



HAL
open science

Expectation violation and cognitive dissonance theory: Proposal for an epistemic inconsistency management model

Benjamin Aubert-teillaud, Alexandre Bran, David C Vaidis

► To cite this version:

Benjamin Aubert-teillaud, Alexandre Bran, David C Vaidis. Expectation violation and cognitive dissonance theory: Proposal for an epistemic inconsistency management model. *European Journal of Social Psychology*, 2023, 53 (7), pp.1663-1679. 10.1002/ejsp.2981 . hal-04190471

HAL Id: hal-04190471

<https://hal.science/hal-04190471>

Submitted on 29 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Expectation violation and cognitive dissonance theory: Proposal for an epistemic inconsistency management model

Benjamin Aubert-Teillaud¹  | Alexandre Bran¹  | David C. Vaidis² 

¹Laboratoire de Psychologie Sociale, Université Paris Cité, Boulogne-Billancourt, France

²CLLE, Université de Toulouse, CNRS, Toulouse, France

Correspondence

Benjamin Aubert-Teillaud and David C. Vaidis, Laboratoire de Psychologie Sociale, Université Paris Cité, 71 avenue Edouard Vaillant, F-92100, Boulogne-Billancourt, France. Email: benjamin.aubert92@gmail.com and david.vaidis@univ-tlse2.fr

Abstract

A crucial dilemma for any individual facing contradictory information is whether to maintain or disengage from an expectation. Cognitive dissonance theory offers an influential model of these situations. However, this theory does not provide clear predictions regarding the use of reduction strategies (methods used to resolve inconsistency) and their consequences. We propose a decision rule and parameters to model decisions faced by individuals who encounter epistemic inconsistency. Inconsistency leads individuals to assess the probabilities and costs associated with the error of disengaging from a correct expectation and the error of maintaining an incorrect expectation. Depending on these assessments, individuals opt for a reduction strategy allowing for maintenance of or disengagement from the expectation. We propose a categorization of reduction strategies derived from this process and oriented towards the explanation of expectation bolstering, contributing to the understanding of radicalization. This model offers new predictions and avenues for empirical work.

KEYWORDS

cognitive dissonance, consistency, inconsistency, prediction error, radicalization

1 | INTRODUCTION

Psychological scientists have studied extensively the behavioural consequences of a perceived contradiction between existing knowledge of the state of the world and new information. This examination of inconsistency is the foundation of cognitive dissonance theory (CDT) (Festinger, 1957; Harmon-Jones, 2019), one of the most influential models in psychology (Devine & Brodish, 2003; Gawronski & Strack, 2012; Haggbloom et al., 2002). However, the concept of inconsistency at the heart of CDT needs more theoretical integration, especially in regard to other fields of cognitive science dealing with less abstract prediction errors, such as errors at the perceptual level (e.g., Rescorla & Wagner, 1972; Sutton & Barto, 1998). In this regard, few connections have been drawn between the concepts of inconsistency and prediction error (Friston, 2018; Kaaronen, 2018; Kruglanski et al., 2018;

Proulx et al., 2012). Moreover, although several reduction strategies (i.e., means used to resolve inconsistency) have been identified, reviews have indicated that CDT lacks a model of decision rules, leading it to favour a certain type of reduction in specific inconsistency situations (McGrath, 2017; Vaidis & Bran, 2018, 2019). Finally, researchers have focused on very specific paradigmatic situations. This has impeded understanding of the decision-making process as a whole and has limited the examination of reduction strategies other than traditionally assessed attitude change (Batson, 1975; Devine et al., 1999; Vaidis & Bran, 2018). In this article, we suggest a theoretical formulation of a decision model for regulation in a situation of epistemic inconsistency.

The following model is based on an integration of CDT (Festinger, 1957) with the concept of prediction error (Friston, 2010; Rescorla & Wagner, 1972; Sutton & Barto, 1998). We approach CDT from the perspective of the epistemic violation, which involves an inconsistency

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *European Journal of Social Psychology* published by John Wiley & Sons Ltd.

between an expectation regarding the state of the world and contradictory information¹. This integration contributes to a comprehensive understanding of the regulation of high-epistemic-complexity inconsistencies, such as inconsistencies related to opinions (e.g., Adams, 1961), religious beliefs (e.g., Festinger et al., 1956), or scientific theories (e.g., Kuhn, 1970).

Recent integrative models have also suggested that inconsistency is related to expectancy violation (Heine et al., 2006; Proulx & Inzlicht, 2012). However, these models focused particularly on indirect regulation and did not propose a decision rule and parameters regarding the choice of direct regulation. The objective of this article is to suggest a decision rule regarding inconsistency reduction and a categorization of reduction strategies to generate new predictions. For this purpose, we focus on four fundamental points: (i) defining inconsistency clearly; (ii) describing the decision-making process and important parameters in the choice of a reduction strategy; (iii) categorizing reduction strategies according to the reduction process described, and (iv) identifying the consequences of using reduction strategies. We emphasize the bolstering of expectations after an inconsistency. This can contribute to the process of radicalization.

2 | REDEFINITION OF THE CONCEPT OF INCONSISTENCY

Building on the original definition of dissonance (Festinger, 1957; Harmon-Jones, 2019), the concept of inconsistency refers to a logical incompatibility between two cognitions. It is assumed that 'the relation between the two elements is dissonant if, disregarding the others, the one does not, or would not be expected to, follow from the other' (Festinger, 1957, p. 15). From this perspective, considering two cognitions x and y , there is inconsistency when y derives from not- x . For example, if one considers the relation x , *Paul is polite so his behaviour is expected to be nice*, and observes y , *Paul's behaviour is rough*, there is inconsistency because y derives from the negation of x . Starting from the initial formulation that inconsistency implies an incompatibility between two cognitions, we consider that the two cognitions do not have the same epistemic status. Indeed, for y to derive from not- x , it is necessary that x exists prior to y . For example, if I have no knowledge concerning Paul (hence no expectation), then Paul's rude behaviour cannot generate inconsistency. At a psychological level, inconsistency thus always presupposes prior knowledge about the state of the world that is invalidated by later information. According to this reasoning, if inconsistency presupposes prior knowledge about the state of the world, which implies how the world is and must be with a certain

probability—because a state of the world can never be totally certain (Hume, 2007)—then this corresponds to the definition of the concept of *expectation*. Furthermore, according to the above, for inconsistency to exist it is necessary that later information contradicts an expected state of the world. In the remainder of this paper, this later cognition will be called *contradictory information*. Let us specify that if the logical sequence described by Festinger is necessary to generate inconsistency, it is not sufficient. Indeed, for an inconsistency to occur, the situation must be relevant for the individual, which implies adherence to the premises, that is, the expectation. In the case where the individual does not adhere to the premises, there would only be a formal contradiction, not an inconsistency.

Following this reasoning, we suggest specifying the definition of inconsistency by adding a difference in epistemic status between the cognitions involved in the inconsistency. As we will see later, this refinement of the notion of inconsistency from Festinger's original concept leads to important consequences in terms of regulation modelling. The integration of the concept of expectation in CDT in our proposal has already been proposed in the meaning maintenance model (MMM) Proulx & Inzlicht, 2012). We distinguish *expectation* from *contradictory information*, so this terminology could require a clear identification of these two cognitions in the different CDT paradigms. To help explain its relevance, Supporting Information 1 summarizes the identification of *expectation* and *contradictory information* in the main CDT paradigms.

From the previous reasoning, we assume an overlap between the notion of inconsistency and that of prediction error, which also describes a cognitive discrepancy between *what is expected* and *what is observed* (Rescorla & Wagner, 1972; Sutton & Barto, 1998). This overlap between the concepts of prediction error and inconsistency enables us to position our proposal within the broader framework of predictive processing (PP) theories. According to PP models, perception, cognition and action are driven by the same inferential processes, allowing a reciprocal dialogue between top-down predictions about the state of sensory input and ascending error signals (Clark, 2013; Friston, 2010). Gaps between predictions and incoming information allow us to update predictions at each hierarchical level (e.g., Friston, 2010). Beyond the interest in providing a unified model of the mind associated with increasing empirical evidence (Walsh et al., 2020), PP models make an ontological proposition by considering biological agents as mainly organized around a motivation for reducing entropy (Friston, 2010).

