ANOTHER ERROR OF DESCARTES?
IMPLICATIONS FOR THE “THIRD WAVE”
COGNITIVE-BEHAVIORAL THERAPY

Daniel DAVID* & Stefan G. HOFMANN

1Babeș-Bolyai University, Cluj-Napoca, Romania
2Boston University, Boston, USA

Abstract
In this article we explore the fundamentals of the “third wave” cognitive-behavioral therapies (i.e., Acceptance and Commitment Therapy; ACT), based on the debate between Descartes and Spinoza about our mental representation and understanding. We argue that although ACT techniques might work in various clinical conditions, their underlying theory is debatable and therefore, the efficacy/effectiveness of ACT techniques could be better conceptualized based on classic cognitive-behavioral theory.

Keywords: Descartes vs. Spinoza debate, cognitive-behavioral therapy, cognitive therapy, rational emotive behavioral therapy, acceptance and commitment therapy

Cognitive-behavioral therapy (CBT) is the gold standard for the evidence-based movement in psychotherapy (see David, Lynn, & Ellis, 2010; Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012). More recently, Acceptance and Commitment Therapy (ACT) has been enthusiastically championed by some of its creators as a third wave behavioral treatment in the CBT paradigm (e.g., Hayes, 2004; but see Hofmann & Asmundson, 2008 for a contrary view). More specifically, it has been proposed that in the CBT paradigm the first wave is represented by behavioral therapies (e.g., behavior modification, exposure-based therapies, systematic desensitization), focused on classical conditioning, operant conditioning, and other behavioral learning principles. The second wave is represented by cognitive therapies (e.g., Rational Emotive Behavioral Therapy/REBT; Cognitive Therapy/CT) that are focused on cognitive appraisal and restructuring of dysfunctional/irrational cognitions (i.e., classic CBT). The third wave is primarily represented by Acceptance and Commitment Therapy (ACT). Some authors (Hayes, Luoma, Bond, Masuda, & Lillis, 2006) also count Meta-cognitive approaches (MCT; Wells, 2000), Dialectic Behavior Therapy (DBT; Linehan, 1993), and Mindfulness-Based Cognitive Therapy (MBCT; Segal, Williams, & Teasdale, 2001) to be part of this third wave movement.

*Correspondence concerning this article should be addressed to:
E-mail: danieldavid@ubbcloj.ro
However, none of the authors of these treatments consider themselves to be part of this *third wave* movement. For example, like other classic CBT approaches, MCT is based on a model that has a cognitive architecture representing the interplay between levels of cognition and types of cognition in the control of conscious experience. Similarly, Marsha Linehan (personal communication, August 28, 2007) does not consider DBT to be part of the *third wave* but, instead, views DBT as a form of classic CBT that includes acceptance strategies. Finally, MBCT includes traditional CBT techniques and adds mindfulness-based techniques.

Accordingly, ACT remains the sole representative of this so-called *third wave* movement. This approach views cognitive restructuring (i.e., cognitive reframing) as non-essential in producing therapeutic change. The objective of therapy is not to change the actual content of dysfunctional/irrational cognitions, but to alter their function in determining psychopathology (e.g., changing the individual’s relationship to cognitions).

**The problem**

At this moment, the prominence of CBT in the clinical field and in the evidence-based movement in psychotherapy is mainly grounded in the first two waves. However, recent studies also support the *third wave* approaches as evidence-based interventions for various clinical conditions. Therefore, the *third wave* might be able to further strengthen the position of CBT as the *gold standard* in the evidence-based movement in psychotherapy (e.g., see the APA Division 12’s List of Research-Supported Psychological Treatments; http://www.div12.org/PsychologicalTreatments/index.html).

David and Montgomery (2011) recently proposed a new classification system for evidence-based psychotherapies, which simultaneously takes into account both the efficacy/effectiveness of the intervention package (i.e., the clinical protocol/psychological treatment) and the scientific support of the underlying theory of the specific clinical protocol (i.e., theory/mechanisms of change). Indeed, mechanisms of change might clarify the distinction between science and pseudoscience in psychotherapy (see Table 1 for details).

