

# Researcher Allegiance as Hidden Moderator in Psychotherapy Outcome Research

Heike Gerger · Jens Gaab

Clinical Psychology and Psychotherapy, Department of Psychology, University of Basel, Basel, Switzerland

## Keywords

Psychotherapy outcome research · Researcher allegiance · Meta-analysis

## Summary

Randomized controlled trials that compare 2 rival treatments are used as the standard design to identify those constituents of complex psychotherapeutic treatments that critically impact symptom improvement. This paper examines the impact of unbalanced researcher allegiance, i.e. the researchers' preference regarding a particular treatment, as an extra-therapeutic moderator of the treatment effect. As researcher allegiance can only be measured at study level, meta-analytic strategies are required to investigate the association between researcher allegiance and outcome. Our review shows that researcher allegiance has a considerable impact on outcome. Despite the finding that researcher allegiance may impact outcome particularly in studies of poorer quality, other causal pathways need to be considered when trying to understand the relationship between unbalanced researcher allegiance and outcome. Although the association between unbalanced researcher allegiance and outcome is well established, the causal pathways that mediate this association are not well understood to date. Accordingly, we conclude that explicit statements by researchers regarding their preferences in a comparative outcome trial (e.g. by stating clear hypotheses or by including a conflict of interest statement) appear to be the most straightforward response to our findings in order to protect the validity of conclusions from psychotherapy outcome research.

## Schlüsselwörter

Psychotherapiewirksamkeitsforschung · Allegianz · Metaanalyse

## Zusammenfassung

Der direkte randomisierte Vergleich zwischen Behandlungen gilt als Standarddesign zur Ermittlung derjenigen Bestandteile komplexer psychotherapeutischer Behandlungen, die die Symptomverbesserung entscheidend beeinflussen. Die vorliegende Arbeit untersucht den Einfluss einseitiger Allegianz von Forschenden – d.h. deren Präferenz für eine bestimmte Behandlung – als extra-therapeutischen Moderator des Behandlungseffekts. Da die Allegianz von Forschenden nur auf Studienebene gemessen werden kann, sind metaanalytische Strategien erforderlich, um mögliche Zusammenhänge zwischen einseitiger Allegianz und Studienergebnissen zu untersuchen. In der Übersichtsarbeit soll gezeigt werden, dass eine einseitige Allegianz von Forschenden bedeutsam mit Studienergebnissen korreliert. Auch wenn dieser Zusammenhang bei verschiedenen Behandlungsansätzen und Störungen beobachtet werden kann, sind die Ursachen und verantwortlichen Prozesse noch nicht weitgehend bekannt. Aufgrund der Vielfalt an potenziellen kausalen Erklärungen für den beobachteten Zusammenhang zwischen Allegianz und Studienergebnissen schlussfolgern wir, dass die explizite Nennung der Allegianz ein möglicher und sinnvoller Schritt zur besseren Beurteilung von Schlussfolgerungen in der Psychotherapieforschung sein kann.

© 2016 S. Karger GmbH, Freiburg

## Introduction

The goal of clinical research is to ‘separate the therapeutic wheat from the chaff’ in order to ensure the safety and effectiveness of clinical treatments [Jones and Podolsky, 2015]. Based on a theoretical classification of treatment constituents, the aim of this article is to analyze the effect of extra-therapeutic ‘chaff’ based on the example of how the researchers’ preference of a particular treatment – the so-called researcher allegiance [Luborsky et al., 1975] – may impact outcome. As researcher allegiance can only be observed at study level, meta-analytic strategies are required to estimate the association between unbalanced researcher allegiance and outcome, thus defining individual studies included in the meta-analysis as the unit of analysis [Staines and Cleland, 2007].