Settling this difference between the nature of the two cognitions implies an epistemic change in the characterization of the regulation process. For CDT, the regulation process has been considered to be driven mainly by the question of which cognition is more resistant to change (e.g., Vaidis & Bran, 2018). In our proposal, because we distinguish the initial expectation from contradictory information, the regulation process aims at deciding whether the expectation should be abandoned or maintained, despite contradictory information. This distinction helps us clarify and suggest a new perspective on inconsistency regulation. For instance, in our proposal, the regulation strategy by adding cognition (Festinger, 1957) corresponds solely to a regulation

¹ This perspective partially excludes other qualitatively different situations of cognitive dissonance such as the free choice paradigm (Brehm, 1956). The effect of the choice paradigm, namely *spreading of alternatives*, has been methodologically discussed in recent decades, considering that it could only reflect an initial preference for one of the alternatives (see Chen & Risen, 2010; for review, Izuma & Murayama, 2013). From a theoretical perspective, this paradigm is more or less compatible with some conceptions of CDT. Festinger's initial 1957 theory is compatible with an explanation based on expectations but as far as we are aware this has not been developed or even conceptually discussed (see for instance the MMM). To avoid misunderstanding, we thus preferred to put this paradigm to the side in the current development of our model.

aimed at maintaining expectations against contradictory information. For example, if one believes that the Earth is flat and receives contradictory information that invalidates this expectation then a solution is to change the expectation; adding cognitions does not make sense in order to reduce inconsistency here. Conversely, if one considers that the Earth is still flat despite contradictory information then one can add alternative causal explanations to justify contradictory evidence while maintaining the expectation. Only in this latter situation is adding cognitions a means of reduction. To sum up, the concept of regulation by adding cognition makes sense within the framework of the original definition of inconsistency (Festinger, 1957), which does not involve differences in epistemic statuses among the cognitions. In our proposal, adding cognitions no longer has the same meaning because it can only be carried out as a regulation strategy aimed at maintaining the expectation in the face of contradictory information. Our proposal suggests that adding cognitions aims to save the expectation by explaining why the contradictory information has to be false. An example can be found again in Festinger et al.'s study (1956, p. 162): Several sect members explained the absence of a midnight apocalyptic event because of a clock that would have been incorrectly set. We propose to refer to this regulation as *ad hocness*, in reference to an epistemological concept that encompasses the search for alternative causality to explain contradictory information while preserving the expectation (Boudry, 2013; Chalmers, 2013; Lakatos & Musgrave, 1970).

3 | DECISION RULE AND PARAMETERS OF AN INCONSISTENCY MANAGEMENT MODEL

The initial proposition of our model of epistemic inconsistency management (EIM) can be expressed as follows: in the empirical reality, given the impossibility of objectively ruling out the refutation of an expectation from a single contradictory piece of information, an individual must establish an *epistemic bet* on the status of the contradictory information. We use the concept of *bet*—in the vein of Pascal's wager in philosophy—in that it refers to a choice under uncertainty between options with asymmetrical error costs. Concerning inconsistency, there can only be two outcomes of this bet: either the expectation is considered to be invalidated by contradictory information or it is concluded that the expectation remains valid despite contradictory information. Before developing our model, we briefly discuss the necessity of making a bet on the epistemic status of contradictory information and of deploying defensive strategies to maintain the expectation.

3.1 | Pragmatic bet versus naive falsificationism

A simple decision rule facing inconsistency would be to reject the expectation following any contradictory information. However, this rule, called *naive falsificationism*, would be doomed to adaptive setbacks due to the *fallibility of the observation statements* (Chalmers, 2013). As

empirical observations are fallible, this excludes the possibility of a formal and conclusive refutation via a decisive experiment (see also Kuhn, 1970; Lakatos, 1968; Popper, 2005).

The impossibility of deciding from a naive falsification rule has two consequences. First, individuals have to make a bet on the status of contradictory information either by considering that the expectation is invalidated by contradictory information or by considering that the expectation remains valid despite contradictory information. Second, because contradictory information is fallible, it may be adaptive to develop ad hoc hypotheses—that is to preserve an initial expectation with an alternative explanation. Indeed, a cognitive mechanism unable to consider that inconsistent data could be false or explainable by a hidden cause would lead to excessive abandonment of correct expectations, which would be problematic for the individual's adaptation. In science, the development of empirically testable ad hoc hypotheses for protecting a theory from refutation is adaptively legitimate due to the need not to over-reject theories at first sight (theory protection) and to seek hidden causes, themselves potential vectors of scientific discoveries (theory extension). For example, the inconsistency generated between the Newtonian theory and the observed trajectory of Uranus led scientists to look for the hidden cause of this discrepancy, formulating a testable hypothesis of an unknown planet. In a nutshell, the formulation of empirically testable ad hoc hypotheses as a defensive strategy is important for the development of science, protecting new theories from initial difficulties while giving them the opportunity to expand (Chalmers, 2013; Feyerabend, 1975/1993; Lakatos & Musgrave, 1970).

Through this example, we demonstrate that the mechanism we are about to describe for managing inconsistency should employ strategies that enable the defence of the expectation. To sum up, an individual must be able to disengage from expectations (non-defensive strategy) but also to build up defensive strategies to maintain them. Which of these inconsistency reduction strategies is chosen depends on the outcome of the bet, which is based on both subjective probabilities and the relative costs of errors in the choice to maintain or disengage the expectation.

3.2 | Decision rule for inconsistency processing

Decisions are made systematically with a certain amount of uncertainty and thus errors are always possible. The processing of inconsistency therefore requires subjective probabilities and relative cost of errors to be taken into account. For example, incorrect disengagement relative to an important expectation for survival would have a crucial impact and thus should reduce the likelihood of disengagement.

When faced with contradictory information, individuals confront two possibilities: on the one hand, they can maintain their initial expectation despite contradictory information and, on the other hand, they can disengage from the expectation. The uncertainty and the fallibility of contradictory information result in the decision maker facing two potential errors. They can consider contradictory information to be

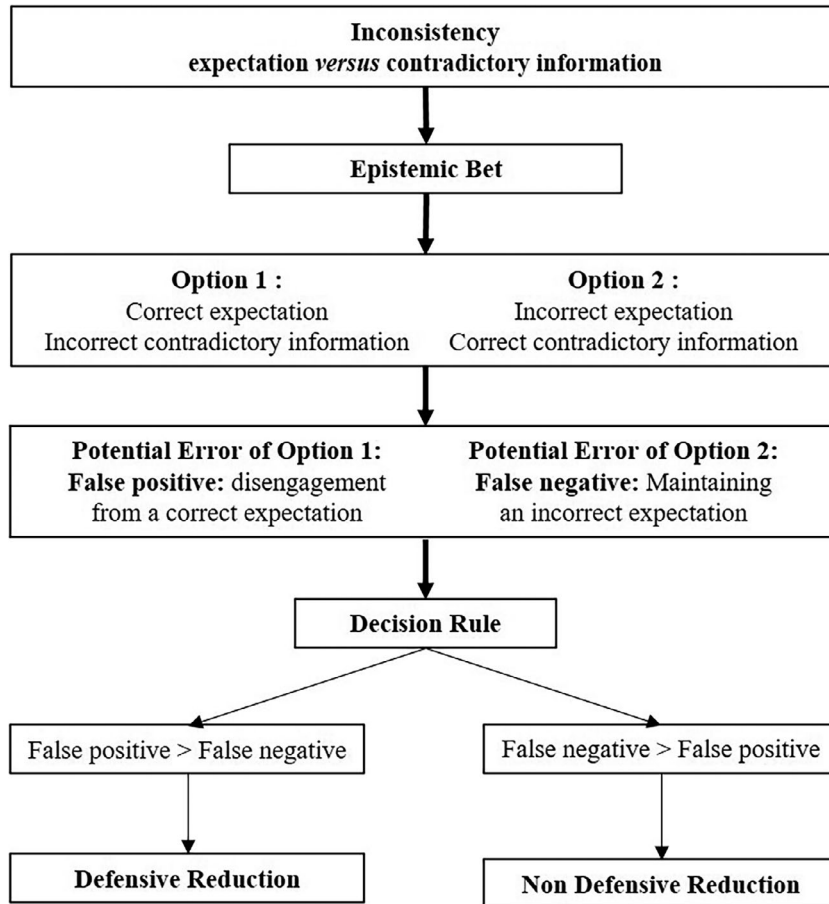


FIGURE 1 Decision process for epistemic inconsistency management.