We think that through the lenses of basic research in cognitive sciences, this second component – *the support for the underlying theory of a specific clinical protocol* – could prove to be an Achilles’ heel (i.e., weakness) for some of the *third wave*’s approaches. Therefore, a critical analysis and understanding of the core of *third wave* approaches and their underlying theories is fundamental to ground *third wave* treatments on firm psychological mainstream, and thus, enhance the central role of CBT in the evidence-based movement. Next, we will briefly analyze this problem as related to ACT.
Table 1. New Psychotherapies Classification Framework: Categories I-IX (first published in David & Montgomery, 2011)

<table>
<thead>
<tr>
<th>Theory of Change/ Clinical Protocol</th>
<th>Well-Supported</th>
<th>Equivocal – No, Preliminary, or Mixed Data</th>
<th>Strong-Contradictory Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-Supported</td>
<td>Category I</td>
<td>Category II</td>
<td>Category V</td>
</tr>
<tr>
<td>Equivocal – No, Preliminary or Mixed Data</td>
<td>Category III</td>
<td>Category IV</td>
<td>Category VII</td>
</tr>
<tr>
<td>Strong-Contradictory Evidence</td>
<td>Category VI</td>
<td>Category VIII</td>
<td>Category IX</td>
</tr>
</tbody>
</table>

Acceptance and Commitment Therapy (ACT)

We focus our analysis on ACT (see Hayes et al., 2006), because it is seriously opposed to the cognitive model and its main techniques (i.e., cognitive restructuring). Indeed, ACT strongly argues that the cognitive restructuring of a dysfunctional/irrational cognition is not needed in order to get clinical benefits. Instead, one should accept the dysfunctional/irrational cognition by various techniques (e.g., acceptance, cognitive defusion, mindfulness). Thus, the core ACT techniques aim to correct people’s tendencies to become fused with their thoughts (e.g., to see them as literally true or false). The acceptance process will then lead to a reduction in the believability of cognitions (e.g., “a thought is just a thought” is a prototypical saying of the ACT practitioners), often not changing their frequency, form, and/or content, but their function. In the end, a dysfunctional/irrational cognition becomes just a “neutral” cognition (i.e., noticing thoughts without agreement or disagreement).

In classic CBT, patients are typically encouraged to detect and monitor their thoughts and to see them as hypotheses rather than objective facts about the world (Hollon & Beck, 1979). This is the first step in the cognitive restructuring process (i.e., cognitive distancing). However, the idea of cognitive distancing is to see the thoughts as hypotheses to be (re)tested, not to see them as simple thoughts without agreement and disagreement.

This fundamental view of the ACT theory (e.g., be it theory/mechanisms of change associated with various specific clinical protocols and/or part of Relational Frame Theory - as a more general theory underlying ACT practice) seems to be based on an outdated Cartesian model of human cognition (i.e., mental representation and understanding). Our intention is not to discuss the core processes of ACT. We are aware that this approach is grounded on functional contextualism (e.g., Gifford & Hayes, 1999; Pepper, 1942), which is an attempt to
offer a way to integrate cognition and language into a behavioral analytic framework (Hayes et al., 2006). Instead, our goal is to put this model in the context of the Cartesian model of human cognition (i.e., mental representation and understanding).

**Descartes’ potential error**

Before discussing the relevance of the Cartesian model, we will briefly introduce some key constructs in cognitive sciences, which will then be used in our analysis and in the interpretation of its results.

In current terms, cognitions refer to information processing events, which can be conscious or unconscious (i.e., implicit) in nature. If cognitions are conscious, we typically call them thoughts; however, some authors may also consider unconscious cognitions thoughts. Beliefs, whether explicit or implicit, refer to an individual's judgment as to whether a thought is true or not (obviously, beliefs are (meta)cognitions themselves). Knowledge is often defined as “justified true beliefs” (however, see the criticism of this definition based on the Gettier problem and the defense of it by Robert Nozick).

