View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by espace@Curtin

Running head: TRANSPORTABILITY OF IMAGERY-ENHANCED CBT

Transportability of imagery-enhanced CBT for social anxiety disorder

Peter M. McEvoy[a, b], David M. Erceg-Hurn[a, b], Kevin C. Barber[c], Jessica R. Dupasquier[c], & David A. Moscovitch[c]

[a] Centre for Clinical Interventions, Perth, Australia

[b] School of Psychology, Curtin University, Perth, Australia

[c] Department of Psychology, Centre for Mental Health Research, University of Waterloo, Canada

ACCEPTED MANUSCRIPT

Tables: 4

Figures: 4

Correspondence concerning this article should be addressed to Peter M. McEvoy, Ph.D., School of Psychology, GPO Box U1987, 6845, Australia. Phone: +61 8 9266 5110. Fax: +618 9226 2464. Email: peter.mcevoy@curtin.edu.au

Abstract

Pilot and open trials suggest that imagery-enhanced group cognitive behaviour therapy (CBT) is highly effective for social anxiety disorder (SAD). However, before being considered reliable and generalisable, the effects of the intervention need to be replicated by clinicians in a setting that is independent of the protocol developers. The current study compared outcomes from clients with a principal diagnosis of SAD at the Australian clinic where the protocol was developed (n = 123) to those from an independent Canadian clinic (n = 46) to investigate whether the large effects would generalise. Trainee clinicians from the independent clinic ran the groups using the treatment protocol without any input from its developers. The treatment involved 12 two-hour group sessions plus a one-month follow-up. Treatment retention was comparable across both clinics (74% vs. 78%, ≥9/12 sessions) and the between-site effect size was very small and non-significant on the primary outcome (social interaction anxiety, d = .09, p = .752). Within-group effect sizes were very large in both settings (ds = 2.05 vs. 2.19), and a substantial minority (41%-44%) achieved clinically significant improvement at follow-up. Replication of treatment effects within an independent clinic and with trainee clinicians increases confidence that outcomes are generalisable.

Key Words: imagery; cognitive behaviour therapy; social anxiety disorder; effectiveness; group; transportability

Transportability of imagery-enhanced CBT for social anxiety disorder

Social anxiety disorder (SAD) is one of the most common and earliest onset anxiety disorders (McEvoy, Grove, & Slade, 2011), and it is characterised by marked and persistent anxiety when exposed to potential scrutiny by others (American Psychiatric Association, APA, 2013). Cognitive behavioural therapy (CBT) is the recommended psychological treatment for SAD (National Institute for Health and Care Excellence, NICE, 2013) because it is both efficacious in research settings (Clark et al., 2003) and effective in real world clinical practice (McEvoy, Nathan, Rapee, & Campbell, 2012). Recent pilot and open trials of a novel imagery-enhanced group CBT protocol in a community mental health clinic yielded very large effect sizes that were comparable to individual CBT (McEvoy & Saulsman, 2014; McEvoy, Erceg-Hurn, Saulsman, & Thibodeau, 2015). However, before clinicians adopt a new treatment, it is important to demonstrate that the effects are generalisable to independent clinical services without direct supervision or input from the protocol developers. Evidence that research teams who report novel findings more frequently replicate such findings compared to independent researchers (Makel, Plucker, & Hegarty, 2012) suggests it is critical that effects are independently replicated before concluding that they are transportable to other settings (Tackett et al., 2017).

Cognitive theorists suggest that individuals with SAD hold assumptions that others are likely to be critical and judgmental, and that there is a high cost to negative evaluation (Heimberg, Brozovich, & Rapee, 2014). Central to these models is that individuals with SAD form a mental representation of the self, creating a vivid impression of how others view them from an observer perspective. This mental representation is guided by internal cues (e.g., symptoms of anxiety such as warmth in cheeks), memories of previous social experiences, and observable feedback (e.g., ambiguous feedback such as yawns, perceived as signs of boredom rather than tiredness). When this internal self-representation is then compared to beliefs about the standards expected by the audience, individuals with SAD

invariably believe their performance falls woefully short. The perceived discrepancy between actual and expected performance then leads to a cascade of negative thoughts and emotions, as well as avoidance behaviours, which ultimately reinforce and maintain the negative self-image and consequently the perceived social threat. A distorted negative self-image is therefore a key maintenance factor of social anxiety (Heimberg et al., 2014), which is targeted in imagery-enhanced CBT by encouraging clients to first identify negative predictions related to the self and others within mental imagery. Negative mental images are then used as predictions within behavioural experiments which, in turn, are designed to maximally violate the vivid expectancies (McEvoy, Saulsman, & Rapee, 2018). Other key therapy components include video feedback, which is used to directly challenge mental imagery of the anxious self, and imaginal (or imagery) rescripting, during which clients are guided within their imagination to relive and reappraise socially painful memories that encapsulate core negative beliefs about self and others.

Heimberg et al.'s (2014) model of SAD also emphasises the potency of modifying affect by working within the imagery mode rather than the verbal mode. Imagery involves multisensory-perceptual representations that can have visual, somatic, auditory, olfactory, and/or gustatory elements, and which have particularly strong links to both positive and negative emotions (Holmes & Mathews, 2010). Negative mental imagery is common in socially anxious individuals and increases anxiety, avoidant behaviours, self-focused attention, negative self-appraisals, and social performance deficits (e.g., Hirsch, Meynen, & Clark, 2004; Makkar & Grisham, 2011). Experimental research has also demonstrated that mental imagery activates positive and negative affect more potently than verbal-linguistic activity (Holmes & Matthews, 2010), suggesting that working within the imagery mode in therapy may result in larger affective change. Therefore, imagery-enhanced CBT incorporates mental imagery into all therapy components that are designed to modify six key

maintaining factors: negative thoughts and images, avoidance, safety behaviours, negative self-images, self-focused attention, and negative core beliefs (McEvoy & Saulsman, 2014).

Evidence for imagery-enhanced CBT for SAD is currently limited to a pilot study (N = 19, McEvoy & Saulsman, 2014) and an open trial (McEvoy et al., 2015) comparing a sample receiving imagery-enhanced CBT (n = 53) to historical controls (n = 129) who received the protocol without the imagery-enhancements. These preliminary studies demonstrated high retention (around 90% of clients receiving ≥9 of 12 sessions), very large effect sizes (*d*s~2.0), and a substantial minority of clients achieving normative functioning (~40%, McEvoy et al., 2015). While these early findings raise hopes for improved outcomes compared to alternative protocols, and are particularly impressive given that they were achieved through group CBT which is half as costly per patient to deliver as individual CBT (Mavranezouli et al., 2015), the transportability of imagery-enhanced group CBT to independent clinical settings is currently unknown.

