M = 1.2$) than the former ($M = 3.8$).

Bibliotherapy Effectiveness

In intervention studies with less than 50 participants, standardized effect sizes with confidence limits convey the same information as tests of statistical significance but without confounding the size of the difference with the size of the sample. Effect sizes represent a simple way of quantifying the size of the difference between two groups, and thus for interventions studies can be defined as “. . . a standardized, scale-free measure of the relative size of the effect of an intervention” (Turner & Bernard, 2006, p. 42). For small samples, the Hedges g effect size is recommended as it corrects for the small sample bias inherent in Cohen’s d (Cumming, 2012; Turner & Bernard, 2006), and in fact can be relied upon when total sample sizes (including intervention and control combined) are less than 20 (Lakens, 2013; Turner & Bernard, 2006). Nevertheless, the effect size is still an estimate of an interventions’ effect in the population from which the sample is drawn and thus an effect calculated from a large sample may be more accurate than one calculated from a small sample. However, to “disentangle effect size and sample size” (Turner & Bernard, 2006, p. 44), confidence intervals for effect sizes can be used, as they represent the likely range of the true population mean effect size. If a 95% confidence interval does not include zero, the observed effect size can be concluded to not have been obtained by chance (Lakens, 2013; Turner & Bernard, 2006). Compared to inferential statistics, standardized effect sizes with confidence lim-

its: a) more directly answer questions about intervention effects; b) provide information about what is likely to happen on replication of an experiment; c) facilitate comparison; d) can be used in meta-analyses; e) tend to be intuitively more understandable; and f) provide information on practical significance (Cohen, 1988; Cumming, 2012; Kline, 2013). We thus employ Hedges g effect sizes and confidence limits and interpret effect sizes with Cohen's (1988) cautious rule of thumb: .20 = small, .50 = medium, .80+ = large. We also report common language effect sizes (McGraw & Wong, 1992), a percentage expressing ". . . the probability that a person from one group will have a higher observed measurement than a randomly sampled person from the other group" (Lakens, 2013, p. 4).

Immediate effectiveness. Table 1 presents the means and standard deviations for the three groups at pretest and posttest. As recommended by Lipsey et al. (2012), to examine the effect of the interventions on the outcomes, we employ both within-group pretest to posttest effect sizes (see Table 1) and between-group posttest effect sizes (see Table 2). In Table 1 we report Hedges' g_{av} (correlation between pretest and posttest accounted for and denominator is average standard) and in Table 2, we report Hedges g_s (denominator is pooled standard deviation). To further facilitate

group comparisons, Table 1 includes notations indicating if the magnitude (i.e., small, medium or large, using Cohen's, 1988 conventions) of each intervention pretest to posttest effect size is greater than the WLC group effect size, and if one intervention effect size is larger than the other. Both Table 1 and Table 2 also present common language effect sizes.

Follow-up effectiveness. Although effect sizes with confidence limits are not confounded with sample size, these results should still be considered exploratory. Table 3 presents the means, standard deviations, and pretest to posttest, posttest to follow-up, and pretest to follow-up Hedges g_{av} effect sizes for those who completed 6-week follow-up measures. Because of relevance in examining maintenance of gains, we report the pretest to follow-up common language effect size. Figure 1 presents a visual of the HISD for those in the intervention groups across all three time points and for those in the WLC group across pretest and posttest.

Discussion

Hypotheses of this study were that a previously studied self-help book (Mintz, 2009) and a similar self-help book (Hall, 2004) would be more effective than a wait-list control group in enhanc-

