Citation for published version (APA):
Establishing how psychological therapies work: The Importance of Mediation Analysis.

Authors

Windgassen S*, Goldsmith K, Moss-Morris R, Chalder T.

* address for correspondence

1. Department of Psychological Medicine, King’s College London.

2. Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King’s College London.

3.Division of Health Psychology, King’s College London.

Corresponding address and contact details:

Sula Windgassen
Chronic Fatigue Unit 1st Floor,
Mapother House,
De Crespigny Park,
Denmark Hill,
SE5 8AZ
Email: sula.1.windgassen@kcl.ac.uk

Tel: 0203 228 3557

Declarations of Interest

No conflicts of interest are declared.

Acknowledgements
TC and KG were in part supported by the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, Kings College London.

**Abstract**

This editorial reviews the importance of psychological studies that are designed to address the question of how psychological interventions create change. The practical use and implications of assessing mechanisms of treatments are considered with examples from existing psychological research. The potential for elucidating theoretical mechanisms, developing new theoretical models and modifying treatment approaches are described. In addition an overview of different types of statistical methods available to researchers for assessing mediation is given, with a recommendation of Structural Equation Modelling (SEM). The review concludes with a summary of optimum study conditions to be adopted by researchers to establish mediating effects.

**Keywords**

Mediation; Psychological Interventions; Statistical Analysis.

**Main Text**

**What is Mediation?**

Psychological studies generally focus on measuring whether an intervention works or not. More specifically does the therapy have an effect on a certain outcome and to what extent? It is less often that researchers investigate *how* interventions exert their effects on an outcome. The investigation into 'how' is an investigation into mediation.
When the effect of one variable, often an intervention, has its effect on an outcome through change in a third variable, mediation is said to occur. This third variable is called the mediator.

Investigating mediation is important both for the advancement of psychological theory and refinement of interventions. The study of psychological mediators can allow us to capitalise upon key processes involved in generating positive outcomes. This editorial aims to facilitate a basic understanding of mediation, provide some sense of its scope and illustrate that despite such analyses potentially requiring some specialist knowledge, it is critical to advancing understanding of psychological therapies. Examples of research into mediational processes from psychosocial intervention studies are provided to demonstrate the importance and value of the study of mediation.

**Study design considerations**

Randomised Controlled Trials (RCTs) are considered to be the gold standard of assessing therapeutic change (Evans, 2003) and they are also the optimum form of study design for establishing mediation. Ideally the incorporation of mediation should be an integral part of the design phase of an RCT. This would allow careful consideration of variables to be measured before during and after the intervention. The variables to be included in mediation analysis should be informed by theory and/or empirical studies, to avoid “fishing” which may cloud theoretical understanding (Johansson & Høglend, 2007).

*How is Mediation Established?*
The literature concerning mediation in psychology is growing, as evidenced by the increasing number year on year of citations of Baron & Kenny’s seminal paper published in 1986 (Baron & Kenny, 1986; MacKinnon et al, 2007; MacKinnon et al, 2002; Zhao et al, 2010). The Baron and Kenny article presents mediation in a three variable path model. The three variables are:

- **A** = the intervention variable
- **B** = the mediator variable
- **C** = the outcome variable (an appropriate measure of therapeutic change)

The model includes two paths leading to the outcome variable: One is the direct effect of the intervention on the outcome, whilst the other is the indirect effect of the intervention on the outcome via the impact of a third variable. This third variable is the mediator; the mechanism through which the intervention variable influences the outcome variable.

Baron and Kenny’s methodology asserts that a series of regressions be conducted to establish the statistical significance of relationships between the variables in the different paths.