The aim of the present study was to benchmark outcomes achieved from imagery-enhanced group CBT administered at an independent service to those observed at the clinic in which the protocol was originally developed. To maximise independence, the protocol was shared with clinicians who were based in a different country and who did not receive supervision, guidance, or consultation from any of the protocol developers. Graduate students, predoctoral residents, and postdoctoral fellows in clinical psychology administered the protocol at the independent site. The use of trainee therapists, compared to the more experienced therapists within the development clinic, provided a particularly rigorous test of whether the effects would replicate. Outcomes from the independent clinic were hypothesised to compare favourably to those from the development clinic with respect to patient retention, effect sizes, trajectories of change, and the proportion of clients achieving reliable and clinically significant change.

Method

Participants

Development clinic. Participants were 123 consecutive referrals to an Australian community mental health clinic (Centre for Clinical Interventions) from health professionals (general medical practitioners, psychiatrists, psychologists). Inclusion criteria: (a) principal Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 1994) SAD diagnosis, (b) not currently actively suicidal, self-harming, or psychotic, and (c) non-treatment-interfering substance use. The Mini International Neuropsychiatric Interview (MINI PLUS 5.0; Sheehan et al., 2001) was administered by masters- or doctorate-level clinical psychologists. Written informed consent was provided and approval was received from the Hospital's Human Research Ethics Committee (QI 2014_05). Development clinic participants were mostly single (79%) and 52% were unemployed. Around one quarter (27%) never completed high school, 36% had finished high school, 7% had a trade certificate, and 30% had completed a University degree. The most common comorbidities were Major Depression (45%), Generalised Anxiety Disorder (33%) and Dysthymic Disorder (7%).

Independent clinic. Participants were 46 clients who presented for treatment for social anxiety to the University of Waterloo Centre for Mental Health Research (CMHR), an outpatient psychology training clinic that provides mental health services to students and members of the general public. Participants were referred by mental health providers in the community or university counselling services, or self-referred in response to advertisements. Inclusion criteria: (a) principal diagnosis of SAD based on the Mini Neuropsychiatric Interview for DSM-5 (MINI 7.0.0; Sheehan, 2014) administered by trained graduate-level or postdoctoral clinicians under the supervision of the last author (DM), and (b) not endorsing active and interfering suicidality or self-harm, mania, psychosis, or substance use. Written informed consent was provided and approval was received from the University's Human Research Ethics Committee (#21956). The independent clinic sample primarily comprised

university students (85%). The most common comorbidities were Major Depressive Disorder (17%), Generalised Anxiety Disorder (17%), and Persistent Depressive Disorder (9%). See Tables 1 and 2 for more details.[1]

Outcome Measures

Social Interaction Anxiety Scale (SIAS) and Social Phobia Scale (SPS). The SIAS and SPS (Mattick & Clarke, 1998) are 20-item measures of interaction and performance anxiety, respectively. The SIAS was the primary outcome and assesses cognitive, affective, and behavioural reactions to interaction situations. The SPS describes situations in which the person is observed by others. The 5-point response scale for both scales is *Not at all*, *Slightly*, *Moderately*, *Very*, or *Extremely* characteristic of me. Total scores range from 0 to 80. These scales have high twelve-week test-retest reliabilities (SIAS r = .92; SPS r = .93, Mattick & Clarke, 1998) and are sensitive to change (Cox, Ross, Swinson, & Direnfeld, 1998). In the current sample, McDonald's omega composite reliability coefficients[2] were high (SIAS ω= .84, SPS ω= .90).

Brief Fear of Negative Evaluation-Straight Forwardly Worded (BFNE-S, Rodebaugh et al., 2004). The BFNE-S is an 8-item self-report measure of fear and concern about negative evaluation from others on a 5-point response scale, *Not at all, Slightly, Moderately, Very,* or *Extremely* characteristic of me. Total scores range from 8 to 40. The BFNE-S has demonstrated high reliability (α = .92) and construct validity in clinical samples (Weeks et al., 2005). In the current sample the composite reliability was high (ω= .83).

Depression, Anxiety and Stress Scale (DASS-21). The DASS-21 (Lovibond & Lovibond, 1995) measures emotional symptoms over the previous week and has excellent psychometric properties in psychiatric settings (Page, Hooke, & Morrison, 2007). Total scores range from 0 to 126. Composite reliability was very high (ω = .88).

Procedure & Treatment

The mean number of clients per group was 8.05 (*SD* = 1.99), with 14 and 7 groups run at the development and independent sites, respectively. Treatment comprised 12 weekly 2-hour sessions plus a one-month follow-up. The independent clinic started groups in April and October of each calendar year. October groups (n = 3) included a four-week break over December and January when the clinic closed. Treatment integrity was facilitated by a manual with comprehensive therapist instructions, patient handouts, and worksheets. Development clinic groups were co-facilitated by two masters- or doctoral-level clinical psychologists or a clinical psychologist and intern. Independent clinic groups were run by two to three doctoral students or postdoctoral fellows in clinical psychology under the direct supervision of a licensed clinical psychologist with expertise in treating SAD (DM). Development clinic supervision involved weekly discussions with co-facilitators, reviews during a weekly clinical team meeting, and ad hoc discussions with a nominated clinical supervisor when required. Independent clinic supervision involved live observation or video review of each session followed by weekly 1-hour supervision meeting with the therapists to discuss group content and process. No formal assessment of protocol adherence was measured. The SIAS and SPS were administered at sessions 1, 4, 8, 12 and 1-month follow-up, whereas the BFNE and DASS were administered prior to each session.

The imagery-enhanced CBT protocol was modified from Rapee, Gaston, and Abbott's (2009) group CBT manual by incorporating imagery-based strategies throughout (see McEvoy et al., 2018, for a detailed description). The protocol targets negative social-evaluative thoughts and images, avoidance, safety behaviours, negative self-images, self-focused attention, and negative core beliefs. Early sessions focus on socialisation to the model, and identifying and challenging negative social images about the past, present and future. Sessions then involve behavioural experiments to challenge negative social-evaluative images experientially while reducing avoidance and the use of safety behaviours. Various within- and between-session exercises challenge the probability and cost of negative

evaluation, as well as self-images (via video-feedback). Attention training and focusing exercises reduce self-focused and environment-focused attention, and increase task-focused attention. Later sessions modify negative core beliefs via imaginal rescripting and promote long-term behavioural change via positive prospective imagery techniques. The one-month follow-up session involves a progress review, relapse prevention, and future goal-setting.