Table 1
Within-Group Pretest to Posttest Effect Sizes by Intervention Arm

Measure	Pretest		Posttest		Hedges' g_{av}	Effect sizes	
	M	SD	M	SD		[95% CI]	Common language
HISD							
WLC	30.57	11.76	32.82	11.40	0.18	[-0.09, 0.48]	0.64
MI	33.10	12.28	50.43	15.05	1.18*	[0.61, 1.89]	0.96
HI	31.56	9.12	42.13	10.41	1.03*	[0.57, 1.58]	0.92
Desire							
WLC	2.19	0.84	2.23	0.80	0.05	[-0.44, 0.54]	0.52
MI	2.26	0.98	2.73	0.68	0.51*†	[-0.15, 1.22]	0.68
HI	2.23	0.79	2.37	0.78	0.16	[-0.46, 0.79]	0.55
Arousal							
WLC	3.11	1.28	3.56	1.56	0.30	[-0.13, 0.75]	0.65
MI	2.65	1.90	3.76	1.24	0.65*†	[0.12, 1.25]	0.78
HI	2.95	1.60	3.32	1.07	0.26	[-0.14, 0.67]	0.62
Lub							
WLC	3.81	1.69	3.54	1.65	-0.16	[-0.47, 0.14]	0.61
MI	2.72	2.05	5.01	0.74	1.39*†	[0.58, 2.35]	0.89
HI	3.23	1.54	4.17	1.30	0.63*	[0.18, 1.12]	0.77
Org							
WLC	4.06	1.58	3.91	1.64	-0.08	[-0.46, 0.28]	0.55
MI	2.68	1.96	4.49	1.41	1.00*	[0.41, 1.70]	0.88
HI	2.62	1.83	4.14	1.37	0.90*	[0.42, 1.44]	0.86
Pain							
WLC	4.63	1.75	4.74	1.38	0.07	[-0.45, 0.60]	0.53
MI	3.05	2.66	4.71	1.61	0.71*†	[0.00, 1.48]	0.73
HI	4.67	1.66	4.31	1.18	-0.24*	[-0.86, 0.36]	0.58
Sat							
WLC	3.17	1.09	3.29	1.11	0.10	[-0.32, 0.52]	0.55
MI	3.05	1.35	4.77	1.41	1.17*†	[0.44, 2.03]	0.86
HI	3.33	1.14	4.11	1.48	0.56*	[-0.01, 1.17]	0.69

Note. MI = Mintz Intervention ($N = 13$); HI = Hall Intervention ($N = 18$); WLC = Wait-List Control Group ($N = 14$). Using Cohen's (1988) conventions of small = .20, medium = .50, large = .80+, * indicates magnitude of treatment pretest to posttest effect size is greater for treatment than WLC group and † indicates magnitude is greater for that intervention than the other intervention.; HISD = Hurlbert Index of Sexual Desire (range 0–100); Desire = Female Sexual Functioning Index Desire Subscale (range 1.2–6); Arousal = Female Sexual Functioning Index Arousal Subscale (range 0–6); Lub = Female Sexual Functioning Index Lubrication Subscale (range 0–6); Org = Female Sexual Functioning Index Orgasm Subscale (range 0–6); Pain = Female Sexual Functioning Index Pain Subscale (range 0–6); Sat = Female Sexual Functioning Index Satisfaction Subscale (range 0–6). For all measures, higher scores indicate higher levels of aspects of sexual functioning.

Table 2
Between-Group Posttest Effect Sizes and Confidence Intervals

Measure	Between-group posttest effect sizes								
	MI vs. WLC			HI vs. WLC			MI vs. HI		
	Hedges' g_s	[95% CI]	Common language	Hedges' g_s	[95% CI]	Common language	Hedges' g_s	[95% CI]	Common language
HISD	1.29	[0.46, 2.12]	0.82	0.84	[0.11, 1.56]	0.73	0.64	[-0.09, 1.38]	0.67
Desire	0.65	[-0.12, 1.43]	0.68	0.17	[-0.53, 0.87]	0.55	0.47	[-0.25, 1.20]	0.64
Arousal	0.14	[-0.62, 0.89]	0.54	-0.18	[-0.88, 0.52]	0.55	0.37	[-0.34, 1.09]	0.61
Lub	1.10	[0.29, 1.91]	0.79	0.42	[-0.29, 1.13]	0.62	0.74	[0.00, 1.48]	0.71
Org	0.37	[-0.39, 1.13]	0.61	0.15	[-0.55, 0.85]	0.54	0.25	[-0.47, 0.96]	0.57
Pain	-0.02	[-0.77, 0.74]	0.51	-0.33	[-1.03, 0.37]	0.59	0.28	[-0.43, 1.00]	0.58
Sat	1.14	[0.32, 1.95]	0.80	0.60	[-0.11, 1.31]	0.67	0.44	[-0.28, 1.16]	0.63