According to Descartes’ model of human mental representation and understanding (see Figure 1), we first understand a message (e.g., coding it as a cognition/thought) and then we decide (e.g., consciously and/or by unconscious information processing) if the cognition/thought is true or false. Thus, we believe a cognition/thought for which we have evidence; this way a cognition/thought becomes a belief. Otherwise, a cognition/thought would be neither believed nor would it be not believed, but in a state of simple thought, without agreement and disagreement. This model would support ACT efforts toward having thoughts without agreement or disagreement.

However, Descartes’ model has a powerful rival. Indeed, according to Spinoza’s model (see Figure 1), understanding also means the automatic attribution of a value of true to an understood thought. Then, if we have enough cognitive resources, we will decide (e.g., consciously and/or by unconscious information processing) if the thought is really true and/or false. Thus, according to Spinoza’s model, the human mind uses a general rule of thumb that anything that is understood is also true (e.g., it might be an evolutionary development favoring “false positive” against “false negative” biases, taking into account that most of the information we are exposed to is indeed true and may be important for our survival) (see Gilbert, Krull, & Malone, 1990). This model would not support ACT theory and efforts toward having thoughts without agreement or disagreement. Both models are summarized in Figure 1.

In a series of studies, Gilbert et al. (Gilbert et al., 1990; Gilbert, 1992; Gilbert, Tafarodi, & Malone, 1993) analyzed and tested these two fundamental assumptions regarding the believability of thoughts: Descartes’ model versus
Spinoza’s model. What Gilbert et al. found is that the human mind seems to work most of the time according to Spinoza’s model!

However, several authors tried to limit the conditions of validity of Spinoza’s model. Hasson, Simmons, and Todorov (2005) argued that the model is correct for uninformative cognitions (i.e., propositions). Informative cognitions refer to propositions that stimulate new meaningful and relevant inferences (e.g., because they are often atypical for the person’s mind) (see Hasson et al., 2005). Some cognitions (e.g., “He is a communist”) are informative both when true and when false. Other cognitions (e.g., “He is a worm eater”) are informative when true, but not when false, while other cognitions (e.g., “He is 50 years old and married”) are informative when false, but not when true. Finally, some cognitions (e.g., “He likes drinking milk”) are uninformative both when true and when false. Hanson et al. (2005) argued that people might have the ability to contemplate cognitions without believing them – according to the Descartes’ model -, but only when these cognitions are informative when they are false (see also Mayo, Schul, & Burnstein, 2004).

![Figure 1. Descartes versus Spinoza’s models of mental representation and understanding (adapted after Gilbert et al., 1990; 1993)](image)

Richter, Schroeder, and Wohrmann, (2009) argued the Spinoza’s model works for general beliefs for which people have weak background beliefs; if the background is strong, individuals will most probably tend to automatically and rapidly evaluate the cognitions as true or false. This logic is consistent with
Bayes’ Theorem (Bayes, 1764), which states that rational belief is guided by the laws of conditional probabilities and base rates.

Nadarevic and Erdfelder (2013) even argued that Spinoza’s model should not be the reference model because the previous support was based on guessing bias and thus the model is valid only when false information is rare.

Mandelbaum (2013) has recently addressed these criticisms and has proposed a re-conceptualization of data apparently limiting the application of Spinoza’s model in accordance to a general application of the model.

Summarizing, Descartes’ model is seriously challenged by Spinoza’s model, not only in terms of philosophical debates, but also in terms of psychological data and their implications. It is true that the verdict is not yet definitive, but there is already strong doubt that Descartes’ model has a general and strong validity in terms of how the human mind works.

Therefore, ACT’s idea that we can change directly, by various techniques (e.g., acceptance, cognitive defusion, and/or mindfulness), a dysfunctional/irrational cognition into a “neutral” thought, neither believed nor unbelieved, might be fundamentally wrong and a weakness of the model! Moreover ACT research has often found that by using ACT techniques for distressing thoughts, one changes both the frequency and the believability of the thoughts (see Zettle & Hayes, 1986) thus, again, challenging the ACT idea that patients would maintain their previous distressing thoughts as “neutral” thoughts, neither believed nor unbelieved.