The series of regressions seeks to infer the mechanistic process of B (mediator) by ascertaining whether: (1) in a regression of the outcome variable C on the intervention variable A, the effect of the intervention variable is statistically significant (2) in a regression of mediator (B) on intervention (A), the effect of the intervention variable is statistically significant (3) in a regression of outcome (C) on intervention (A) and mediator (B), the effect of the mediator is statistically significant (4) in the regression in (3), the intervention effect is no longer statistically significant when controlling for the mediator. Baron and Kenny say
the strongest evidence for mediation is when the effect of the intervention in regression (3) is reduced to zero, which is generally referred to as “full mediation”. If the effect of the mediator and intervention are significant in (3), or if the effect of the mediator is significant and the effect of the intervention is not zero, but is lessened when controlling for the mediator, this is generally referred to as partial mediation, and the assumption is that there are other mediators influencing the effect of the intervention on the outcome. This method is sometimes referred to as the causal steps approach to mediation (MacKinnon et al, 2007; MacKinnon et al, 2002).

Although this method of testing mediation is now widespread in the literature, there are limitations associated with the method that have been widely discussed elsewhere (Emsley et al, 2010; Kazdin, 2007; MacKinnon et al, 2002; Zhao et al., 2010). One issue with the method is the requirement for a significant intervention effect as stated in (1). Along with others (Emsley et al, 2010; MacKinnon, 2008), we do not agree that mediation should only be investigated when there is a significant intervention effect. It may be even more important to study mediation in this case; in order to determine if the intervention does not have the desired effect on the mediator, the mediator does not have an effect on the outcome, or if there is evidence of suppression (MacKinnon et al, 2000; MacKinnon, 2008). Suppression occurs when the indirect and direct effects oppose one another. Another difficulty with this method is that it does not directly quantify the indirect effect through the mediator. Instead it relies on a number of hypothesis tests to make inferences about the indirect effect. In addition, this method has been shown to have low power to detect mediated effects. In other words, researchers using this method may miss effects even
when they are present (Hayes, 2009; MacKinnon et al, 2007; MacKinnon et al, 2002; Preacher & Hayes, 2008). Finally, Baron & Kenny did not address the possibility of biased results due to unmeasured confounding variables i.e. variables which may influence both the mediator and the outcome. The omission of such confounding variables could bias the results of mediation analysis (Emsley et al, 2010; MacKinnon, 2008). This issue was presented in the earlier and less referenced paper by Judd & Kenny (Judd & Kenny, 1981), and can be at least partially dealt with by measuring potential confounders and including them in the regression models.

Another issue that was brought to the fore in both the Baron & Kenny and Judd & Kenny papers (Baron & Kenny, 1986; Judd & Kenny, 1981) was measurement error in variables, which could also lead to biased effect estimates. Measurement error is likely to be of particular concern in psychology and psychiatry, where we are often interested in unobservable or latent traits. Structural Equation Modelling (SEM) allows modelling of relationships between underlying latent trait variables, each quantified by several scale items or measures. For example maladaptive cognitions may be measured using a questionnaire, which includes several items that could be indicative of such cognitions (“I am no good”, “they probably think I can’t do this”). In this example the latent trait is maladaptive cognitions. By utilising multiple items or measures, SEM account for measurement error and elucidate relationships between latent traits (Bollen & Pearl, 2013; MacKinnon, 2008). SEM utilising only observed variables are sometimes referred to as Path Analysis Models (MacKinnon, 2008).
Like the use of Baron & Kenny’s Framework, the use of SEM for investigating mediation has also been criticised. However, it has been noted that rather than an issue with the method itself (Bollen & Pearl, 2013; Emsley et al, 2010; MacKinnon, 2008), this is more to do with improper or non-specification of theoretical models and disregarding assumptions in interpreting results (Bollen & Pearl, 2013; Emsley et al, 2010; MacKinnon, 2008). Another advantage to the SEM approach to mediation is that SEM can simultaneously model multiple outcomes/regressions, such as those described in the Baron & Kenny framework. Importantly, this also allows for longitudinal modelling of multiple measures of mediators and outcomes (MacKinnon, 2008). This may be necessary where a theoretical model implicates multiple mediators, which may act independently or in conjunction with one another. SEM allows modelling of the relationships between intervention, multiple mediators and outcome(s). To summarise, two main benefits of SEM are: the ability to allow for measurement error (Little et al, 2007), and the ability to investigate more complex models of mediation.