Note. MI = Mintz Intervention ($N = 13$); HI = Hall Intervention ($N = 18$); WLC = Wait-List Control Group ($N = 14$); HISD = Hurlbert Index of Sexual Desire (range 0–100); Desire = Female Sexual Functioning Index Desire Subscale (range 1.2–6); Arousal = Female Sexual Functioning Index Arousal Subscale (range 0–6); Lub = Female Sexual Functioning Index Lubrication Subscale (range 0–6); Org = Female Sexual Functioning Index Orgasm Subscale (range 0–6); Pain = Female Sexual Functioning Index Pain Subscale (range 0–6); Sat = Female Sexual Functioning Index Satisfaction Subscale (range 0–6). For all measures, higher scores indicate higher levels of aspects of sexual functioning.

ing sexual functioning, with both hypothesized to result in positive change over time and neither more effective than the other. In terms of the main outcome of sexual desire, hypotheses were confirmed. In terms of outcomes in sexual functioning beyond desire, results were mixed. Whether or not the two books performed similarly is equivocal.

In terms of the effectiveness of the MI, the results of this study mainly replicated the results of a previous RCT. When comparing

those reading the book to those in a wait-list control group, large between-group posttest effect sizes were found on a measure of sexual desire (HISD), as well as on measures of lubrication and satisfaction. Additional statistics reported in this study, but not in the prior RTC, bolster these results. First, confidence limits indicated that these results were likely not due to chance. Second, common language effect sizes indicated that the likelihood that those reading the book would score higher at posttest than those in

Table 3
Within-Group Effect Sizes for Participants Completing Follow-Up Measures Across Two Interventions

Measure	Effect sizes												
	Pretest		Posttest		Follow-up		Pre to post		Post to follow-up		Pre to follow-up		Common language
	M	SD	M	SD	M	SD	Hedges' g_{av}	[95% CI]	Hedges' g_{av}	[95% CI]	Hedges' g_{av}	[95% CI]	
HISD													
MI	30.45	8.03	47.96	12.69	48.43	13.05	1.54	[0.75, 2.52]	0.03	[-0.50, 0.57]	1.54	[0.73, 2.56]	0.88
HI	35.00	7.99	45.38	10.30	44.13	12.12	1.05	[0.44, 1.79]	-0.10	[-0.41, 0.19]	0.83	[0.25, 1.50]	0.74
Desire													
MI	2.10	0.83	2.65	0.65	3.40	1.10	0.69 [†]	[-0.10, 1.55]	0.77	[0.18, 1.46]	1.25 [†]	[0.28, 2.36]	0.83
HI	2.40	0.65	2.35	0.83	3.14	1.10	-0.06	[-0.81, 0.69]	0.75	[0.25, 1.34]	0.77	[0.07, 1.54]	0.72
Arousal													
MI	2.45	1.83	3.68	1.25	5.13	0.85	0.73 [†]	[0.15, 1.39]	1.26 [†]	[0.55, 2.12]	1.75 [†]	[0.83, 2.89]	0.91
HI	3.32	1.43	3.37	1.18	3.56	2.12	0.03	[-0.30, 0.37]	0.10	[-0.35, 0.57]	0.12	[-0.31, 0.56]	0.54
Lub													
MI	2.63	2.11	4.95	0.74	4.60	1.69	1.37 [†]	[0.53, 2.37]	-0.25	[-0.92, 0.40]	0.96 [†]	[0.25, 1.78]	0.77
HI	3.41	1.51	4.08	1.49	3.18	1.95	0.42	[0.08, 0.80]	-0.49	[-1.05, 0.03]	-0.12	[-0.53, 0.27]	0.54
Org													
MI	2.68	2.04	4.47	1.47	4.70	1.74	0.96 [†]	[0.36, 1.68]	0.14	[-0.29, 0.57]	1.02 [†]	[0.25, 1.90]	0.77
HI	3.05	1.92	4.35	1.52	3.23	2.22	0.70	[0.24, 1.25]	-0.55	[-1.16, 0.01]	0.08	[-0.41, 0.59]	0.52
Pain													
MI	2.97	2.76	4.63	1.66	5.43	1.73	0.68 [†]	[-0.05, 1.49]	0.44	[0.11, 0.83]	1.00 [†]	[0.27, 1.84]	0.77
HI	4.78	1.26	4.12	1.20	4.77	1.81	-0.50	[-1.15, 0.11]	0.39	[-0.18, 1.00]	-0.01	[-0.53, 0.52]	0.50
Sat													
MI	3.00	1.40	4.73	1.47	4.17	1.38	1.12 [†]	[0.37, 2.01]	-0.37	[-0.70, -0.09]	0.78 [†]	[-0.04, 1.69]	0.72
HI	3.63	1.05	4.31	1.46	3.57	1.65	0.50	[-0.08, 1.13]	-0.44	[-0.98, 0.05]	-0.04	[-0.59, 0.50]	0.51