Data analysis

Baseline Comparisons and Treatment Retention. The extent to which participants at the two sites differed on continuous variables was assessed with Welch *t*-tests, and by evaluating the size of mean differences and standardised mean differences (i.e., Cohen's *d*). For categorical variables, differences in proportions were examined using Chi-square tests and by calculating confidence intervals using the Newcombe (2013) Hybrid-Score interval. Treatment retention was compared with respect to mean number of sessions attended, and the proportion of clients attending ≥9 sessions.

Outcome Analyses. Unadjusted analyses were used to examine whether the size of treatment effects and treatment trajectories were comparable across sites, irrespective of any baseline differences in client characteristics. These analyses were conducted using linear mixed-effect regression models. Focused contrasts were used to evaluate whether treatment effect sizes were comparable across sites. We estimated the mean change from baseline to the final time point (1 month-follow up) at the development and independent sites, and then tested whether there was a significant difference in mean change between the sites. Standardised effect sizes (i.e., Cohen's *d* values) were computed by dividing the unstandardised effects by the pooled pre-treatment standard deviation (Morris, 2008). The extent to which trajectories of change were comparable across sites was evaluated by examining plots of the trajectories and *p*-value for the time x site interaction. For reliable change (RC) and clinically significant change (CSC) analyses, we used the same criteria as McEvoy et al. (2015). The changes required to achieve RC on the SIAS and SPS were 8.84

and 10.66, respectively. The cutoff for CSC was defined as the mid-point between the means of clinical and normative samples (SIAS = 40.56, SPS = 31.61, Carleton et al., 2014).

Adjusted analyses were conducted to evaluate the comparability of treatment effect sizes and trajectories after controlling for baseline differences in client characteristics across sites. This was achieved by through a propensity score analysis (West et al., 2014). This involved matching the 46 independent site participants with 46 participants from the development site in such a way as to minimise baseline differences. Analyses were then rerun using a similar approach to the unadjusted analyses. See supplementary materials for a full explanation of unadjusted and adjusted analyses and for details about software, data, and code.

Results

Baseline Comparisons

Development clinic clients were older, more likely to be married and have at least one comorbidity, and had higher DASS scores. Scores on the SIAS, SPS, BFNE, and gender, did not significantly differ across sites (see Tables 1 and 2).

Treatment Retention

The average number of sessions attended was similar across sites (Independent site M = 9.98, SD = 3.03; Development site M = 9.63, SD = 3.63, difference = 0.35, 95% CI [-0.76, 1.44], d = 0.10). The percentage of clients receiving a high treatment dose (≥9 sessions) was also similar (Development site = 78.26%; Independent site = 73.98%, difference = 4.28%, 95% CI [-11.20, 16.94], p = .567).

Unadjusted Analyses

Within-site standardised changes were large, between-site differences were small and non-significant (Table 3, Figures 1 and 2), and the p-values for the time x site interaction were large (SIAS p = .989, SPS p = .845). BFNE-S and DASS trajectories were also similar across sites and the time x site interactions were non-significant (BFNE-S p = .051, DASS p

= .097). Eight clients had a pre-treatment SIAS score below the CSC cutoff and were therefore excluded from the RC and CSC analyses. Reliable improvement rates were high and similar across the sites (Development site = 79.47%, Independent site = 79.19%, difference = 0.28%, 95% CI [-14.04, 14.62], p = .948). A substantial minority of clients achieved CSC (Development site = 44.42%, Independent site = 40.77%, difference = 3.66%, 95% CI [-12.39, 19.70], p = .533). On the SPS, 30 clients had pre-treatment scores below the CSC cutoff and were excluded from analyses. Reliable improvement was high across the development site (90.04%) and independent clinic (87.83%), difference = 2.21%, 95% CI [-13.01, 17.43], p = .578. Rates of CSC were also high and similar across sites (Development site = 72.64%, Independent site 77.57%, difference = -4.93%, 95% CI [-18.03, 8.18], p = .331).

Adjusted Analyses

Propensity score matching was successful as standardised mean differences at baseline after matching were ≤ 0.08 on all variables (ps ≥ .72). Within-site standardised changes were large, between-site differences were small (Table 4), treatment trajectories were very similar (Figures 3 and 4), and p-values for the time x site interaction were large (p = .893 for SIAS and p = .706 for SPS). The treatment trajectory for the BFNE-S was similar across sites, except that there was a more pronounced drop in symptoms between sessions 11 and 12 in the independent sample (time by site interaction, p = .032). The DASS trajectories were similar (time x site interaction p = .336). Six clients were excluded from analyses for having pre-treatment SIAS scores below the CSC cutoff. Rates of reliable change were 80.07% at the development clinic and 79.40% at the independent site, difference = 0.67%, 95% CI [-14.72, 16.07], p = .887. CSC rates were 43.91% at the development clinic and 40.30% at the independent site, difference = 3.60%, 95% CI [-12.62, 19.83], p = .551. On the SPS, 21 clients were excluded due to pre-treatment scores below the CSC cutoff. Of the remaining clients, 87.92% reliably improved at the development clinic

and 88.20% at the independent clinic, difference = -.28, 95% CI [-19.59, 19.02], p = .948. Around three-quarters experienced CSC (Development clinic = 75.69%, Independent clinic = 77.34%, difference = -1.65%, 95% CI [-18.70, 15.40], p = .767.

Discussion

Imagery-enhanced group CBT integrates established evidence-supported techniques that target theory-driven maintaining factors for SAD (Heimberg et al., 2014; Rapee et al., 2009) with evidence from the emotion literature suggesting that modifying cognitions within the imagery mode may be more potent than predominantly working in the verbal mode (Holmes & Matthews, 2010). Preliminary outcomes have been promising, with high retention, large effect sizes, and improvements in the proportion of individuals achieving clinically significant change (McEvoy & Saulsman, 2014; McEvoy et al., 2015). However, it is critical to demonstrate that comparable effects can be achieved by clinicians who are independent from the protocol developers, which was the main aim of the current study. The hypothesis that the effectiveness of imagery-enhanced CBT would be transportable to an independent clinic when delivered by trainee therapists was supported.