The present paper does not answer the question of which treatment approach is most beneficial for a particular type of mental disorder. We also do not intend to evaluate individual randomized controlled trials (RCTs) with respect to their risk of bias. Rather, we focus on presenting observed associations on an aggregate level. We do not want to criticize any researchers or therapists for preferring a particular treatment – how could one not prefer a treatment if one, for instance, invented it or has experienced it to be effective? However, we would like to draw attention to the fact that such preferences, if they are not well controlled from the beginning of an RCT, may hamper the discovery of true differences between treatments [Lipsey and Wilson, 2001, p 158 ff]. Finally, although some of the research presented here focused on posttraumatic stress disorder (PTSD) outcome research, in which the diversity of available treatments with strong preferences of researchers makes the occurrence of allegiance effects very likely, it is important to note that the problems associated with unbalanced researcher allegiance may occur in any field where rival hypotheses are being tested [Cuijpers and Cristea, 2015; Ioannidis, 2005].

## Theoretical Framework for the Classification of Treatment Constituents

Treatment effects are typically described as the effects of the supposed verum plus the effects of the context in which the treatment is administered. The latter is usually controlled for by the placebo condition, while the former constitutes the treatment under investigation. The definition of what is to be considered verum or placebo treatment in medical research is facilitated by the treatment’s qualitative properties, with the former, for instance, being an active drug and the latter being a physiologically inert pill with ‘only’ psychological properties. In contrast, the operationalization of placebo control conditions in psychotherapy research is theoretically and practically flawed [Gaab et al., 2015; Kirsch, 2005; Weinberger, 2014]. It is important to note that although the outlined assumption of additivity is still the rule in outcome research, it has not yet been confirmed for as complex a treatment as psychotherapy. Furthermore, the exclusively additional understanding of placebo and verum effects has been challenged by recent evi-

dence [e.g. Lund et al., 2014], and most likely interaction effects exist between individual constituents of psychotherapy [Barber et al., 2006]. Thus, the classification of verum and placebo or ‘wheat’ and ‘chaff’ is difficult in psychotherapy research and may be at the root of ‘the culture wars in psychotherapy, (which) dramatically pit the treatment method against the therapy relationship’ [Norcross and Lambert, 2011, p. 4]. Nevertheless, the RCT design requires some definition of ‘wheat’ in order to experimentally manipulate its presence.

For matters of clarification, it is therefore helpful to employ a theory-based definition of what is to be considered ‘wheat’ and ‘chaff’ in psychotherapy. Here, Grünbaum proposed a definition that reflects the outlined dichotomy between different constituents of a complex treatment such as psychotherapy [Grünbaum, 1981, 1986]:

‘The therapeutic theory...that advocates the use of a particular treatment...to remedy (a given target disorder) demands the inclusion of certain *characteristic* constituents...in any treatment process that the (therapeutic theory) authenticates as an application of the (treatment). Any such process...will typically have constituents...other than the characteristic ones...which I shall denominate as ‘*incidental*’. Indeed, it may even maintain that one or another of the incidental factors affects the disorder’ [Grünbaum, 1981, p. 159].

For instance, exposure to the traumatic event in PTSD treatments, as described by Foa and colleagues [Foa et al., 1991], may be seen as *characteristic* (‘wheat’) because a treatment theory explains how exposure will reduce PTSD symptoms. In contrast, talking about daily problems, as in the placebo control condition by Foa and colleagues [Foa et al., 1991], would be considered *incidental* (‘chaff’) due to the lack of a theory-based link between such a treatment and symptom improvement. Nevertheless, *incidental* factors could influence outcome in a 2-way fashion:

First, *incidental* factors could be mediators of treatment effects. These factors have previously been described as essential but not unique treatment constituents [Waltz et al., 1993] and include, among others, empathic communication [Elliott et al., 2011] or patients’ outcome expectancies [Constantino et al., 2011; Kirsch, 2005].

Second, *incidental* factors could be moderators of treatment effects. Such factors are neither essential nor unique treatment constituents and may include intra-therapeutic aspects (e.g. meeting with a professional therapist in a secure context), as well as extra-therapeutic factors which are unrelated to the treatment or the therapeutic encounter. The latter may include a whole variety of factors related to the patient (e.g. symptom severity [Driessen et al., 2010]) and the therapist (e.g. therapist’s adherence to a manual [Barber et al., 2006]). In the context of clinical research, characteristics of the study itself (e.g. quality of outcome assessment) and the researcher (e.g. theoretical background) may function as extra-therapeutic moderators [Ioannidis, 2005].