Note. MI = Mintz Intervention ($N = 12$), HI = Hall Intervention ($N = 13$). Using Cohen's (1988) conventions of small = .20, medium = .50, large = .80+, [†] indicates magnitude of effect size is greater for that intervention than the other intervention. HISD = Hurlbert Index of Sexual Desire (range 0–100); Desire = Female Sexual Functioning Index Desire Subscale (range 1.2–6); Arousal = Female Sexual Functioning Index Arousal Subscale (range 0–6); Lub = Female Sexual Functioning Index Lubrication Subscale (range 0–6); Org = Female Sexual Functioning Index Orgasm Subscale (range 0–6); Pain = Female Sexual Functioning Index Pain Subscale (range 0–6); Sat = Female Sexual Functioning Index Satisfaction Subscale (range 0–6). For all measures, higher scores indicate higher levels of aspects of sexual functioning.

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

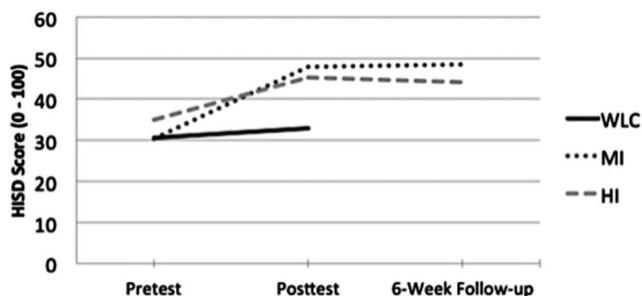


Figure 1. Hurlbert Index of Sexual Desire scores for Mintz intervention (MI), Hall intervention (HI), and wait-list control group (WLC).

the WLC group was 82% for desire, 79% for lubrication, and 80% for satisfaction. Third, large pretest to posttest within group effect sizes not likely obtained by chance were found for these same three dimensions, with large and medium effects also found for orgasm and arousal, respectively. Finally, the percent likelihood that those reading the book would score higher at posttest than pretest ranged from 78% (arousal) to 96% (desire), with remaining percentages being: 89% for lubrication; 88% for orgasm; and 86% for satisfaction. To summarize, when compared directly to a WLC group, the MI resulted in greater changes in desire, lubrication, and satisfaction, whereas when simply examining changes over time for those undertaking this intervention, changes were evident in desire, lubrication, satisfaction, orgasm and arousal.