Clients at the development site, which is a community mental health clinic, were more likely to have comorbid disorders and symptoms, and to be married and older, than clients the independent site, which is a clinical unit within a university setting. However, the samples did not differ in gender distribution or on social anxiety symptom severity. Despite the differences across the samples and the absence of any input from the development site, the outcomes in terms of treatment retention and symptom change were remarkably similar. Approximately three-quarters of clients across both settings received a ‘high dose’ of treatment. Within-group effect sizes at both sites were very large on changes in social interaction anxiety (Cohen’s ds ~ 2.0), and were large on changes in performance anxiety, general anxiety and depression, and fear of negative evaluation. Trajectories of change were very similar across both sites on all outcomes, as were rates of reliable and clinically

significant change. Approximately 80% of clients across both sites achieved reliable change and around 40% were within the normative range of social interaction anxiety at follow-up. These rates were even higher for performance anxiety (88% and 76%, respectively). The unadjusted analyses indicated that outcomes were comparable despite differences in features of the treatment context or clients. The pattern of findings when using propensity score matching indicated that comparable outcomes were achieved from similar clients across the treatment settings.

These findings strongly suggest that the outcomes from the development site are generalisable to other clinics, including psychology training settings with relatively inexperienced treating clinicians. It is important to note that the supervisor at the independent clinic is an expert in SAD (DM), but his role was restricted to that of clinical supervisor and, thus, it is impressive that the outcomes were virtually identical when the treatment manual was administered by trainee therapists from a geographically distinct team who received no input from the protocol developers[3]. It is important for future research to evaluate whether similar outcomes could be achieved when therapists deliver the intervention without supervision from a SAD specialist and in other contexts (e.g., private practice).

The within-group effect sizes are comparable to some of the largest effect sizes in the literature from both group and individual therapy. A recent meta-analysis found that group and individual CBT yielded standardised mean differences of 0.92 and 1.19 compared to waitlist control, respectively (Mayo-Wilson et al., 2014). Within-group effects for waitlists in CBT trials tend to be very small, thus the effect sizes from imagery-enhanced CBT appear to compare very favourably to treatments within the meta-analysis. Imagery-enhanced group CBT required an average of around six therapist hours per client, compared to up to around 20 hours for individual therapy (Clark et al., 2003). Although the evidence to date suggests that outcomes from imagery-enhanced group CBT compare favourably to individual CBT, a

direct comparison across group and individual formats is required to answer this question more definitively.

Several limitations must be considered. First, neither site included a waitlist control group nor a credible alternative intervention to control for confounds (e.g., regression to the mean) or non-specific factors of therapy (e.g., expectancies). The fact that clients in both samples met diagnostic criteria for social anxiety disorder and reported severe symptoms suggests that they were unlikely to remit over the 12-week intervention period, but controlled effects are likely to be somewhat smaller than the within-group effect sizes reported here. Second, although regular supervision was provided, treatment fidelity and therapist competence were not formally assessed. Although this issue means that we cannot ensure that all treatment components were delivered as prescribed in the manual, it is nonetheless reassuring that comparable effects were achieved across sites under these conditions, which provides some external validity for the findings. The manual includes very detailed therapist instructions and scripts, along with detailed client handouts and worksheets to guide each technique, which likely facilitated rapid training of new therapists at the independent clinic as well as a high degree of fidelity. Third, inter-rater reliability for pre-treatment diagnoses was not assessed, although self-reported symptom severity was very similar across sites at pre-treatment. Fourth, self-reported outcomes are susceptible to social desirability biases, so clinician-administered, behavioural, and psychophysiological measures would be informative.

The present study replicated the retention rates and large effect sizes from imagery-enhanced group CBT observed in the clinic that developed the treatment protocol within an independent clinic that received no input from the protocol developers. These findings increase confidence that imagery-enhanced CBT produces strong and transportable effects. Future research that directly compares imagery-enhanced group CBT to a credible alternative treatment is required to build the case that imagery-enhancements improve

outcomes beyond protocols that include some targeted imagery-based interventions (e.g., video-feedback) but do not comprehensively modify negative social-evaluative images more broadly within the imagery mode rather than the verbal mode. Comparisons between individual and group imagery-enhanced CBT would also be informative to ensure that comparable outcomes are achieved across treatment formats, or indeed whether effects may be even larger individually, and to directly compare cost-effectiveness.

Footnotes

[1] Ethnicity data were only collected for around half the patients in the Development Clinic, 76% of whom identified as Anglo/European Australian, 10% as Asian Australian, and 14% as "another ethnicity". Ethnicity data were not collected at the Independent Clinic.

[2] McDonald's omega is a generalisation of Cronbach's alpha psychometricians recommend be reported instead, primarily because alpha tends to underestimate the reliability of psychological measures. For an overview, see McNeish, D. (in press). Thanks coefficient alpha, we'll take it from here. *Psychological Methods.* doi: 10.1037/met0000144.

[3] DM contributed to the design of a current randomised controlled trial comparing imagery-enhanced to verbally-based group CBT, and is a co-author on that protocol paper (McEvoy et al., 2017). However, to date, DM has not met or spoken with the protocol developers. The protocol was provided electronically, delivered without supervision from the protocol developers, and all communication has been via email.

McEvoy, P. M., Moulds, M. L., Grisham, J. R., Holmes, E. A., Moscovitch, D. A., Hendrie, D., et al. (2017). Assessing the efficacy of imagery-enhanced cognitive behavioral group therapy for social anxiety disorder: study protocol for a randomized controlled trial. *Contemporary Clinical Trials, 60,* 34-41. doi.org/10.1016/j.cct.2017.06.010

Funding

This work was partially supported by the National Health and Medical Research Council [project grant number APP1104007]. The funding body had no influence on the study design, analysis, interpretation, writing, or decision to submit the article for publication.

Acknowledgements

The authors wish to express their gratitude to the clinicians who facilitated groups and to the clients who consented to their data being used for quality improvement and research purposes.

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington DC: APA.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington DC: APA.

Carleton, R. N., Thibodeau, M. A., Weeks, J. W., Teale Sapach, M. J. N., McEvoy, P. M., et al. (2014). Comparing Short Forms of the Social Interaction Anxiety Scale and the Social Phobia Scales. *Psychological Assessment, 26,* 1116-1126.

Clark, D. M., Ehlers, A., McManus, F., Hackmann, A., Fennell, M., Campbell, H., et al. (2003). Cognitive therapy versus fluoxetine in generalized social phobia: A randomized placebo-controlled trial. *Journal of Consulting and Clinical Psychology, 71*, 1058–1067.