correct when it is not and thus disengage from a correct expectation (i.e., *false positive*). On the other hand, they can consider contradictory information to be incorrect when it is not and thus maintain an incorrect expectation (i.e., *false negative*). For example, a fisherman who expects a fishing spot to be rich and who receives information that there are no more fish must make a decision. The fisherman can either consider his initial expectation to remain correct and the contradictory information to be incorrect or he can decide that his expectation is no longer correct and that the contradictory information is correct. Faced with this decision, the two possible errors are to disengage from a still abundant fishing spot (false positive) or to stay at an actually depleted fishing spot (false negative) (Figure 1). This decision making is reminiscent of the exploitation-exploration dilemma (Bland & Schaefer, 2012; Courville et al., 2006; Daw et al., 2006) where a prediction error leads the individual to regulate either by maintaining the exploitation of a known alternative or by favouring the exploration of a new alternative (Cohen et al., 2007). Here again, the choice of maintaining exploitation or of exploring can generate the same type of error, by continuing to exploit a suboptimal alternative or by abandoning a profitable alternative in favour of uncertain exploration.

These considerations of the costs of errors are included in error management theory (EMT) (Haselton & Buss, 2000). According to EMT, individuals bias their decisions and adopt beliefs according to the asymmetry of the costs of possible errors (Haselton & Nettle, 2006;

Haselton et al., 2016; Johnson et al., 2013). In a situation involving uncertainty, decision making could lead to Type I errors (false positive: adopting a belief or a decision when it turns out to be incorrect) or Type II errors (false negative: not adopting a belief or a decision when it turns out to be correct). Depending on costs and probabilities relating to these two types of errors, the individual's decision will be biased in favour of the error that is the least likely and with the least negative consequences (Haselton & Nettle, 2006). Regarding the regulation of inconsistency, if individuals must take into account probabilities of a false positive and a false negative, they must also weigh these probabilities according to the asymmetry of the cost of errors linked either to an incorrect disengagement from their expectation or to an incorrect maintenance of the expectation. To implement this decision rule, individuals must: (a) evaluate the probability of a false positive (disengagement from a correct expectation based on incorrect contradictory information), (b) evaluate the probability of a false negative (maintaining an incorrect expectation on the basis of correct contradictory information), (c) weigh the probability of a false positive based on the anticipated consequences of a false positive, and (d) weigh the probability of a false negative according to the anticipated consequences of a false negative.

We can summarize the decision criterion taking into account the probabilities of errors and their costs based on the following equation, where C represents the decision criterion, P_{fp} the probability of a false positive, P_{fn} the probability of a false negative, v_{fp} the cost of a false

positive and v_{fn} the cost of a false negative:

$$C = (P_{fn} * v_{fn}) - (P_{fp} * v_{fp})$$

From there, we can determine a *disengagement threshold*, which is the point at which individuals decide to disengage from their expectations. This decision rule is thus associated with different strategies for coping with inconsistency. Indeed, according to this calculation, if criterion C is positive, it indicates that the probability of a false negative (maintaining an incorrect expectation despite correct contradictory information), weighted by its consequences, exceeds the probability of a false positive (changing a correct expectation based on incorrect contradictory information), weighted by its consequences. In this case, individuals should choose a non-defensive reduction strategy, allowing them to disengage from their expectations. For CDT, these strategies correspond to *attitude change* (e.g., Festinger & Carlsmith, 1959) or *cognitive restructuring* (Hardyck & Kardush, 1968; Leippe & Eisenstadt, 1994). Conversely, if criterion C is negative, it implies that the probability of a false positive (changing a correct expectation based on incorrect contradictory information), weighted by its consequences, exceeds the probability of a false negative (maintaining an incorrect expectation despite correct contradictory information). In this case, individuals should opt for the maintenance of their expectations and for defensive reduction strategies. For CDT, this includes strategies such as *trivialization* (e.g., Simon et al., 1995), *denial of responsibility* (Gosling et al., 2006), *ad hocness* (Burriss et al., 1997; Festinger et al., 1956), or *selective exposure* (e.g., D'Alessio & Allen, 2002). Our model suggests two types of regulation strategies: non-defensive strategies that result in disengagement from the expectation and defensive strategies that defend the expectation in the face of contradictory information. These align with the concepts of *accommodation* and *assimilation* suggested by the MMM (Proulx & Inzlicht, 2012) and grounded in Piagetian theory (1954). Accommodation adapts the initial framework to incorporate contradictory information, whereas assimilation modifies the meaning of a contradictory experience to align with the expectation.

3.3 | Overview of the parameters

To make a decision when faced with contradictory information, individuals must take into consideration several parameters (see Figure 2) that allow them to evaluate the probability and the cost of the two types of errors (false positive and false negative). Concerning the probability of a false negative and a false positive, individuals must take into account the *certainty* of their expectations and the *reliability* of contradictory information. *Certainty* refers to the quantity of evidence accumulated (e.g., Hume, 2007; Kiani et al., 2014). The more regularly and consistently an expectation has been reinforced, the greater the certainty should be and the greater the probability of a false positive (disengagement from a correct expectation on the basis of incorrect information). For PP models, *reliability* depends on the precision of the error signal (e.g., Clark, 2018). The more contradictory information is accurate and informative, the more the individual should consider it to be reliable

and the greater is the probability of a false negative (i.e., maintaining an incorrect expectation).

In addition to the probabilities of the two types of error, individuals must also consider their costs. The cost of a false positive can be defined as the consequences of individuals incorrectly disengaging from their expectation, whereas the cost of a false negative refers to the consequences of the individual maintaining an incorrect expectation. For example: consider a sedentary person and a professional athlete. They both believe that regular consumption of meat enhances athletic performance. If they are presented with information that contradicts this belief, if they change their minds, the cost of being wrong would differ: giving up meat would be minor for the sedentary person whereas it would be catastrophic for the athlete. In other words, the greater the negatives consequences of incorrectly disengaging from the expectation are, the more individuals should be biased towards false negatives (maintaining an incorrect expectation) to avoid false positives (disengaging from a correct expectation) and thus invest in a defensive strategy (e.g., denial or ad hocness). Conversely, the greater the negative consequences of incorrect expectation maintenance, the more the individual should be biased towards false positives (correct expectation disengagement) to avoid false negatives (incorrect expectation maintenance) and thus opt to a defensive strategy (e.g., attitude change).

To sum up, we can predict whether individuals will adopt a defensive or non-defensive strategy by considering how certain they are about their belief, the reliability of conflicting information and the anticipated cost of the two types of error (see Figure 2). Furthermore, although we have suggested that the different parameters of the model may be related to metacognitive judgements of certainty or reliability, we do not consider that it is necessary for the decision process to be conscious. Indeed, as we have presented it, certainty is a function of the amount of evidence accumulated and reliability of contradictory information is a function of the precision of the error signal.