**Mediation in Psychological Research**

As already discussed, the benefit of mediation research in psychological studies lies within the potential for therapeutic approaches to be enhanced. By identifying mediating mechanisms, therapeutic processes may be refined to focus on specific aspects of therapy that lead to improvements in outcomes, with the possibility of discarding aspects that are less relevant (Kazdin, 2007). This could lead to more efficient delivery of therapy.
Investigations of the mechanistic processes of psychotherapies are increasing, although currently the literature remains limited. Psychotherapies that have been more widely subjected to mediation analysis include cognitive behavioural therapy (CBT) (Maric et al, 2013; Odondi et al, 2013; Turner et al, 2007; Whisman, 1993) and mindfulness based therapies (MBTs) (Bränström et al., 2010; Coffey & Hartman, 2008; Sears & Kraus, 2009). CBT is designed to alter negative patterns of thinking and behaving that are considered to cause and/or maintain symptoms and disability. This is applicable to CBT across varying disorders. The proposed mechanisms of change at a general level are cognitions and behaviours. Investigation into whether these processes are indeed responsible for improved outcomes as a result of CBT have been conducted in the context of chronic pain (Turner et al., 2007), panic disorder (Hofmann et al., 2007), chronic fatigue (Moss-Morris et al., 2005, Chalder et al., 2015) and irritable bowel syndrome (Lackner et al., 2007; Miklowitz & Scott, 2009; Reme et al., 2011) amongst others.

**Investigating Proposed Theoretical Mechanisms**

Research into mediators in CBT may be conducted utilising a more general model of CBT i.e. identifying change in outcome to be a result of changes in cognitions and/or behaviours generally. This was the case in a study conducted by Turner et al., (2007). Here they conducted mediation analysis using data from an RCT that reported the effectiveness of CBT compared with education or attentional control conditions in improving pain severity and disability in the context of Temporomandibular Disorder (TMD). The CBT intervention was designed to target participants’ negative beliefs regarding pain, tendencies to catastrophize and to increase participants’ adaptive coping strategies and
control beliefs. To assess the effects of individual mediator variables, they applied the Baron & Kenny framework. In order to assess the amount of variance in the outcome accounted for by all mediators as a group, they also used structural equation modelling. They found that when all mediators were included in the model, changes in perceived disability and self-efficacy were found to mediate change in activity interference at one year.

Alternatively, research into mediators in CBT may be informed by CBT models developed for specific conditions. Here, specific cognitions and behaviours targeted by the specified model may be identified as potential mediators. In the context of irritable bowel syndrome for example, the CBT model postulates that reduction in symptom severity and impact on life is due to treatment-induced changes in conceptualisation of bowel symptoms (Toner et al., 2000). This reconceptualization should involve changes in beliefs about IBS being an uncontrollable medical problem and increases in behavioural strategies that can be employed in the face of symptoms. One study examining cognitive and behavioural mechanisms in IBS found evidence supportive of this (Reme et al., 2011). The mediator variables tested included the Cognitive Scale for Functional Bowel Disorders (CS-FBD) (Toner et al., 1998) study and the Behaviour Scale for IBS (IBS-BRQ) (Reme et al., 2010). The CS-FBD includes items relating to specific beliefs about functional bowel disorders such as “It is embarrassing to keep going to the toilet”. The IBS-BRQ includes specific items for behaviour implicated in IBS such as “I avoid exercise when I have stomach pains”.
The analysis found that change in both cognitions and behaviours mediated the reduction in symptom severity and impact on life. SEM was used to apply a sequential mediator model indicating that behaviours changed prior to cognitions (Reme et al., 2011).