Cox, B. J., Ross, L., Swinson, R. P., & Direnfeld, D. M. (1998). A comparison of social phobia outcome measures in cognitive-behavioral therapy. *Behavior Modification, 22,* 285-297.

Heimberg, R. G., Brozovich, F. A., & Rapee, R. M. (2014). A cognitive-behavioral model of social anxiety disorder. In S. G. Hofmann & P. M. DiBartolo (Eds.), *Social anxiety: Clinical, developmental, and social perspectives* (3rd ed., pp. 705–728). Waltham, MA: Academic Press.

Hirsch, C. R., Meynen, T., & Clark, D. M. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory, 12,* 496-506.

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review, 30,* 349-362.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Lovibond, S.H. & Lovibond, P.F. (1995). *Manual for the Depression Anxiety Stress Scales*. (2nd. Ed.) Sydney: Psychology Foundation.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: how often do they really occur? *Perspectives on Psychological Science, 7*, 537-542. DOI: 10.1177/1745691612460688

Makkar, S. R., & Grisham, J. R. (2011). Social anxiety and the effects of negative self-imagery on emotion, cognition, and post-event processing. *Behaviour Research and Therapy, 49,* 654–664.

Mallinckrodt, C., & Lipkovich, I. (2017). *Analyzing Longitudinal Clinical Trial Data: A Practical Guide*. Boca Raton: FL: CRC Press.

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36,* 455-470.

Mavranezouli, I., Mayo-Wilson, E., Dias, S., Key, K., Clark, D. M., Ades, A. E., et al. (2015). The cost effectiveness of psychological and pharmacological interventions for social anxiety disorder: a model-based economic analysis. *PLOS ONE*;10:e0140704

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., et al. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *Lancet Psychiatry, 1,* 368–376.

McEvoy, P. M., Erceg-Hurn, D. M., Saulsman, L. M., & Thibodeau, M. A. (2015). Imagery enhancements increase the effectiveness of cognitive behavioural group therapy for social anxiety disorder: a benchmarking study. *Behaviour Research and Therapy, 65*, 42-51. Doi: 10.1016/j.brat.2014.12.011.

McEvoy, P. M., Grove, R., & Slade, T. (2011). Epidemiology of Anxiety Disorders in the Australian General Population: Findings of the 2007 Australian National Survey of

Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry, 45,* 957-967.

McEvoy, P. M., Nathan, P., Rapee, R. M., & Campbell, B. N. C. (2012). Cognitive behavioural group therapy for social phobia: Evidence of transportability to community clinics. *Behaviour Research and Therapy, 50,* 258-265.

McEvoy, P. M., & Saulsman, L. (2014). Imagery-enhanced cognitive behavioural group therapy for social anxiety disorder: A pilot study. *Behaviour Research and Therapy, 55,* 1-6.

McEvoy, P. M., Saulsman, L. M., & Rapee, R. M. (2018). *Imagery-Enhanced CBT for Social Anxiety Disorder.* Guilford Press.

Morris, S. B. (2008). Estimating Effect Sizes From the Pretest-Posttest-Control Group Designs. *Organizational Research Methods, 11,* 364–386.

Newcombe, R. G. (2013). *Confidence intervals for proportions and related measures of effect size.* Boca Raton, FL, CRC Press.

National Institute for Health and Care Excellence (NICE). (2013). Social anxiety disorder: recognition, assessment and treatment. guidance.nice.org.uk/cg159.

Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology, 46,* 283-297.

Rapee, R. M., Gaston, J. E., & Abbott, M. J. (2009). Testing the efficacy of theoretically derived improvements in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 77,* 317-327.

Rodebaugh, T. L., Woods, C. M., Thissen, D. M., Heimberg, R. G., Chambless, D. L., & Rapee, R. M. (2004). More information from fewer questions: the factor structure and item properties of the original and brief fear of negative evaluation scale. *Psychological Assessment, 16,* 169-181. doi:10.1037/1040-3590.16.2.169

Sheehan, D., Janavus, J., Baker, R., Harnett-Sheehan, K., Knapp, E., & Sheehan, M. (2001) *M.I.N.I. Plus. (Mini International Neuropsychiatric Interview).* Version 5.0.

Sheehan. D. V. (2014). The Mini-International Neuropsychiatric Interview, Version 7.0 for DSM-5 (M.I.N.I 7.0). Medical Outcomes Systems.

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., et al. (2017). It's time to broaden the replicability conversation: thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, *12*, 742-756. Doi: 10.1177/17456916176900

Weeks, J. W., Heimberg, R. G., Fresco, D. M., Hart, T. A., Turk, C. L., Schneier, F. R., et al. (2005). Empirical validation and psychometric evaluation of the Brief Fear of Negative Evaluation Scale in patients with social anxiety disorder. *Psychological Assessment, 17,* 179–190.

West, S. G., Cham, H., Thoemmes, F., Renneberg, B., Schulze, J., & Weiler, M. (2014). Propensity scores as a basis for equating groups: Basic principles and application in clinical treatment outcome research. *Journal of Consulting and Clinical Psychology*, *82*, 906–919. doi:10.1037/a0036387

Table 1

Comparison of Categorical Variables at Pre-Treatment

	Site		Difference		
Variable	Development	Independent	Est	95% CI	*p*
At least 1 comorbidity (%)	73.98	56.52	17.46	1.76, 33.34	.029
Married (%)	9.48	0.00	9.48	0.75, 16.19	.031
Female (%)	58.54	54.35	4.19	-11.94, 20.62	.624

Note. Est = Point-estimate of difference in percentages; 95% CI = lower and upper limits of the 95% confidence interval around the difference; p = *p* value

Table 2

Comparison of Continuous Variables at Pre-Treatment

	Site				Unstandardized		Standardized		
	Development		Independent		Mean Difference		Mean Difference		
Variable	M	*SD*	M	*SD*	Est	95% CI	*d*	95% CI	*p*
Age (years)	28.54	10.76	22.87	5.98	5.67	3.08, 8.26	0.58	0.24, 0.93	<.001
DASS	61.52	21.38	51.98	24.75	9.54	1.20, 17.88	0.43	0.08, 0.77	.025
SPS	44.49	13.24	40.80	15.08	3.70	-1.38, 8.77	0.27	-0.07, 0.61	.151
SIAS	59.04	9.85	58.41	11.15	0.63	-3.13, 4.39	0.06	-0.28, 0.40	.739
BFNE-S	22.76	5.59	22.56	5.23	0.21	-1.66, 2.07	0.04	-0.31, 0.38	.825