4 | CERTAINTY OF EXPECTATION

According to Hume (2007), the attribution of causality depends on the uniformity and repetition of experience. The degree of certainty of an expectation depends on the amount of evidence accumulated previously. This is the first parameter to be taken into consideration to predict reduction. The degree of certainty regarding a belief raises the disengagement threshold and therefore contributes to the occurrence of defensive reduction strategies oriented towards the maintenance of the expectation. Data from decision science support this assumption: the more an expectation receives confirmatory evidence, the more the degree of confidence increases (e.g., Kiani et al., 2014; Nassar et al., 2010). The more certain the expectation is, the less new information affects individuals (e.g., Behrens et al., 2007; Nassar et al., 2010; Vaghi et al., 2017). Starting from our model, this section relies on the CDT literature to illustrate that the greater the epistemic support an expectation receives, the more likely a defensive reduction strategy becomes. In certain contexts favouring the confirmation of

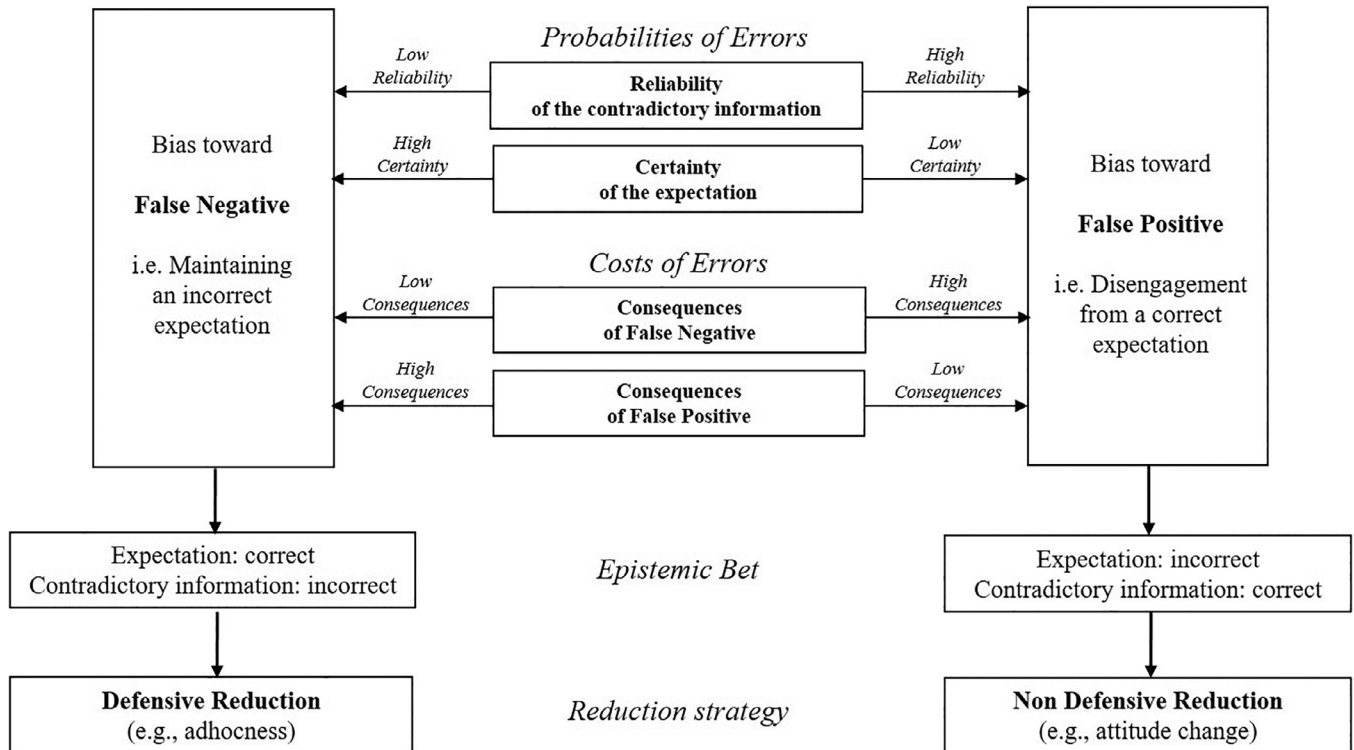


FIGURE 2 Parameters and outcomes of the epistemic inconsistency management model.

expectations, such as sectarian or radicalized groups, defensive reductions are also likely to bolster these expectations and push individuals towards radicalism.

4.1 | Belief disconfirmation: Group isolation increases certainty

Group isolation, which can be defined as the tendency to interact only with individuals sharing the same framework of beliefs, can raise the degree of certainty of expectations. Group isolation is a typical situation in radicalized groups (Doosje et al., 2016; McCauley & Moskalenko, 2008). In a way, group isolation facilitates the validation of expectations thanks to confirmatory information and prevents the existence of error signals. Examples of such occurrences in CDT studies could refer to sect observations (e.g., Hardyck & Braden, 1962) that can constitute echo chambers (e.g., Cinelli et al., 2021) producing confirmatory information that reinforces the expectation certainty and results in defensive regulation strategies facing a prophetic failure. According to our decision rule, increasing the subjective certainty of expectations by adding confirmatory evidence due to group isolation increases the perceived probability of a false positive when facing contradictory information and therefore the probability that the individual will use defensive reduction strategies when faced with contradictory information.

In their seminal study, Festinger et al. (1956) relate the failure of sect members' prophecy, which had announced the end of the world.

Contradictory information was the same for all members of the sect but the authors noted differences in the reduction strategies depending on the subgroups. The Lake City group was described as more cohesive and unanimous, with more social support helping to generate information-validating belief systems. Conversely, the lesser unanimity in the Collegeville group may have shaped expectations that were less certain, providing less confirmatory evidence and more divergent information (Festinger et al., 1956; Weiser, 1974). These observations are consistent with our model. The more isolated and united Lake City group used more defensive strategies than the Collegeville group, most of whose members disengaged from their expectations and left the sect. In the same vein, the observational study of Hardyck and Braden (1962) provides a second example. In this study, the authors observed sect members who anticipated a nuclear attack and who holed-up in bunkers for 42 days. As in the Lake City group, the group members facing prophetic failure maintained their beliefs. Group isolation, the authors note (p. 138), may have helped to bolster their expectations by leading individuals to interact exclusively with peers sharing the same belief system.

In PP models, these elements pertaining to group isolation are also consistent with the fundamental motivation for the long-term reduction of surprise proposed by predictive coding (Friston, 2005, 2010). Even though it would be absurd to consider locking oneself in a dark room to be a good way to completely reduce error (Friston et al., 2012; Sun & Firestone, 2020), it may nevertheless be the case that for certain beliefs considered central, an individual may look for environmental conditions limiting the presence of signals conveying error. In the PP

models, this possibility is suggested as an explanation of confirmation bias (Clark, 2016). This situation seems to correspond particularly to that reported by Hardyck and Braden (1962, p. 138) in which specific circumstances ensuring the absence of contradictory information and the accumulation of confirmatory information were selected intentionally. This motivation to reduce surprise could partly explain the strong tendency for radicalized groups to look for isolation or social support, which contributes to the radicalization process (Doosje et al., 2016; McCauley & Moskaleiko, 2008; Moghaddam, 2005). Sectarian or radicalized groups could constitute *social dark rooms* whose central function is to limit the presence of error signals proactively while continuing to deliver confirmatory information. Group isolation represents a proactive reduction strategy, allowing the protection of central beliefs whose abandonment could produce a psychological collapse. The consequence of group isolation is also to increase the certainty of expectations and thus to raise the probability of a false positive in case of contradictory information and it ultimately raises the likelihood of the occurrence of defensive strategies. In particular, in studies relative to religious or occult beliefs (Burris et al., 1997; Festinger et al., 1956; Hardyck & Braden, 1962), this defensive strategy can involve adding an ad hoc hypothesis to preserve the expectation (i.e., adhocness).