Findings for the Hall (2004) book reveal that when compared directly to the WLC group, the intervention resulted in greater changes in desire (as measured by the HISD), whereas when simply examining changes over time for those reading the book, changes were evident in desire, lubrication, and orgasm. To explain, when comparing the HI to the WLC group, only one effect (desire) likely not obtained by chance was obtained, with this effect being large. The common language effect size indicated a 73% likelihood that those reading the book would score higher on the HISD at posttest than those in the wait-list control group. Additionally, when examining pretest to posttest within group effect sizes, results not likely obtained by chance included a large effect for desire and orgasm, and a medium effect for lubrication. The percent likelihood that those reading the book would score higher at posttest than pretest was 92% for desire, 86% for orgasm, and 77% for lubrication. Thus, this study provided initial evidence of the effectiveness of this book for increasing sexual functioning in women distressed by low desire.

Answers regarding the comparative efficacy of the two books are not clear-cut. In comparing the pattern of results found for the two books, one would conclude that the MI was more effective than the HI. For example, when compared to the WLC group directly, the MI resulted in large effects for three dimensions of sexual functioning (desire, lubrication, and orgasm), whereas the HI resulted in a large effect for desire only. Likewise, when examining pretest to posttest within-group effects, the MI yielded four large effects (desire, lubrication, satisfaction, and orgasm) and one medium effect (arousal) and the HI yielded two large effects (desire and orgasm) and one medium effect (lubrication). Nevertheless, it should be noted that for the primary outcome of the intervention—desire—both books performed equally well, lead-

ing to the conclusion that both books had the intended effect with one book (Mintz) having more additional positive effects than the other. However, this conclusion is not supported when calculating between-group posttest effect sizes comparing the interventions: While both common language and raw number effect sizes indicated that the MI outperformed the HI, confidence limits indicated that these results could have been obtained by chance. In sum, while it does appear that both books were similarly effective for enhancing the main outcome of desire, whether or not there is differential efficacy for other domains of sexual functioning is equivocal.

Despite this, the clinical implications of this study are clear. Results lend credence to the notion that bibliotherapy can be considered a good first choice in a stepped-care model for treating low sexual desire. Additionally, these results lead to the tentative conclusion that scientifically and clinically sound books written from diverse perspectives can be equally helpful in increasing desire. Thus, perhaps it is the act of doing something to help oneself (i.e., reading a book), rather than specifically what is done (i.e., the exact content of the book) that is most helpful. A study comparing one or both books to a placebo medication would test this notion.

Additional study regarding the longer-term effectiveness of these books is also warranted. In both the prior RCT (Mintz et al., 2012) and this RCT, longer-term findings are considered exploratory due to small samples. Nevertheless, in both the prior RCT on the MI and for both the MI and the HI tested in this study, gains in sexual desire were maintained at follow-up. Indeed, there was an 88% (MI) and 74% (HI) likelihood that someone reading these respective books would score higher on a measure of desire at follow-up than at pretest. Although space constraints limit a detailed discussion of maintenance of gains for other dimensions of sexual functioning, a few intriguing findings deserve mention. Among those reading the Mintz (2009) book, for pain and arousal, pretest to follow-up effect sizes were of a larger magnitude than the pretest to posttest effect sizes, indicating that these aspects of sexual functioning continued to improve after finishing the book. Also, results indicate that the MI performed better over time than did the HI. To explain, in terms of results not likely due to chance, the MI yielded six large pretest to follow-up effect sizes (HISD, FSFI desire, arousal, lubrication, orgasm, and pain subscales) and the HI yielded one large and one medium pretest to follow-up effect sizes (i.e., HISD and FSFI desire subscale, respectively).