This example demonstrates that one benefit of conducting mediation analysis is that it allows for the evaluation of theoretically implicated process variables. Establishing whether these processes are important or not allows the therapeutic focus to be modified accordingly. Where the focus of intervention differs within the same school of psychotherapy, mediation can be used to clarify the extent to which different processes produce change in outcomes. For example, within CBT as applied to IBS, different researchers postulate the importance of different mechanistic processes, namely a change in cognitions and/or behaviour versus a reduction in distress. To date findings supporting one key process over the other have been inconsistent. This is demonstrated by the differing results in the mediation studies investigating CBT for IBS, conducted by Lackner et al., (2007) and Jones et al., (2011). Both studies investigated the potential mediating roles of psychological distress on outcome after CBT for IBS, with conflicting findings. Jones et al found that anxiety and to a lesser extent depression had a mediating effect, whereby decreases in both led to a reduction in symptom severity. This contrasted with Lackner’s analysis of a complex mediation model, including psychological distress as both a mediator and an outcome. It was found that psychological distress was not a significant mediator, and that instead CBT had a direct effect on symptom severity independent of distress. Chilcot and Moss-Morris (2013) also later found that distress was not a
significant mediator, whereas cognitions were, strengthening support of a CBT model for IBS in which change is mediated by cognitions rather than distress.

Further empirical studies using SEM to assess the potential mediating roles of these variables could help clarify conclusions as to where focus in therapy should lie. In this context, further mediation studies could provide insight into whether cognitive reframing should be conducted with a view to reducing distress primarily or to address other specific components of the condition that may be more effective in improving outcomes. The example of the Jones and Lackner studies illustrates a number of limitations within the current mediation literature: (a) Results are obviously dependent on which mediation variables are entered in the analysis. Neither of these studies included cognitive or behavioural measures and therefore the effects of these as potential mediators could not be assessed. Consequently it is important for mediation studies to be fully informed by theory to allow for examination of all possible mediators as dictated by the theoretical model. (b) Different measures may be utilised to measure concepts that are the same or similar, which can limit interpretation across studies e.g. a measure of anxiety vs. a measure of psychological distress. (c) Different approaches to mediation analysis may affect findings. This will be considered later in the article and (d) It may be that there is a longer mediation chain that involves more than one mediator variable in the causal path and therefore one mediator may serve to mediate the effects of another mediator (Taylor et al., 2008).
**Developing theory**

The previous section described the use of mediation studies to examine processes guided by psychological theory. Here we consider how mediation analysis can provide opportunities to build theory, which we refer to as "back translation". Mindfulness research provides an opportunity in which to consider the application of back translation. The provision of MBT is growing across different clinical populations, yet the theoretical underpinnings of such practice remains scarce in the context of psychotherapeutic interventions. In 2006, Shapiro et al published a theory by which mindfulness was proposed to exert clinical effects (Shapiro et al., 2006). The theory asserted that four variables identified as changes in self-regulation, values, flexibility (cognitive, emotional and behavioural) and exposure to internal processes, may act as mechanisms responsible for outcomes such as reduction in symptoms or distress. This preliminary theory was tested through the use of mediation analysis and partial support was found for the model (Carmody et al., 2009). Mindfulness was found to have a direct effect on psychological distress, whilst changes in values and flexibility were found to partially mediate the effect of mindfulness. Although no update to the theory proposed by Shapiro has been published, research has continued to investigate cognitive and behavioural flexibility as potential mediators of mindfulness (Baer, 2010; Heeren et al., 2009) and other elements nested within the theoretical model and wider literature, such as compassion have invited further mediation research (Baer, 2010; Hölzel et al., 2011; Kuyken et al., 2010).
Another reason for investigating mediation in the context of psychotherapy is to elucidate mechanisms that are potentially shared across different psychotherapeutic practices. Discovery of such mechanisms could lead to a greater insight into particular elements of therapy that are efficacious for a wide range of outcomes. In a systematic review of psychosocial treatments for bipolar disorder, common mechanisms across treatments were identified. These included enhancement of interpersonal functioning and teaching self-monitoring to allow early self-intervention during relapses (Miklowitz & Scott, 2009). Such processes have also been implicated in interpersonal therapy (IPT) (Lipsitz & Markowitz, 2013) and can be seen to be present in other therapies such as CBT (Livesley, 2007; Steever, 1999), MBT (Epstein et al., 2008) and counselling (Howey & Ormrod, 2002). What was particularly noteworthy in the systematic review was that effective therapies shared a number of common characteristics with regards to how the model of therapy was shared with the patient, how the therapy was delivered and the structuring of treatment (Miklowitz & Scott, 2009). Shared characteristics included individualized formulation, openly sharing the therapy model with the patient, a clear rationale for techniques used that were logical to the patient, an emphasis on psychoeducation and skill development, attributing change to the patient’s efforts and the encouragement of the continued use of illness management techniques for the patient post therapy.