4.2 | Advocacy and attitude saliency increase expectation certainty

The level of certainty of an expectation can also be increased through experimental manipulations such as writing arguments in favour of an expectation, as is the case in the induced hypocrisy paradigm (Aronson et al., 1991). This procedure consists of generating inconsistency by stressing the gap between a normative position defended by an individual and the recall of behaviours that deviated from this expectation (Aronson et al., 1991; Dickerson et al., 1992). We assert that the change in behaviour is characterized by the performance of new behaviours consistent with the initial expectation—that is, the sustained standard. Relative to the expectation, the occurrence of behaviours in line with the normative position, therefore, maintains it and even bolsters it (Yousaf & Gobet, 2013).

In the hypocrisy paradigm, the individual is led to bolster an expectation by supplying information supporting it. For instance, participants generate arguments to encourage others to wear condoms (Aronson et al., 1991; Stone et al., 1994) or plead for pro-ecological positions (Fointiat et al., 2013). During the advocacy, this accumulation of evidence could increase the certainty of the expectation. Consequently, contradictory information in the form of reminders of transgressions is more likely to be below the disengagement threshold and thus promote a defensive reduction to maintain the expectation. The defence of the expectation against the recalled contradictory information is manifested in this paradigm by the performance of behaviours consistent with the normative behavioural expectation. For instance, in a religious context, Yousaf and Gobet (2013) led their participants to contradict their dogma by recalling inconsistent past behaviours. The authors observed a bolstering of religious behaviours and attitudes in

the hypocrisy condition relative to the control condition. In our view, this study is interesting because its results suggest that the behavioural change observed within the framework of this paradigm could be based a priori on reinforcing the expectation.

The hypocrisy paradigm often presents a situation involving simple recall of a transgression and a situation involving recall of a transgression in conjunction with advocacy in favour of the expectation (e.g., writing arguments in favour of the environment; Aronson et al., 1991; Stone et al., 1994). Dissonance theorists assume that individuals initially agree with the position they have to defend and this is why transgression recall produces dissonance. This means that the reminder of a transgression without accompanying advocacy already constitutes a situation of inconsistency. The difference between the two conditions involving inconsistency is therefore that, for one of these conditions, the expectation is bolstered by adding confirmatory information, whereas it is not for the other condition. In terms of our model, this suggests that, in advocacy conditions, adding information raises the certainty of the expectation and increases the likelihood of defensive reduction compared to the other condition. Research within this paradigm supports this proposal, with more behaviours in defence of the expectation in conditions with advocacy and transgression (Stone & Fernandez, 2008).

To sum up, an increase in the amount of confirmatory information tends to strengthen the certainty of expectation, making the use of defensive strategies more likely. Some of these strategies, such as adhocness, ultimately reinforce the expectation certainty and thereby contribute to radicalization of beliefs. The use of defensive strategies can be problematic and lead to radicalization but it can also serve pro-social interests, particularly when it concerns pro-normative positions (e.g., Aronson et al., 1991).

4.3 | Importance of the expectation

To weigh the costs associated with the potential errors effectively in the decision process, it seems relevant for individuals to consider the importance of their expectations in their interactions with the environment. For instance, if someone is stranded on a deserted island, being proven wrong about the sphericity of Earth, whatever its associated certainty, would not be as crucial as being incorrect about the location of a fishing spot. This example stresses the distinction between importance and certainty. Cognitive dissonance theory has also considered the importance of attitude as a factor in resistance to change, which modifies the process of regulation (Festinger, 1957).

Many studies examined the role of the importance of attitudes but the concept of importance has not yet been defined clearly. Indeed, the importance of an attitude is often defined as being personally preoccupied or involved with the object of the attitude (Eaton & Visser, 2008; Howe & Krosnick, 2017; Krosnick, 1988). This definition of importance is flawed because it merely substitutes one term with its synonym, without offering a regression to the simpler elements corresponding to the essence of the importance judgement. To address this issue we suggest defining importance as the utility judgement of an object

depending on the degree of its positive or negative consequences for individuals. For example, the consequences of gender equality should appear more significant for a staunch defender of feminism than for a non-feminist: when exposed to inconsistency, the former regulates more defensively than the latter (Sherman & Gorkin, 1980). In the same way, if an empirical anomaly endangers a hypothesis, its impact would differ depending on the centrality of this hypothesis for the field and its applications. If the hypothesis is central, revising the theory due to an anomaly can be costly, particularly if the anomaly turns out to be a false positive. Mechanically, an expectation, the consequences of which have a significant impact on individuals' adaptation to their environment, should increase the cost of a false positive in the event of confrontation with contradictory information and should increase the probability of a defensive reduction. To sum up, an important expectation increases the evaluated cost of the error linked to an incorrect disengagement and pushes the individual towards defensive reduction.

Several CDT studies have examined the impact of attitude importance on reduction strategy selection specifically within the induced compliance paradigm (Cooper & Mackie, 1983; Eisenstadt & Leippe, 2005; Sherman & Gorkin, 1980; Simon et al., 1995; Starzyk et al., 2009). For example, Simon et al. (1995, Study 4) showed that when the topic of a counterattitudinal advocacy was particularly important, the attitude change was blocked and participants regulated defensively by reducing the importance of contradictory information (i.e., trivialization). Likewise, Sherman and Gorkin (1980) showed that performing a counterattitudinal behaviour strengthened central attitudes while non-central attitudes were altered.

The effect of attitude importance on regulation has also been demonstrated in studies where participants were exposed to contradictory information (Chaiken et al., 1995; Jacks & Devine, 2000). For example, Zuwerink and Devine (1996) showed that when the topic was important to the individuals, being presented with contradictory information led to more resistance and less attitude change. According to the authors, when an attitude is important, resistance is linked to a process of cognitive elaboration associated with the search for counterarguments. This is a hypothesis close to the proposed adhocness regulation.

5 | RELIABILITY OF CONTRADICTORY INFORMATION: A SECOND-ORDER PREDICTION

When faced with contradictory information, individuals must be able to assess the reliability of the information to distinguish between relevant and trivial error signals. We define *reliability* as the degree of precision in contradictory information, or in other words, the extent to which the contradictory information informs individuals about their expectations. For example, if an environmentalist is opposed to the use of planes but uses one anyway, the information conveyed by this behaviour will be more reliable for the evaluation of the environmentalist's attitude if the purpose of the journey is a freely chosen vacation than if it is an urgent trip to see a seriously ill family member. This proposal is supported by PP models for which the influence of a prediction

error on expectation updating depends on the reliability of the error signal (Clark, 2018; Feldman & Friston, 2010). The precision assigned to an error is related to a second-order prediction concerning the reliability of incoming evidence. The more reliable the evidence, the more weight an individual assigns to the error signal (Haarsma et al., 2020; Walsh et al., 2020). Conversely, an error that is considered to be unreliable should have little or no impact on expectation updating (Diederer et al., 2016; Mathys et al., 2011) as a mechanism that incorrectly predicts the reliability of an error message would turn out to be adaptively deleterious (Corlett et al., 2007; Haarsma et al., 2020).

Inconsistency between complex pieces of knowledge should also lead to a second-order prediction regarding the reliability of evidence in order to weigh contradictory information. With the use of our decision rule, a piece of information assessed to be reliable should increase the probability of a false negative and thus favour disengagement from the expectation. In contrast, contradictory information assessed to be unreliable should increase the probability of a false positive and thus elicit defensive strategies. Furthermore, it could be the case that the error signal is so unreliable that it generates no processing at all. The absence of a reduction strategy in such situations was suggested within the framework of CDT (e.g., Hardycy & Kardush, 1968).

For CDT, the notion of reliability, when it concerns information relating to the production of a counterattitudinal behaviour, results in classic manipulations of external justifications. For example, inconsistent behaviour performed for low rewards (e.g., Aronson & Carlsmith, 1963; Festinger & Carlsmith, 1959) or with a high degree of choice (e.g., Brehm & Cohen, 1962; Harmon-Jones et al., 1996) should represent more reliable information for the individual. Inconsistent behaviour occurring in the absence of external justifications should provide a more reliable error signal (i.e., a more informative one) for updating expectations. Conversely, all external justifications can be considered as noise, reducing the reliability of information. Minimizing external justifications (e.g., low reward or high choice) could increase the reliability of contradictory information, thus increasing the perceived probability of a false negative. This, in turn, could increase the likelihood of non-defensive reduction (e.g., attitude change).