Note. Est = Point-estimate of difference in percentages; 95% CI = lower and upper limits of the 95% confidence interval around the difference; p = *p* value for the difference in means; DASS= Depression, Anxiety and Stress Scale total score; SPS = Social Phobia Scale; SIAS = Social Interaction Anxiety Scale; BFNE-S = Brief Fear of Negative-Evaluation Straightforwardly worded version

Table 3

Unadjusted Analyses - Mean Change from Pre-treatment to Follow-Up

Measure		M	SE	*d*	*p*
SIAS					
	Development	21.06	1.49	2.07	
	Independent	20.18	2.37	1.98	
	Difference	0.89	2.80	0.09	.752
SPS					
	Development	22.82	1.42	1.65	
	Independent	19.90	2.24	1.44	
	Difference	2.91	2.65	0.21	.273
BFNE-S					
	Development	8.74	0.64	1.60	
	Independent	9.44	1.02	1.72	
	Difference	-0.69	1.21	-0.13	.565
DASS					
	Development	29.33	3.17	1.29	
	Independent	23.75	5.11	1.05	
	Difference	5.58	6.01	0.25	.353

Note. SIAS = Social Interaction Anxiety Scale; SPS = Social Phobia Scale; BFNE-S = Brief Fear of Negative-Evaluation Straightforwardly worded version; DASS= Depression, Anxiety and Stress Scale total score; Development = Development site; Independent = Independent site; Development and Independent site effects are within-group effects; Difference = Between-group effects; *d* = standardized effect sizes; *p* = *p* value for test of mean difference in outcome between sites. Pooled pre-treatment standard deviations used to compute the standardized effect sizes were 10.181 (SIAS), 13.808 (SPS), 5.474 (BFNE), and 22.689 (DASS).

Table 4

Adjusted Analyses - Mean Change from Pre-treatment to Follow-Up

Measure		M	SE	*d*	*p*
SIAS					
	Development	22.25	2.20	2.19	
	Independent	20.86	2.30	2.05	
	Difference	1.39	3.18	0.14	.663
SPS					
	Development	20.48	1.81	1.48	
	Independent	20.35	1.87	1.47	
	Difference	0.13	2.60	0.01	.961
BFNE-S					
	Development	9.11	0.94	1.66	
	Independent	9.62	0.97	1.76	
	Difference	-0.51	1.35	-0.09	.705
DASS					
	Development	22.25	3.05	0.98	
	Independent	24.42	3.16	1.08	
	Difference	-2.17	4.39	-0.10	.621

Note. SIAS = Social Interaction Anxiety Scale; SPS = Social Phobia Scale; BFNE-S = Brief Fear of Negative-Evaluation Straightforwardly worded version; DASS= Depression, Anxiety and Stress Scale total score; Development = Development site; Independent = Independent site; Development and Independent site effects are within-group effects; Difference = Between-group effects; *d* = standardized effect sizes; *p* = *p* value for test of mean difference in outcome between sites. Pooled pre-treatment standard deviations used to compute the standardized effect sizes were 10.181 (SIAS), 13.808 (SPS), 5.474 (BFNE), and 22.689 (DASS).

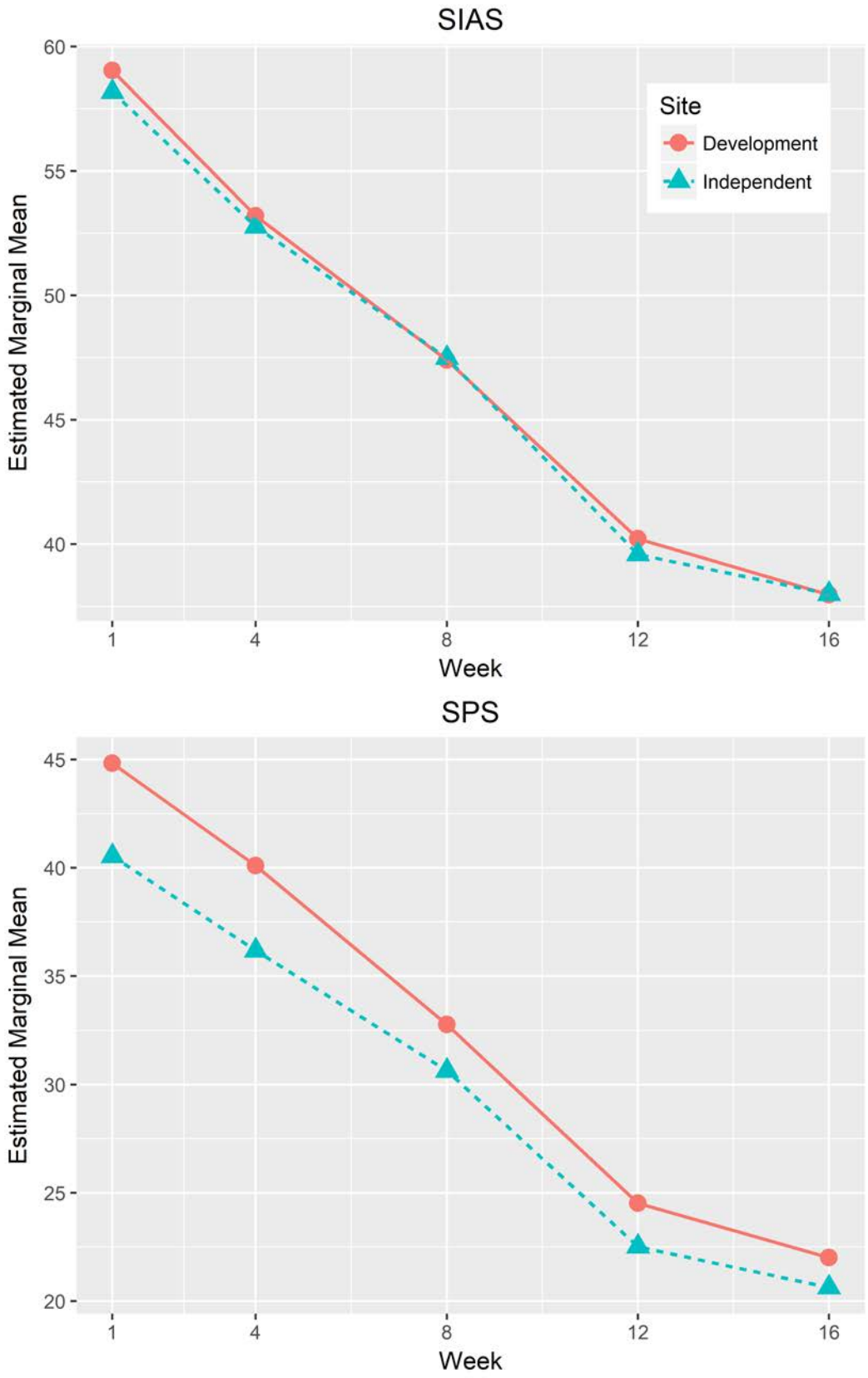

Figure 1. Unadjusted Analyses - Trajectories of Symptom Change for Social Interaction Anxiety Scale (SIAS) and Social Phobia Scale (SPS)