While medium to large pretest to follow-up effects for the FSFI Desire subscale were found among those completing follow-up measures, all effect sizes for this same subscale among those only completing pretest and posttest measures were found to be obtained by chance. This may be because the FSFI assesses responses over the last four weeks, and thus may have been insensitive to changes following a 6-week reading period. If this is the case, then one might speculate that for the other FSFI subscale effect sizes not obtained by chance, the effects of the intervention were obtained within the first 2 weeks of the 6-week intervention. It thus may be wise for future intervention studies using the FSFI to change the timeframe from “in the last 4 weeks,” to “Since reading this book” or the like. Of course, such changes to the stem would require additional psychometric data collected on this scale. Additionally, since the FSFI is the most widely used measure of sexual functioning in intervention studies, these findings under-

score the importance of including confidence limits for effect sizes when using this measure, as without them it is uncertain as to whether results were obtained by chance or not.

Along with this measurement issue, other methodological issues need to be considered. For example, as outlined by van Lankveld (2009), the effects of pretreatment assessment may have influenced the results, leading both books to perform better in this study than they might if read in a naturalistic context. As is the case with most studies on bibliotherapy (van Lankveld, 2009), the sample size was small and diversity was lacking; as noted, however, the former was mitigated by use of Hedges g effect sizes and confidence limits. Also, in this study, there was a differential attrition rate of the two books from pretest to posttest and from posttest to follow-up. Specifically, more of those in the MI dropped out between pretest and posttest and more of those in the HI dropped out between posttest and follow-up. Nevertheless, by the end of the study, the sample sizes were roughly equivalent, leading to the conclusion that both books were equally well-received or held equal attention among participants over the course of the study. Nevertheless, future studies could provide additional incentives for study completion, so that differences between conditions and longer term effectiveness can be more carefully examined.

Other potentially useful studies to follow-up on the results obtained here include comparing one or both of these bibliotherapy interventions to face-to-face treatment. Also informative would be a study comparing bibliotherapy to Internet-based self-help. Similarly, a study comparing reading with no support (pure self-help) versus reading with therapist support (guided self-help) would be of interest. Nevertheless, even without such additional studies, psychologists can now feel comfortable recommending both books to women struggling with low sexual desire, with perhaps additional confidence in the Mintz book when improving aspects of sexual functioning beyond desire is an additional goal. Whether these books would be equally helpful to women with clinician-diagnosed sexual disorders, rather than with self-reported low sexual desire, is an empirical question. Nevertheless, given that most women struggle with low desire in the context of good overall mental health and good relationships (Leiblum, 2010), that these women are likely to seek out self-help (Rosen et al., 2009), and that this is the population for which both books were written, establishing efficacy in this context is most relevant.

The Mintz book can be purported to be “probably efficacious” given that it now has “two experiments showing the treatment is more effective than a waiting-list control group” (Chambless et al., 1998, p. 8). The Hall book is one step closer to this same designation, as there is now one study showing it is more effective than a WLC group. However, due to small samples in both studies and despite the use of statistics that disentangle sample and effect size, it would be wise to continue to study both interventions for added certainty of efficacy designations. Statements about efficacy designations are being made despite criteria for empirically supported treatments only previously being applied to face-to-face treatment. We contend that it is time to evaluate bibliotherapy, and other forms of self-help, by the same standards with which we evaluate face-to-face treatments. The time has come to apply standards for evidence based practice to self-help. It is hoped that this study will set the tone for this important step.

Counseling psychologists are uniquely qualified to lead this charge. First, a substantial percentage of self-help materials focus

on a mainstay of counseling psychology: “. . . treating normal or developmentally appropriate challenges and transitions” (Mallen & Vogel, 2005, p. 917). Second, counseling psychology’s strength-based, developmental model in which psychoeducation is a key intervention (Gerstein, 2006) is consistent with the stepped care model, in which self-help is often the first step (Norcross, 2006). Third, we have a long history of providing scientifically validated clinical services to a wide range of populations (Mallen & Vogel, 2005). Fourth, a benefit of self-help is the provision of services to those who might otherwise go unserved (Norcross, 2006), thus aligning with the social justice orientation of our field. Finally, counseling psychologists are already integrating self-help into their work: Norcross and colleagues (2000) found that out of 1,229 clinical and counseling psychologists surveyed, 85% had recommended self-help books to clients in the last year. In short, the philosophical underpinnings and interventions of counseling psychology map well onto the study of the efficacy of self-help in general and bibliotherapy in specific, yet our voices have been hitherto relatively silent in this arena.