Other disease specific systematic reviews conducted have further demonstrated the existence of shared mechanisms across different therapeutic approaches. In a recent review of mechanisms in psychosocial interventions for cancer, self-
efficacy to use coping strategies and changes in cognitions mediated treatment outcomes in CBT, psycho-education and relaxation training (Stanton et al., 2013). However, it should be noted that the outcomes across studies varied. The outcomes were broadly classified into the following domains: Psychosocial adjustment, self-reported physical health indicators and biological health indicators. Future reviews establishing mediation across the same or similar outcomes would provide an opportunity to test the hypothesis that the effects of mediators on such outcomes are the same across therapies.

**Scope of mediation analysis**

Mediation can be a complex process to conceptualise theoretically as well as to approach statistically. Some particular areas for researchers to focus on are discussed below.

*Different approaches to mediation*

The "product of coefficients" approach (MacKinnon, 2001; MacKinnon & Dwyer, 1993), is an extension of the Baron & Kenny causal steps approach, which draws from, and can be applied using, the SEM framework. This approach is preferred (Hayes, 2009; MacKinnon et al., 2007) over causal steps. Product of coefficients calculates the indirect (mediated) effect by multiplying the intervention regression coefficient in regression (2), by the mediation regression coefficient in regression (3).

*Temporal Precedence*

Simply studying mediation in the context of an RCT may not be sufficient to establish mediation in the absence of the collection of measures at multiple and appropriate time points in order to explore the timeline of mediator and
outcome change (Cole & Maxwell, 2003). To demonstrate temporal precedence of mediation it is necessary to assess both mediator and outcome measures before, during and after treatment, with assessment at additional time points being ideal. This can allow the capture of change in the mediator prior to change in the outcome as well as potentially allowing for the investigation of bidirectional changes whereby two variables (or more) may influence each other at differing time points. For example in CBT initially behaviour may change cognitions (during an intervention), but once cognitions have changed, they may further change behaviour and so on.

**Summary**

The full complexity of investigating mediation is beyond the scope of this article, however the issues raised suggest reasons why findings across mediation studies within the same area may be inconsistent (different approaches, measurement error, etc.). Furthermore considering the intuitive and comparatively simplistic nature of the causal steps approach, it also makes sense that this has been the predominant method of establishing mediation. Nevertheless it would be useful if researchers undertaking mediation analysis could adopt more sophisticated methods where appropriate, to increase the validity of findings (Hayes, 2009; MacKinnon et al., 2007; MacKinnon et al., 2002; Zhao et al., 2010). The widely prevalent use of the causal steps approach to mediation has resulted in a number of misconceptions, including the idea that a statistically significant effect of an intervention is necessary before it is advisable to test for mediation; and that a decrease in variance accounted for in the A-C relationship after inclusion of a mediator is sufficient to conclude there is
mediation (Stanton et al., 2013). Happily, researchers interested in mediation analysis can now refer to a widening literature on best practices in the modern study of mediation (Hayes, 2009; MacKinnon, 2008).

In the context of psychological studies, the study of mediation is critical if we are to understand how therapies exert effect, test psychological models of therapeutic mechanisms, and most importantly, improve outcomes for our patients. More studies of mediation that are conducted in a more rigorous fashion will provide insight into therapeutic mechanisms, which may be transdiagnostic (Murphy et al., 2009). Future randomised studies of psychological therapies should therefore include mediation analysis in their design wherever possible, with the inclusion of potential mediator measures informed by theory. Measurements should be taken early and at multiple time points, with concomitant measurement of potential confounders in order to allow for detailed and robust assessments of mediational processes.
References


**Figures**

![Simple mediation model](image)

**Figure 1:** Simple mediation model