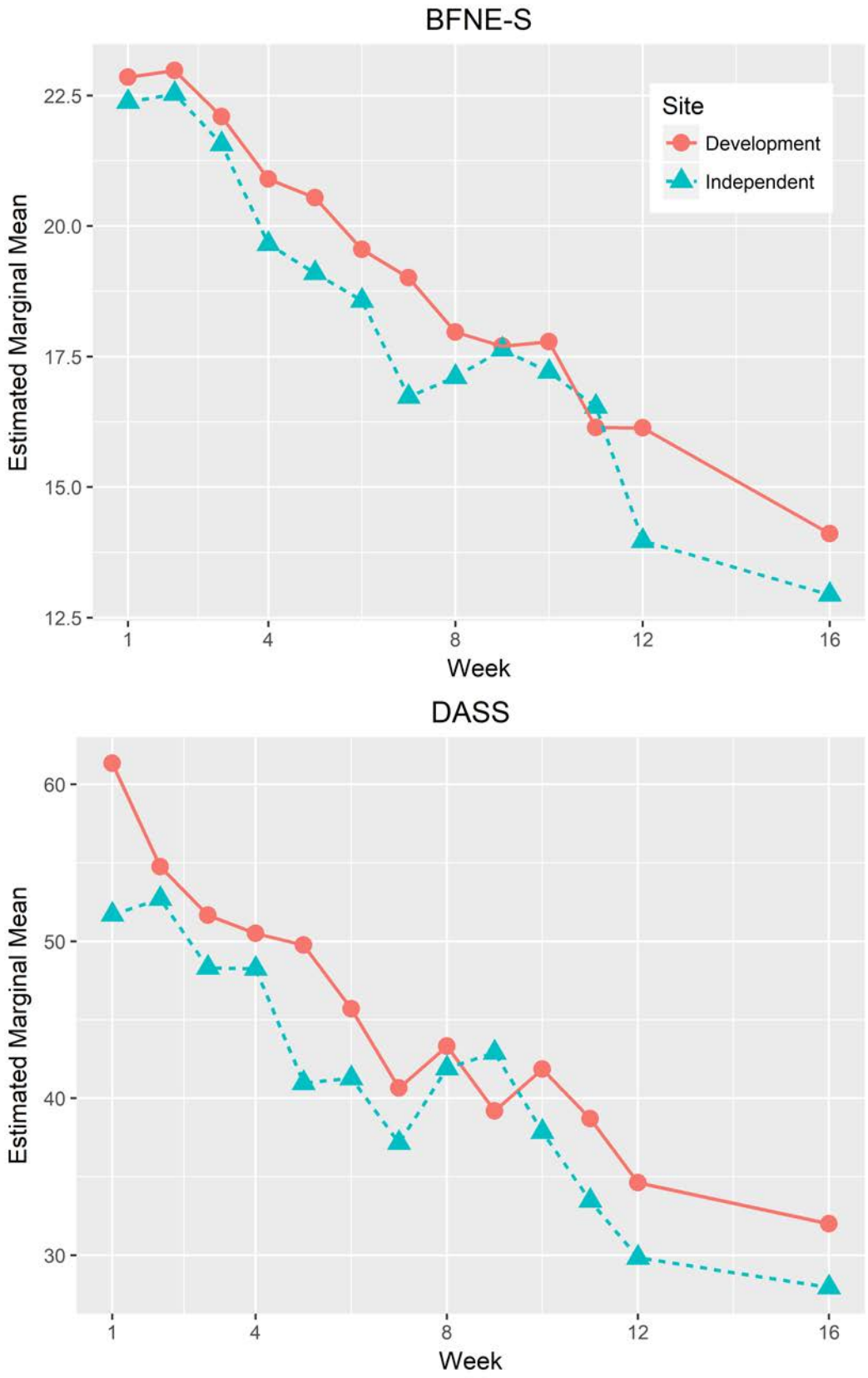

Figure 2. Unadjusted Analyses - Trajectories of Symptom Change for Brief Fear of Negative Evaluation (BFNE-S) and Depression, Anxiety, and Stress Scale (DASS)