It is hoped that this research will spawn additional investigation on the efficacy of self-help for a multitude of concerns, as well as comprise the first step in applying evidence-based standards to self-guided treatments.

References

- American Psychological Association. (2002). Criteria for evaluating treatment guidelines. *American Psychologist*, *57*, 1052–1059. <http://dx.doi.org/10.1037/0003-066X.57.12.1052>
- Annon, J. S. (1974). *Behavioral treatment of sexual problems*. Honolulu, HI: Enabling Systems.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*, 271–285. <http://dx.doi.org/10.1037/0003-066X.61.4.271>
- Apt, C., & Hurlbert, D. F. (1992). Motherhood and female sexuality beyond one year postpartum: A study of military wives. *Journal of Sex Education & Therapy*, *18*, 104–114.
- Basson, R., & Brotto, L. A. (2009). Disorders of sexual desire and subjective arousal in women. In R. Balon & R. T. Segraves (Eds.), *Clinical manual of sexual disorders* (pp. 119–159). Arlington, VA: American Psychiatric Publishing.
- Beck, J. G. (1995). Hypoactive sexual desire disorder: An overview. *Journal of Consulting and Clinical Psychology*, *63*, 919–927. <http://dx.doi.org/10.1037/0022-006X.63.6.919>
- Brotto, L. A., Basson, R., & Luria, M. (2008). A mindfulness-based group psychoeducational intervention targeting sexual arousal disorder in women. *Journal of Sexual Medicine*, *5*, 1646–1659. <http://dx.doi.org/10.1111/j.1743-6109.2008.00850.x>
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., . . . Woody, S. R. (1998). Update on empirically validated therapies. *Clinical Psychologist*, *51*, 3–16.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Taylor & Francis.
- Gerstein, L. H. (2006). Counseling psychology’s commitment to strengths: Rhetoric or reality? *The Counseling Psychologist*, *34*, 276–292. <http://dx.doi.org/10.1177/0011000005283518>
- Hall, K. (2004). *Reclaiming your sexual self: How you can bring desire back into your life*. Hoboken, NJ: Wiley.
- Hurlbert, D. F. (1993). A comparative study using orgasm consistency training in the treatment of women reporting hypoactive sexual desire.