While mediators of treatment effects may have a direct impact on treatment outcome, moderators presumably do not directly affect treatment outcome. At best, moderators indicate under which circumstances particular effects do or do not occur [as in Lynch et al., 2010], but at the worst they contribute to biased estimates of

treatment effects. Therefore, it is important to identify moderators of treatment effects and to control for the potential impact of such moderators in order to draw valid conclusions from study results.

### Researcher Allegiance as Hidden Moderator in PTSD Outcome Research

Two types of rival treatments are typically compared in one study in order to demonstrate the superiority of one over the other. In PTSD outcome research, for example, prolonged exposure plus cognitive restructuring was compared with exposure alone in 1 RCT in order to estimate the incremental effect of adding cognitive restructuring to the established exposure treatment [Foa et al., 2005]. This particular RCT failed to demonstrate superiority of adding cognitive restructuring to an exposure treatment; also, meta-analyses that summarized comparative RCTs of individual PTSD treatments found no statistically significant differences between the effects of 2 different types of PTSD treatments [Benish et al., 2008; Bisson and Andrew, 2007; Bradley et al., 2005; Davidson and Parker, 2001; Gerger et al., 2014; Mendes et al., 2008; Powers et al., 2010; Seidler and Wagner, 2006; Watts et al., 2013].

However, several of the above-mentioned meta-analyses reported a considerable amount of between-study heterogeneity [Bisson and Andrew, 2007; Bradley et al., 2005; Davidson and Parker, 2001; Gerger et al., 2014; Powers et al., 2010; Watts et al., 2013], which indicates that in some studies the observed treatment effects may differ from the observed effects of other studies. A concrete example of such contradicting findings from individual studies can be found in the review by Bisson and Andrew [2007], where out of 6 studies that compared cognitive behavioral therapy (CBT) and eye movement desensitization and reprocessing (EMDR) 3 studies reported moderate to large superiority of CBT on clinician-rated PTSD scores, while the remaining 3 studies reported the exact opposite effect, namely moderate to large superiority of EMDR over CBT. Overall, the authors concluded that there was no difference between the effects of the 2 treatments [Bisson and Andrew, 2007].

In order to explain such heterogeneity between individual study estimates, 2 meta-analyses [Gerger et al., 2014; Watts et al., 2013] included different types of PTSD treatments but found no evidence for different types of psychological PTSD treatments to explain between-study heterogeneity. One explanation for the observed differences in the direction of effects between 2 treatments as in the EMDR-CBT comparison is the presence of unbalanced researcher allegiance [Luborsky et al., 1975]. Accordingly, the observed pattern of results in the EMDR-CBT meta-analysis by Bisson and Andrew [2007] could simply be explained by the fact that in one half of the studies researchers preferred CBT and in the other half researchers preferred EMDR. While, by chance, in this particular case the distribution of researcher allegiance appeared to be balanced across the 6 included studies, an unbalanced preference for one particular treatment in a meta-analysis or RCT could be more problematic. In fact, a meta-analysis on trauma-focused PTSD treatments found researcher allegiance to significantly correlate

with effect size differences between treatments ( $r = 0.35$ ) and to explain a substantial amount of between-study heterogeneity [Munder et al., 2012]. Further, the association between researcher allegiance and outcome appears to be indicative of bias rather than of true differences between the effects of individual treatments [Munder et al., 2012; Munder et al., 2011]. Thus, these findings from comparative PTSD RCTs suggest that researcher allegiance – a factor that is *incidental* to the treatment – is of importance in explaining differences between individual study results.

### Dealing with Researcher Allegiance as a Hidden Moderator

Since the impact of an unbalanced researcher allegiance has been observed across a variety of disorders (e.g. depression [Cuijpers et al., 2012], see [Munder et al., 2013] for an overview), this extra-therapeutic moderator has the potential to influence the validity of RCTs and meta-analyses [Staines and Cleland, 2007] and could, if left unattended, result in misleading recommendations for researchers, practitioners, and policy makers, who base their treatment decisions on empirical findings.