Exposure to information contradicting our expectations represents another situation involving inconsistency that requires the estimation of the reliability of the signal. For instance, someone who strongly believes in the effectiveness of a medical treatment should evaluate the reliability of a communication relating to its ineffectiveness to avoid mistakenly abandoning a drug that is actually effective. If the source of contradictory information is considered to be unreliable, this should reduce the weight of the error signal and thus increase the probability of a false positive. In this case, a defensive reduction strategy could be to consider the source to be delivering a false message purposefully, for instance in the framework of a conspiracy theory (i.e., an ad hoc theory). Several CDT studies have explored the impact of source credibility when exposed to contradictory information (Bergin, 1962; Bochner & Insko, 1966; Eagly, 1969). For example, Aronson et al. (1963) demonstrated that a source with low credibility decreases attitude change and promotes the adoption of defensive reduction strategies.

Finally, the misattribution paradigm (Zanna & Cooper, 1974) offers another example of the manipulation of the quality of the error signal. Indeed, as PP theories suggest, the presence of a noise source obscuring the error signal is likely to impact the updating of expectations by reducing the weight of the error signal (e.g., den Ouden et al., 2012). In CDT paradigms, exogenous sources manipulated in the misattribution protocols are likely to reduce the reliability of the error signal, resulting in less attitude change (e.g., Zanna & Cooper, 1974) and more defensive strategies (e.g., Joule & Martinie, 2008).

6 | CONCLUSION REGARDING THE INCONSISTENCY BET

Starting from the adaptive challenge relating to the decision to maintain or abandon an expectation in the face of contradictory information, we proposed a decision rule in order to model the reduction of inconsistency. Efficient behaviour relies on estimating the probabilities of two potential error types: incorrectly disengaging from a valid expectation (false positive) and incorrectly maintaining an invalid expectation (false negative). These probabilities are determined by three parameters: the certainty of the expectation, depending on the amount of evidence reinforcing the expectation, the reliability of contradictory information, which indicates the quality of the error signal, and the cost asymmetry between false positives and false negatives. When the weighted probability of a false positive exceeds that of a false negative, individuals are motivated to maintain their expectations and employ defensive reduction strategies. Conversely, if the weighted probability of a false negative is higher, individuals are more likely to opt for a non-defensive reduction strategy, such as attitude change. Based on this model, we present in the following section a typology of regulation strategies that could guide future research into phenomena such as radicalization or adherence to conspiracy theories.

7 | CHARACTERISTICS OF REDUCTIONS STRATEGIES

Previous works have attempted to list the different reduction strategies of CDT (McGrath, 2017; see also Hardyck & Kardush, 1968). However, they did not include a clear conceptual framework to enable the classification of the reduction strategies according to a set of characteristics (Vaidis & Bran, 2018). Some integrative models that include CDT, such as the MMM, have proposed a typology that distinguishes between two types of regulation strategy (Proulx & Inzlicht, 2012). For the MMM, direct strategies focus on the regulation of inconsistency whereas palliative strategies do not respond directly to inconsistency but allow for the restoration of the sense of meaning. However, it lacks clear predictions regarding which strategy will occur (see Proulx et al., 2012). The most recent models explore the proposed regulation strategies and the *general process model of threat and defence* (GPMTD) (Jonas et al., 2014; Reiss et al., 2021) offers compelling proposals.

Nonetheless, very little has been done regarding direct regulation and this is why the current model focuses on them. The typology that is offered is thus different because here we develop a typology of direct regulation with the identification of specific characteristics and their consequences, whereas both MMM and GPMTD focus on indirect or palliative regulation. Based on our model, three specific characteristics can be used to categorize direct reduction strategies and their consequences: *action*, *temporality* and *immunity*.

7.1 | Action: Defensive versus non-defensive reductions

This first characteristic depends directly on the decision to disengage or maintain expectations according to the decision rule that we have introduced in this article. The *action* distinguishes defensive strategies from non-defensive strategies according to the effect on the expectation (i.e., maintenance vs disengagement).

Concerning defensive strategies, adhocness is a primary defensive reduction because it concerns saving expectations by adding information to explain the contradictory elements. Adhocness can take several forms and we can find it in particular in CDT studies on belief disconfirmation. One intriguing example of adhocness is the concept of transcendental explanation (Burris et al., 1997). In this study, participants were presented with an article detailing the tragic death of a child due to gunshot wounds, despite the father's fervent prayers for his survival. The inconsistency highlighted by this experiment raises a fundamental question within the realm of theism (as explored by Leibniz, 2010): How can misfortune exist in a world that is supposedly entirely determined by a good and merciful god? To address this inherent inconsistency, participants were introduced to transcendental arguments that emphasized the inscrutability of divine will (i.e., an ad hoc theory). By invoking a higher principle, this particular instance of adhocness aimed to reconcile the belief in a benevolent god with the presence of misfortune in the world.

Besides adhocness, several other defensive strategies exist. For instance, trivialization, as defined by Simon et al. (1995), corresponds to the reduction of the importance of contradictory information to maintain the expectation. Finally, the denial of responsibility in the case of carrying out a problematic behaviour (Gosling et al., 2006) or the derogation of the source (Aronson et al., 1963) can also help maintain the expectation.

On the other hand, non-defensive strategies correspond to strategies that allow individuals to disengage from the expectation. This disengagement can be expressed, for example, through attitude change (e.g., Festinger & Carlsmith, 1959), corresponding, in CDT studies, to reduced agreement with the expectation. Disengagement can also have substantial consequences with cognitive restructuring of a set of attitudes linked to the inconsistency (e.g., Leippe & Eisenstadt, 1994).

We argue that the choice between a defensive and a non-defensive strategy is related to the decision rule described above. Regarding the parameters of our model, we argue that individuals are more

likely to choose a non-defensive strategy (e.g., attitude change) when the expectation is less certain, contradictory information is more reliable and the consequences of incorrectly maintaining the expectation (regarding to incorrectly disengaging from the expectation) are greater. Conversely, individuals are more likely to choose a defensive strategy (e.g., adhocness) when the expectation is more certain, contradictory information is less reliable and the consequences of incorrectly disengaging from the expectation (with regard to incorrectly maintaining the expectation) are greater.

7.2 | Temporality: Curative versus prophylactic reductions

Temporality characterizes a strategy according to its moment of occurrence in relation to the error. For CDT, reduction strategies usually reflect curative responses to inconsistency—that is, responses that reduce contradictory information. Thus, they can only intervene once the inconsistency has arisen. Alternatively, *prophylactic strategies* aim to prevent the occurrence of inconsistency in the first place. These means minimize the likelihood of encountering an error by taking preventive measures.

The term *prophylactic* was chosen to place inconsistency management in the context of long-term motivation to reduce cognitive surprise (e.g., Clark, 2018; Friston, 2010). Interactions between an individual and the environment entail potential exposure to error signals, which justifies the use of strategies to prevent their appearance. We refer directly to the predictive coding, according to which any biological system is driven by cognitive entropy reduction—that is, the long-term minimization of cognitive surprise (Friston, 2009). When individuals interact with an environment that protects their belief system from error signals, they effectively minimize cognitive surprise in the long term by choosing predictable contexts. As introduced previously, these strategies can take the form of group isolation along with individuals sharing the same beliefs (Dein & Dawson, 2008; Festinger et al., 1956; Hardycy & Braden, 1962). From an individual perspective, this strategy can be expressed simply through the avoidance of contradictory information (e.g., D'Alessio & Allen, 2002).

A second prophylactic strategy consists of modulating the expectation's specificity, for example, by under-specifying its conditions of realization (Boudry & Braeckman, 2011; Gilovich, 1991, p. 30, p. 133). This strategy avoids inconsistency proactively by deliberately vaguely defining the conditions of validity of the expectation, making it possible to accommodate any new information.