- Journal of Sex & Marital Therapy*, 19, 41–55. <http://dx.doi.org/10.1080/00926239308404887>
- Joannides, P. (2012). *Guide to getting it on: A book about the wonders of sex*. Waldport, OR: Goofy Foot Press.
- Kline, R. B. (2013). *Beyond significance testing: Statistics reform in the behavioral sciences* (2nd ed.). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14136-000>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <http://dx.doi.org/10.3389/fpsyg.2013.00863>
- Leiblum, S. R. (2010). *Treating sexual desire disorders: A clinical casebook*. New York, NY: The Guilford Press.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education intervention into more readily interpretable forms* (NCSE, 2013–3000). Washington, DC: National Center for Special Education Research, U.S. Department of Education.
- Mallen, M. J., & Vogel, D. L. (2005). Online counseling: The need for discovery. *The Counseling Psychologist*, 33, 910–921. <http://dx.doi.org/10.1177/0011000005280182>
- McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, 111, 361–365. <http://dx.doi.org/10.1037/0033-2909.111.2.361>
- McKendree-Smith, N. L., Floyd, M., & Scogin, F. R. (2003). Self-administered treatments for depression: A review. *Journal of Clinical Psychology*, 59, 275–288. <http://dx.doi.org/10.1002/jclp.10129>
- Meyer-Bahlburg, H. F. L., & Dolezal, C. (2007). The female sexual function index: A methodological critique and suggestions for improvement. *Journal of Sex & Marital Therapy*, 33, 217–224. <http://dx.doi.org/10.1080/00926230701267852>
- Mintz, L. (2009). *A tired woman's guide to passionate sex*. Avon, MA: Adam's Media.
- Mintz, L. B., Balzer, A. M., Zhao, X., & Bush, H. E. (2012). Bibliotherapy for low sexual desire: Evidence for effectiveness. *Journal of Counseling Psychology*, 59, 471–478. <http://dx.doi.org/10.1037/a0028946>
- Morokoff, P. J., & LoPiccolo, J. (1986). A comparative evaluation of minimal therapist contact and 15-session treatment for female orgasmic dysfunction. *Journal of Consulting and Clinical Psychology*, 54, 294–300. <http://dx.doi.org/10.1037/0022-006X.54.3.294>
- Norcross, J. C. (2006). Integrating self-help into psychotherapy: 16 practical suggestions. *Professional Psychology: Research and Practice*, 37, 683–693. <http://dx.doi.org/10.1037/0735-7028.37.6.683>
- Norcross, J. C., Santrock, J. W., Campbell, L. F., Smith, T. S., Sommer, R., & Zuckerman, E. L. (2000). *Authoritative guide to self-help resources in mental health*. New York, NY: Guilford Press.
- Rosen, R., Brown, C., Heiman, J., Leiblum, S., Meston, C., Shabsigh, R., . . . D'Agostino, R., Jr. (2000). The Female Sexual Function Index (FSFI): A multidimensional self-report instrument for the assessment of female sexual function. *Journal of Sex & Marital Therapy*, 26, 191–208. <http://dx.doi.org/10.1080/009262300278597>
- Rosen, R., Shifren, J., Monz, B., Odom, D., Russo, P., & Johannes, C. (2009). Correlates of sexually related personal distress in women with low sexual desire. *Journal of Sexual Medicine*, 6, 1549–1560. <http://dx.doi.org/10.1111/j.1743-6109.2009.01252.x>
- Teachman, B. A. (2014). No appointment necessary: Treating mental illness outside the therapist's office. *Perspectives on Psychological Science*, 9, 85–87. <http://dx.doi.org/10.1177/1745691613512659>
- Trudel, G., Marchand, A., Ravart, M., Aubin, S., Turgeon, L., & Fortier, P. (2001). The effect of a cognitive-behavioral group treatment program on hypoactive sexual desire in women. *Sexual and Relationship Therapy*, 16, 145–164. <http://dx.doi.org/10.1080/14681990120040078>
- Turner, H. M., & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders*, 33, 43–55.
- van Lankveld, J. (2009). Self-help therapies for sexual dysfunction. *Journal of Sex Research*, 46, 143–155. <http://dx.doi.org/10.1080/00224490902747776>
- van Lankveld, J. J., ter Kuile, M. M., de Groot, H. E., Melles, R., Nefs, J., & Zandbergen, M. (2006). Cognitive-behavioral therapy for women with lifelong vaginismus: A randomized waiting-list controlled trial of efficacy. *Journal of Consulting and Clinical Psychology*, 74, 168–178. <http://dx.doi.org/10.1037/0022-006X.74.1.168>
- Wampold, B. E., Lichtenberg, J. W., & Waehler, C. A. (2002). Principles of empirically supported interventions in counseling psychology. *The Counseling Psychologist*, 30, 197–217. <http://dx.doi.org/10.1177/0011000002302001>
- West, S. L., D'Aloisio, A. A., Agans, R. P., Kalsbeek, W. D., Borisov, N. N., & Thorp, J. M. (2008). Prevalence of low sexual desire and hypoactive sexual desire disorder in a nationally representative sample of U.S. women. *Archives of Internal Medicine*, 168, 1441–1449. <http://dx.doi.org/10.1001/archinte.168.13.1441>

Received June 16, 2014

Revision received January 8, 2015

Accepted January 8, 2015 ■