It is noteworthy that the moderating role of researcher allegiance is complicated, as multiple causal pathways may explain the established association between unbalanced researcher allegiance and treatment outcome in psychotherapy outcome research [Cuijpers and Cristea, 2015; Leykin and DeRubeis, 2009; McLeod, 2009]. A researcher's allegiance to a particular treatment under investigation may affect methodological decisions which a researcher has to make in the course of a comparative outcome study (e.g. selecting a credible control treatment). Methodological quality has indeed been shown to partly mediate the allegiance outcome association [Munder et al., 2011]: the association between researcher allegiance and outcome was significantly lower in treatment comparisons with high internal validity. But, researcher allegiance might also motivate researchers to operationalize treatments in a way that favors the preferred treatment and/or reduces effects of the non-preferred treatment (e.g. better training of therapists in the preferred treatment) [Spielmans et al., 2010]. However, a recent analysis of depression trials confirms that factors other than study quality (e.g. allocation concealment, selective reporting, training of therapists) may contribute to the association between researcher allegiance and outcome [Chen et al., 2014]. In this analysis, researchers have shown that over the years the amount of studies with a low risk of bias have increased or remained more or less similar for all considered quality criteria with the exception of researcher allegiance. The risk of bias from unbalanced researcher allegiance increased in more recent RCTs. Thus, the mere improvement of study quality does not seem to prevent the risk of bias which occurs due to unbalanced researcher allegiance. Besides introducing bias in clinical trials, researcher allegiance may also interact with patient and therapist variables [Cuijpers and Cristea, 2015]. Theoretically, researcher allegiance could impact patients' outcome expectancies. But a researcher's preference for a particular treatment may also affect the

therapists' behavior, which may result in a more convincing presentation of the treatment rationale or a stronger adherence to the treatment manual. In this sense, researcher allegiance could interact with other moderators or mediators and improve the overall treatment effect. However, the therapists' stronger adherence to a treatment manual may also enhance the effects of *characteristic* treatment constituents if the treatment was delivered with more rigor and competence [McLeod, 2009]. Further, depending on their allegiance, researchers may interpret the available evidence differently. For instance, based on the same study pool, 2 different research teams came to contradicting conclusions regarding the necessity of focusing on the trauma during PTSD treatment [Ehlers et al., 2010; Wampold et al., 2010]. Finally, the preference for a particular treatment may also affect the decision whether or not to publish results in case they are not supporting the effectiveness/superiority of the preferred treatment. Researcher allegiance would thus contribute to the file drawer problem or publication bias [Turner et al., 2008]. For instance, the meta-analysis of Bisson and Andrew [2007] demonstrated for several comparisons an increased risk of missing studies with negative effects.

The variety of possible causal pathways that may explain the association between unbalanced researcher allegiance and outcome poses a substantial problem to the validity of RCTs that aim at estimating the effects of characteristic treatment constituents. In this sense, researcher allegiance should be considered as conflict of interest and thus be fully disclosed [The PLoS Medicine Editors, 2008]; however, to date conflict of interest statements are still rare in psychotherapy RCTs and meta-analyses. For instance, only 1 out of 793 RCTs in which researcher allegiance was rated to be present included a respective conflict of interest statement in a recent review [Dragioti et al., 2015]. Alternatively, the formulation of clear hypotheses (e.g. as in [Foa et al., 1991]) could help to overcome the

issues associated with post-hoc subjective ratings of researcher allegiance, because they allow the reader to directly infer the presence of researcher preferences for one of the compared rival treatments.

## Conclusion

While the often passionate and heated debates about which treatment components are necessary for symptom improvement are and should be part of scientific progress (with promising signs of rapprochement [Hofmann and Barlow, 2014; Schnyder et al., 2015]), the downside may be that the same preferences and beliefs open the door for bias. Researcher allegiance is not a *characteristic* component of any psychotherapeutic treatment approach, nor is it one of the so-called common factors of psychotherapy. Instead, it is an extra-therapeutic factor in the sense that it has no overt association with the investigated treatment nor the context of treatment. The neglect of the impact of such hidden extra-therapeutic factors on treatment effects in psychotherapy research most likely hampers the detection of treatment constituents that critically impact outcome. Based on the correlative nature of the presented meta-analytic research, clearly, further research is strongly needed to confirm the presented conclusions.