Expectation certainty moderates the confirmation bias (Zhou & Shen, 2022), which can be interpreted as a prophylactic strategy, so we suggest that the more certain the expectation, the more prophylactic strategies individuals should use. Moreover, in line with our model, we also expect that the use of prophylactic strategies will ultimately increase the certainty of expectation and thus the likelihood of using a defensive strategy when confronted with contradictory information. Indeed, as prophylactic strategies lead to a reduction in the quan-

tity of contradictory information in the environment, this mechanically increases the proportion of information confirming the expectation.

7.3 | Immunity level

The notion of *immunity* comes from epistemology and in particular from the falsificationist approach (Boudry, 2013; Boudry & Braeckman, 2011; Popper, 2002). We define an immunizing strategy as a reduction that suppresses inconsistency in subsequent encounters with any information conveying the same meaning as the initial contradictory information. For example, by adding a causal explanation to contradictory information, adhocness is likely to remove the possibility for new information sharing the same semantic content to generate inconsistency.

Conversely, defensive reductions such as trivialization (Simon et al., 1995) or denial of responsibility (Gosling et al., 2006) should have a low level of immunity. For trivialization, reducing the weight of contradictory information does not imply a semantic modification of this piece of information and thus the same piece of information can generate inconsistency anew in subsequent confrontations. If participants writing an essay in favour of selective admission at their university can regulate by considering their behaviour unimportant, the written arguments remain in contradiction with their expectation and thus subsequent information in favour of selectivity is likely to generate inconsistency again.

An immunizing strategy removes the possibility for a piece of information sharing the same semantic content to generate inconsistency, so we suggest that it leads ultimately to an increase in expectation certainty. We suggest that if the individual resorts to an immunizing strategy, the amount of inconsistent information in the environment should decrease and the proportion of expectation-confirming information should increase, which should ultimately increase expectation certainty.

8 | CONSEQUENCES OF ADHOCNESS

This article claims that integrating the novel concept of adhocness into the CDT framework can offer valuable insights into the phenomenon of belief reinforcement when individuals are confronted with inconsistency (e.g., Festinger et al., 1956). Due to the characteristics of adhocness, it functions as a potent defensive strategy that effectively immunizes expectations and significantly contributes to their reinforcement.

As a first consequence, adhocness increases certainty regarding expectations through the addition of confirmatory information that corresponds to an alternative causal explanation. According to EIM, adding evidence increases the expectation certainty and thus increases the likelihood of using defensive strategies when confronted with subsequent contradictory information.

As a second consequence, this strategy could immunize the expectation to any future information conveying the same meaning. The

effectiveness of an ad hoc hypothesis in providing immunity depends on its ability to be tested empirically. Taking the example of Festinger et al. (1956) once again, the sect members generated various ad hoc hypotheses when confronted with the failure of their prophecy. Initially, they proposed that the clock was 4 min fast (1956, p. 162). However, this ad hoc hypothesis had low immunity as it was disproven just 4 min later. Employing additional ad hoc strategies, by the end of the night, the members interpreted the absence of doomsday as a reward for their unwavering devotion, believing that their dedication had saved the world (p. 169). This particular ad hoc strategy was unique in that it could not be tested empirically, making it particularly potent in preserving their expectations. Consequently, the members no longer needed to address the failed prophecy. Following this cognitive adjustment, they became more radicalized and engaged in proselytizing behaviours that were absent before.

In this regard, unfalsifiable ad hoc hypotheses play a pivotal role in immunizing expectations by converting contradictory information into confirmatory evidence of the ad hoc hypothesis. An intriguing illustration of this phenomenon involves individuals regulating an inconsistency by proposing the existence of a conspiracy, which serves as an unfalsifiable explanation to uphold their initial expectation. This specific issue presents a compelling avenue for future empirical studies within the realm of CDT research (see also, Cooper & Pearce, 2023).

9 | MOTIVATIONAL AND EMOTIONAL ASPECTS OF EIM

Although our model mainly focuses on the cognitive side of inconsistency management, it is nevertheless necessary to address the motivational and emotional aspects as well. An aversive state of arousal is expected in the case of inconsistencies (Croyle & Cooper, 1983; Devine et al., 1999; Martinie et al., 2013; Proulx et al., 2012) and prediction errors (Alamia et al., 2019; Braem et al., 2015; Kastner et al., 2017). In this respect, the action-based model (ABM) focuses particularly on the cerebral and physiological characteristic of inconsistency (Harmon-Jones et al., 2015), showing important overlaps with the literature on prediction error at the cerebral level (e.g., Alexander & Brown, 2015; Kennerley et al., 2006) and at the physiological level (Alamia et al., 2019; Braem et al., 2015; Kastner et al., 2017).

From an evolutionary standpoint (Tinbergen, 1963), arousal must be comprehended as a proximal mechanism that signals errors and motivates individuals to allocate attentional resources for resolving inconsistencies (also see Gawronski & Brannon, 2019). Inconsistencies (i.e., prediction errors) represent adaptive challenges (Bar, 2009). Similarly, ABM suggests that inconsistency represents an adaptive challenge for individuals by affecting their capacity for action (Harmon-Jones et al., 2015). Related to this consideration, the greater the adaptive challenge that inconsistency generates for individuals, the stronger the arousal state should be and thus the motivation to regulate inconsistency. In our model, the importance of this adaptive challenge could be determined according to the EIM parameter: the

certainty of the expectation, the reliability of contradictory information and the anticipated consequences of the error in the regulation decision. Consequently, the level of arousal should be adjusted by these parameters.

An expectation with a high level of certainty (i.e., reinforced) confronted with very reliable contradictory information (i.e., a high precision error signal) should generate a significant adaptive challenge for the individual and thus generate strong arousal. In other words, the level of arousal should be maximal when the certainty of expectation and the reliability of contradictory information are strong. Conversely, an inconsistency for which the certainty of expectation is strong and contradictory information is unreliable should elicit a lower level of arousal. In this regard, PP models indicate that individuals should make a second-order prediction about the reliability of the error signal in order to weigh it so as not to update expectations based on unreliable signals (Clark, 2018; Feldman & Friston, 2010). Following this logic, an unreliable error signal should result in less motivation to regulate and thus a weaker arousal state. Referring to Hume's example (2007, p. 80), a mentally healthy person directly confronted with a *resurrection* would experience a conflict between a strongly certain expectation ('humans don't resurrect') and reliable contradictory information. This creates an adaptive challenge that generates high arousal and motivation to maintain the expectation while seeking information to explain the specific case. In contrast, receiving testimony about a resurrection from someone with a psychiatric disorder would produce low arousal due to the high certainty of the expectation and the low reliability of contradictory information from an unreliable source.

Finally, arousal and motivation to regulate should also be a function of the importance of the consequences of errors in regulation. This is explained at a more fundamental level by the risks for individuals' adaptation in cases where they make an error in their decision to maintain their expectation or to abandon it in the face of contradictory information. Arousal motivates individuals to invest attentional resources (Gawronski & Brannon, 2019) so we should thus expect the level of arousal to vary with the importance of the anticipated consequences of an error in decision-making.

Regarding the emotions evoked by inconsistency, surprise plays a significant role, as it corresponds to the negative subjective feeling resulting from the violation of an expectation (Noordewier et al., 2016; Reizenzen et al., 2019). The role of surprise has also been noted in the literature on cognitive dissonance and prediction error, showing once again the value of this overlap (e.g., Friston & Stephan, 2007; Levy et al., 2018).

The level of surprise is related to the discrepancy between the expectation and contradictory information (Meyer, 1988; Reizenzen et al., 2019). In the logic of our model, we should therefore expect maximum surprise when the expectation is highly certain and the contradictory information highly reliable. Moreover, surprise serves adaptive functions, such as generating curiosity (Berlyne, 1960) and motivating the search for information about the causes of error (Gendolla & Koller, 2001; Stiensmeier-Pelster et al., 1995). In our model, inconsistency containing a highly certain expectation and reliable contradictory information can lead to strong surprise, which motivates the

search for ad hoc explanations while preserving the expectation (in the case of a defensive strategy) or a cognitive restructuring (in the case of a non-defensive strategy). With regard to arousal and surprise, our model offers new predictions, which are summarized in Supporting Information 2.