Even if we think in a Popperian framework (Popper, 1959) and try to falsify the Spinozian conclusion (i.e., that there is no contemplation of propositions without believing them) by looking for counter-examples [i.e., by accepting, for the sake of analysis, the constraints of Hasson et al. (2005), Nadarevic & Erdfelder (2013), Richter et al. (2009)], one can note that distorted cognitions (e.g., “I will fail”) would still fit the limited conditions of validity of Spinoza’s model. Indeed, (1) distorted cognitions (“I will fail”) seem to be uninformative when false (e.g., similar to the uninformative value of the cognition “He is a worm eater” when it is false) and (2) the base rate of true and false cognitions favors true cognitions, as distorted cognitions are typically strongly believed by our patients. Moreover, and maybe the strongest argument, because we have a strong background for distorted cognitions (e.g., schema), patients will automatically and rapidly evaluate them as true or false, rather than maintaining them in a state of contemplation without believing them. Indeed, this argument is important not only in challenging the original sequence described by Descartes (i.e., moving from mere understanding a message to tagging the message as true or false; see Figure 1), but also, and maybe more important, in challenging the reversed sequence of Descartes’ model: moving from a message already tagged as true or false – as it is typically the case in the psychotherapy process – to a mere understanding of the message.
Implication of Descartes’ potential error for “third-wave” CBT

The techniques used by ACT seem to work in the clinical field, although much more research is needed to get a steady position in the evidence-based movement. However, the specific mechanisms of change that are supposed to underlie these clinical changes and the efficacy/effectiveness of ACT techniques (included in ACT’s clinical protocols) are less investigated and, taking into account the above analysis of Descartes’ error, might even be fundamentally flawed. Therefore, future research should investigate the mechanisms of change involved in the positive outcomes of ACT techniques of “neutralizing” dysfunctional/irrational cognitions, in order to identify “the real”, not the phenomenological, “causes” of these changes. This clarification is important for ACT’s status in the evidence-based movement in psychotherapy (see David & Montgomery, 2011).

We believe that an old strong debate between the two traditional forms of CBT, namely Rational Emotive Behavior Therapy (REBT of Albert Ellis) and Cognitive Therapy (CT of Aaron T. Beck) might provide context to clarify this issue. Let us briefly analyze this debate (see for details David et al., 2010; Ellis, 2003).

Our cognitions can be examined in a causal chain (see for details David et al., 2010): (1) descriptions (e.g., “My wife is not at home”) prime various (2) inferences (e.g., “She is cheating on me”) that are then (3) evaluated/appraised (e.g., “It must not happen and it is awful if my wife is cheating on me”). This process, related to a negative activating event, can then lead to dysfunctional feelings and behaviors (see for details the ABC model of REBT/CBT - David et al., 2010).

Research in Cognitive Therapy has primarily focused on descriptions/inferences, identifying several classes of dysfunctional cognitions (e.g., “all or nothing thinking”, “jumping to conclusions”, “overgeneralization”) (see Beck, 1995). Rational Emotive Behavior Therapy has focused more on evaluations/appraisal in the form of irrational (e.g., rigid thinking/demandingness, catastrophizing/awfulizing, frustration intolerance, global evaluation of human worth) and rational (e.g., preferences/acceptance, non-catastrophizing/anti-awfulizing, frustration tolerance, unconditional self-, other- and life acceptance) beliefs. Indeed, consistent with Lazarus’ appraisal theory (Lazarus, 2001), Ellis (2003) even argued that a dysfunctional description/inference will not generate emotional problems, unless it is appraised irrationally (see for analysis David et al., 2010). If a dysfunctional description/inference (e.g., “My wife is cheating on me”) is appraised rationally, in terms of “strong preference” (“preference” plus “motivational relevance”) and “acceptance” (e.g., “I would strongly prefer it did not happen, but I can accept that sometimes things do not happen the way I want”), the outcome could be a functional negative feeling rather than a
dysfunctional negative feeling (e.g., sadness, but not depressed mood; concern, but not anxiety).