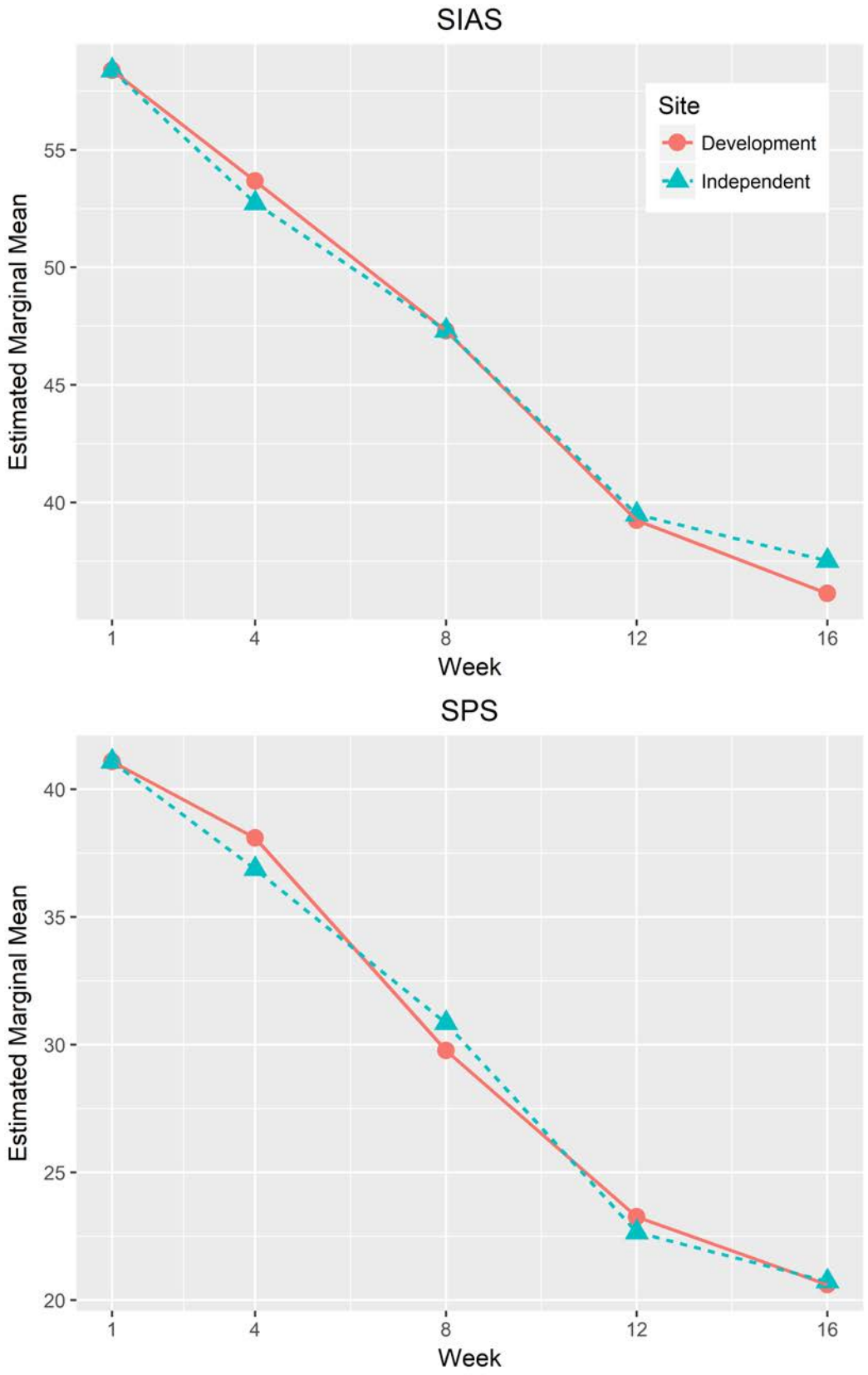

Figure 3. Adjusted Analyses - Trajectories of Symptom Change for Social Interaction Anxiety Scale (SIAS) and Social Phobia Scale (SPS)

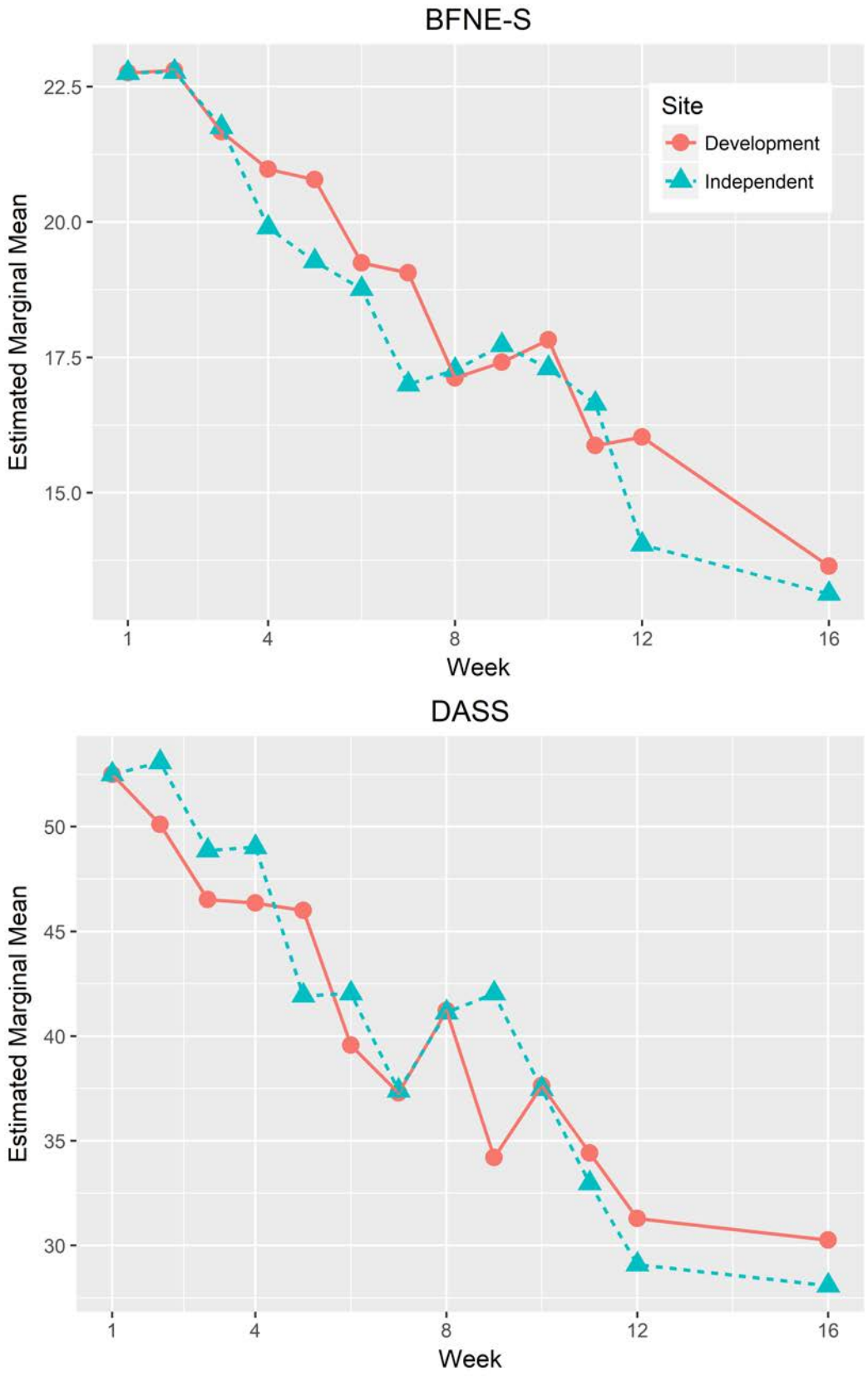

Figure 4. Adjusted Analyses - Trajectories of Symptom Change for Brief Fear of Negative Evaluation (BFNE-S) and Depression, Anxiety, and Stress Scale (DASS)