## Disclosure Statement

The authors declare that there is no financial conflict of interest concerning this paper. However, and according to our recommendation, the authors declare that H.G. has published affirmatively on the topic of researcher allegiance before and both authors have an interest in the study of contextual and extra-therapeutic factors in psychological interventions.

## References

- Barber JP, Gallop R, Crits-Christoph P, Frank A, Thase ME, Weiss RD, Gibbons M: The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychother Res* 2006;16:229–240.
- Benish SG, Imel ZE, Wampold BE: The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: a meta-analysis of direct comparisons. *Clin Psychol Rev* 2008;28:746–758.
- Bisson JI, Andrew M: Psychological treatment of post-traumatic stress disorder (PTSD). *Cochrane Database Syst Rev* 2007;CD003388.
- Bradley R, Greene J, Russ E, Dutra L, Westen D: A multidimensional meta-analysis of psychotherapy for PTSD. *Am J Psychiatry* 2005;162:214–227.
- Chen P, Furukawa TA, Shinohara K, Honyashiki M, Imai H, Ichikawa K, Caldwell DM, Hunot V, Churchill R: Quantity and quality of psychotherapy trials for depression in the past five decades. *J Affect Disord* 2014;165:190–195.
- Constantino MJ, Arnkoff DB, Glass CR, Ametrano RM, Smith JZ: Expectations. *J Clin Psychol* 2011;67:184–192.
- Cuijpers P, Cristea IA: How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiol Psychiatr Sci* 2015;28:1–8..
- Cuijpers P, Driessen E, Hollon SD, van Oppen P, Barth J, Andersson G.: The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. *Clin Psychol Rev* 2012;32:280–291.
- Davidson PR, Parker KC: Eye movement desensitization and reprocessing (EMDR): a meta-analysis. *J Consult Clin Psychol* 2001;69:305–316.
- Dragioti E, Dimoliatis I, Evangelou E: Disclosure of researcher allegiance in meta-analyses and randomised controlled trials of psychotherapy: a systematic appraisal. *BMJ Open* 2015;5:e007206.
- Driessen E, Cuijpers P, Hollon SD, Dekker JJ: Does pre-treatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *J Consult Clin Psychol* 2010;78:668–680.
- Ehlers A, Bisson JI, Clark DM, Creamer M, Pilling S, Richards D, Schnurr PP, Turner S, Yule W: Do all psychological treatments really work the same in posttraumatic stress disorder? *Clin Psychol Rev* 2010;30:269–276.
- Elliott R, Bohart AC, Watson JC, Greenberg LS: Empathy. *Psychotherapy* 2011;48:43–49.
- Foa EB, Hembree EA, Cahill SP, Rauch SA, Riggs DS, Feeny NC, Yadin E: Randomized trial of prolonged exposure for posttraumatic stress disorder with and without cognitive restructuring: outcome at academic and community clinics. *J Consult Clin Psychol* 2005;73:953–964.
- Foa EB, Rothbaum BO, Riggs DS, Murdock TB: Treatment of posttraumatic stress disorder in rape victims: a comparison between cognitive-behavioral procedures and counseling. *J Consult Clin Psychol* 1991;59:715–723.
- Gaab J, Bleas C, Locher C, Gerger H: Go open: a plea for transparency in psychotherapy. *Psychol Conscious (Wash D C)* 2015;DOI dx.doi.org.1037/cns0000063.
- Gerger H, Munder T, Gemperli A, Nüesch E, Trelle S, Jüni P, Barth J: Integrating fragmented evidence by network meta-analysis: relative effectiveness of psychological interventions for adults with post-traumatic stress disorder. *Psychol Med* 2014;44:3151–3164.
- Grünbaum A: The placebo concept. *Behav Res Ther* 1981;19:157–167.
- Grünbaum A: The placebo concept in medicine and psychiatry. *Psychol Med* 1986;16:19–38.