In this context, one should note that ACT focuses its main critical analysis of cognitive restructuring on descriptions/inferences, rather than on rational/irrational beliefs. However, according to REBT one can have a dysfunctional cognition, but if it is not appraised irrationally, it will not generate an emotional problem. Therefore, “neutralizing” a description/inference in ACT terms – which, however, seems to not be possible, taking into account Spinoza’s model - could simply mean that one changes (e.g., nuances and/or reduces) the “motivational relevance” of his/her dysfunctional cognitions. Changing “motivational relevance” could mean that one evaluates the distorted descriptions/inferences more rationally, meaning a change of irrational beliefs into rational beliefs (e.g., preferences/acceptance; anti-awfulizing; frustration tolerance; unconditional self-, other-, and life acceptance). However, contrasting ACT and REBT, ACT may emphasize the acceptance component of the rational belief – thus favoring a reduction in emotional reaction –, while REBT might emphasize both components (i.e., “strong preferences” and “acceptance”), and this could favor a functional negative feeling (i.e., concern and sadness), rather than a dysfunctional negative feeling (e.g., anxiety and depressed mood) and/or a reduction in our emotional reaction. However, ACT compensates the initial reduction in motivational relevance by activating values and their associated actions.

In essence, the ACT process of “neutralizing” a dysfunctional cognition may simply mean a change in its underlying evaluation (i.e., irrational belief). Thus, ironically, the process of “neutralizing” a dysfunctional cognition may amount to cognitive restructuring focused on the underlying evaluation of this distorted cognition, rather than focused directly on the dysfunctional descriptions/inferences. In our example, one might learn to accept the thought “My wife is cheating on me”, but will not be anxious because (1) the demand (e.g., “It should not happen”) and the catastrophizing/awfulizing (e.g., “It is awful if my wife is cheating on me”) have been changed by cognitive restructuring into their rational counterparts [(a) preferences/acceptance (e.g., I would have liked it not to happen, but I can accept that things do not always happen the way I want”) and (b) anti-awfulizing (e.g., “It is bad, but not the worst thing that could happen”)] and thus, he would feel concerned (i.e., functional negative feeling) or (2) the “motivational relevance” of the thought has been reduced (i.e., by cognitive restructuring focused on the reduction of the intensity of “preferences” and the enhancement of the “acceptance” component of the rational beliefs) and thus, he would experience a reduction in the anxiety (up to a “neutral” feeling). A cognitive restructuring of irrational beliefs can also lead to (1) a reduction in the frequency of distorted cognitions evaluated irrationally and/or to (2) maintaining distorted cognitions, but “neutralizing” them (in terms of reducing “motivational relevance”) and therefore, reducing their negative consequences (see for details David et al.,
Thus, the results of the cognitive restructuring of irrational beliefs are mixed and similar to the effect of ACT’s “neutralizing” techniques (e.g., cognitive defusion) on the frequency of distorted cognitions. Future research should clarify these mechanism and effects.

It is interesting to note that ACT practitioners opposed cognitive restructuring process so strongly because in their early writings they often confused cognitive restructuring with suppression, avoidance, and/or control of negative thoughts. Although this confusion seems to be understood and corrected now, the old negative halo against the cognitive restructuring process is still limiting ACT research and development.

Raising these criticisms and debating ACT theory in the context of cognitive science should not prevent us from noting that ACT (and other third wave CBT) contribution to the CBT field is very important in bringing and/or promoting new techniques of cognitive restructuring (sic!) (e.g., cognitive defusion, mindfulness) - useful when classic CBT cognitive restructuring techniques do not work properly -, although the underlying theory is still debatable, based on Descartes' potential error.

REFERENCES


Descartes' Error and CBT