- Hofmann SG, Barlow DH: Evidence-based psychological interventions and the common factors approach: the beginnings of a rapprochement? *Psychotherapy* 2014; 51:510–513.
- Ioannidis JPA: Why most published research findings are false. *PLoS Med* 2005;2:e124.
- Jones DS, Podolsky SH: The history and fate of the gold standard. *Lancet* 2015;385:1502–1503.
- Kirsch I: Placebo psychotherapy: synonym or oxymoron? *J Clin Psychol* 2005;61:791–803.
- Leykin Y, DeRubeis RJ: Allegiance in psychotherapy outcome research: separating association from bias. *Clin Psychol Sci Pract* 2009;16:54–65.
- Lipsey MW, Wilson DB: *Practical Meta-Analysis*. Thousand Oaks, Sage, 2001.
- Luborsky L, Singer B, Luborsky L: Comparative studies of psychotherapies. Is it true that 'everyone has won and must have prizes'? *Arch Gen Psychiatry* 1975;32:995–1008.
- Lund K, Vase L, Petersen GL, Jensen TS, Finnerup NB: Randomised controlled trials may underestimate drug effects: balanced placebo trial design. *PLoS One* 2014; 9:e84104.
- Lynch D, Laws KR, McKenna PJ: Cognitive behavioural therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials. *Psychol Med* 2010;40:9–24.
- McLeod BD: Understanding why therapy allegiance is linked to clinical outcomes. *Clin Psychol Sci Pract* 2009; 16:69–72.
- Mendes DD, Mello MF, Ventura P, de Medeiros Passarela C, de Jesus Mari J: A systematic review on the effectiveness of cognitive behavioral therapy for posttraumatic stress disorder. *Int J Psychiatry Med* 2008;38: 241–259.
- Munder T, Brüttsch O, Leonhart R, Gerger H, Barth J: Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clin Psychol Rev* 2013; 33:501–511.
- Munder T, Flückiger C, Gerger H, Wampold BE, Barth J: Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *J Couns Psychol* 2012;59:631–637.
- Munder T, Gerger H, Trelle S, Barth J: Testing the allegiance bias hypothesis: a meta-analysis. *Psychother Res* 2011;21:670–684.
- Norcross JC, Lambert MJ: Psychotherapy relationships that work II. *Psychotherapy* 2011;48:4–8.
- Powers MB, Halpern JM, Ferenschak MP, Gillihan SJ, Foa EB: A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clin Psychol Rev* 2010; 30:635–641.
- Schnyder U, Ehlers A, Elbert T, Foa EB, Gersons BPR, Resick PA, Shapiro F, Cloitre M: Psychotherapies for PTSD: what do they have in common? *Eur J Psychotraumatol* 2015;6:28186.
- Seidler GH, Wagner FE: Comparing the efficacy of EMDR and trauma-focused cognitive-behavioral therapy in the treatment of PTSD: a meta-analytic study. *Psychol Med* 2006;36:1515–1522.
- Spielmans GI, Gatlin ET, McFall JP: The efficacy of evidence-based psychotherapies versus usual care for youths: controlling confounds in a meta-reanalysis. *Psychother Res* 2010;20:234–246.
- Staines GL, Cleland CM: Bias in meta-analytic estimates of the absolute efficacy of psychotherapy. *Rev Gen Psychol* 2007;11:329–347.
- The PLoS Medicine Editors: Making sense of non-financial competing interests. *PLoS Med* 2008;5:1299–1301.
- Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R: Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252–260.
- Waltz J, Addis ME, Koerner K, Jacobson NS: Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *J Consult Clin Psychol* 1993;61:620–630.
- Wampold BE, Imel ZE, Laska KM, Benish S, Miller SD, Flückiger C, Del Re AC, Baardseth TP, Budge S: Determining what works in the treatment of PTSD. *Clin Psychol Rev* 2010;30:923–933.
- Watts BV, Schnurr PP, Mayo L, Young-Xu Y, Weeks WB, Friedman MJ: Meta-analysis of the efficacy of treatments for posttraumatic stress disorder. *J Clin Psychiatry* 2013;74:e541–550.
- Weinberger J: Common factors are not so common and specific factors are not so specified: toward an inclusive integration of psychotherapy research. *Psychotherapy* 2014;51:514–518.