10 | FINAL CONSIDERATIONS AND FUTURE DIRECTIONS

Based on a conceptual overlap between inconsistency and prediction error, we have suggested that individuals make a bet when faced with contradictory information. We propose that individuals must arbitrate between two possible types of errors: the disengagement from a correct expectation or the maintenance of an incorrect expectation. From an evaluation of the most probable and most costly errors, individuals should bias their decisions by adopting either defensive strategies aimed at maintaining the expectation, or non-defensive strategies aimed at abandoning the expectation. In order to evaluate the probability of the two errors and their cost, we suggested that individuals evaluate three parameters: certainty, reliability and cost asymmetry. We have also derived a typology of regulation strategies from this decision rule.

In addition to accommodating a large number of empirical results, our current model introduces novel directional predictions, particularly about the expected direct regulations, which were absent in the literature (Proulx et al., 2012; Vaidis & Bran, 2018, 2019). To put the EIM to the test, the described three parameters should now be used to predict the effective choice of strategy in the face of contradictory information. A main assumption is that a defensive strategy should be favoured when certainty is high, reliability is low, and the consequences of incorrect disengagement are greater. The assessment of the probability of a false positive and a false negative should mediate the links between the informational parameters and choice regulation. In addition to conducting correlational analyses, it will be essential to manipulate the information parameters in order to challenge the model. For instance, one could modulate certainty by adding confirmatory information prior to inconsistency, or one could manipulate reliability by adding noise in a contradictory message, or highlight negative consequences of an incorrect disengagement or incorrect maintenance of the expectation to manipulate the costs of errors. These are just a few of the potential alternatives to explore but testing the regulation direction depending on the described informational parameters is critical and contradictory data would need to be considered as a falsification of EIM. Finally, with an effective regulation prediction model in place, a multitude of prominent applications could be anticipated, particularly in the realm of public policy communication. These applications would aim to generate effective communication strategies that mitigate the occurrence of backfire effect.

We also believe that our model can serve as inspiration for research on attitude change. Indeed, beside the development of defensive strategies, the EIM offers new predictions about attitude change in

situations of inconsistency. Our model predicts that a highly certain expectation increases the probability of employing a defensive regulation strategy, potentially reinforcing the belief and associated beliefs. Conversely, changing a low-certainty expectation may increase the probability of changing a more certain expectation associated with the same object. To investigate this, we encourage a novel paradigm involving repeated inconsistency that targets multiple beliefs about a specific object. By exposing participants to inconsistent information regarding a belief with a low certainty level, regulating it and then presenting inconsistent information regarding a strongly certain belief, EIM anticipates a greater attitude change in comparison with the reversed order (i.e., inconsistent information against a highly certain belief, then inconsistent information against a weakly certain belief). The type of regulation employed during the first inconsistency (defensive vs non-defensive) is also expected to mediate this attitude change.

Finally, whereas CDT literature has focused almost exclusively on attitude change, EIM offers predictions about resistance to change and belief reinforcement after inconsistency. In the typology of regulations, we have described prophylactic regulation strategies that make it possible to reduce, a priori, the quantity of errors in the environment. Prophylactic strategies, such as group isolation or selective exposure, should be particularly used for central beliefs (highly certain and important). Moreover, as prophylactic strategies reduce errors, we can anticipate an a posteriori increase in the certainty of the expectation, attributed to the greater amount of confirmatory information. Finally, because the use of prophylactic regulation enhances certainty by increasing the quantity of confirmatory information, this should increase the likelihood of employing a defensive strategy in the face of inconsistency. To test these predictions, a manipulation could present texts containing confirmatory and contradictory information concerning highly certain beliefs or weakly certain beliefs before measuring certainty once again. In this case, EIM predicts a greater recall of confirmatory information for highly (vs weakly) certain beliefs. This greater recall should also predict increased certainty. In a third phase, this increase in certainty should generate greater likelihood of a defensive strategy facing contradictory information.

We have also described belief immunization as a possible consequence of the use of certain regulation strategies, in particular adhocness. This also leads to new predictions. Ad hoc hypotheses, such as conspiracy theories, could generate a 'protective belt' (Lakatos, 1976, p.126), transforming subsequent contradictory information into confirmatory evidence (i.e., immunization), thus preventing errors at the level of the expectation. In this respect, EIM predicts, for example, that recruiting a conspiracy theory to regulate inconsistency generates an immunization of the expectation. Consequently, any further contradictory information would be interpreted as confirming the ad hoc hypothesis (i.e., the conspiracy theory), avoiding errors in the expectation. In other words, after an immunizing regulation, we should expect that subsequent contradictory information would no longer generate inconsistency. In this sense, because an inconsistency or error in prediction mobilizes working memory resources (Martinie et al., 2010), one expects that after regulation via an unfalsifiable ad hoc hypothesis,

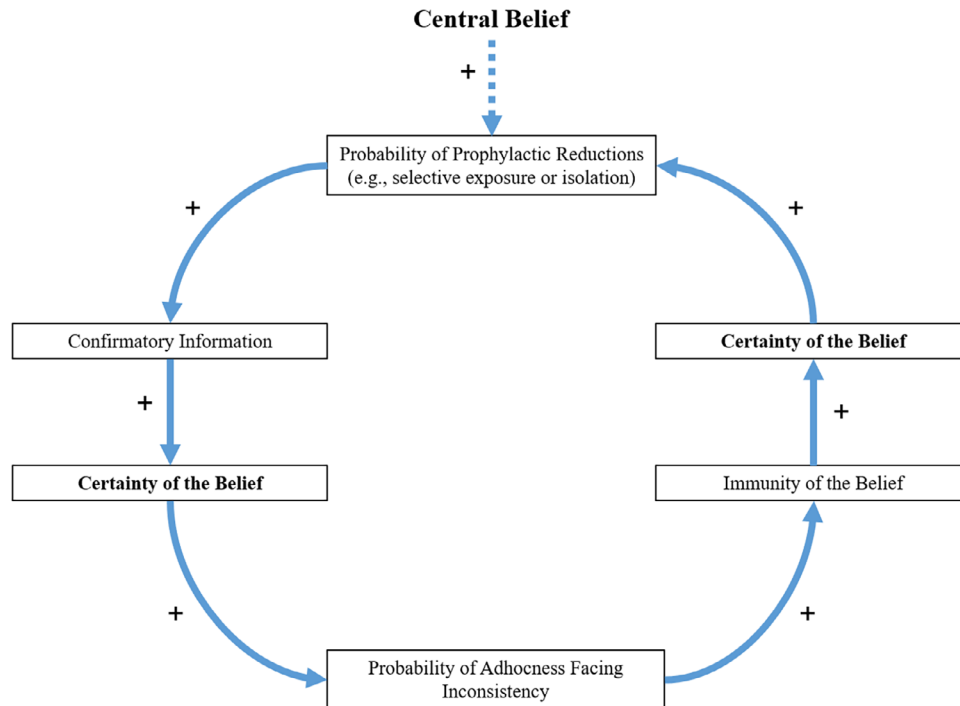


FIGURE 3 Vicious circle of radicalization.

new contradictory information would produce less working memory load.

More generally, we consider that our model and the predictions associated with resistance to change could help to enrich research on radicalization, by offering new avenues for empirical work on important societal phenomena for CDT. The process we described in this article and the various concepts we developed can be integrated to describe a vicious circle of radicalization (Figure 3).

In this dynamic, central beliefs increase the probability of prophylactic regulation (e.g., group isolation) which in turn reduces the errors in the environment and increases the relative proportion of confirmatory information. Following this logic, increasing confirmatory information raises the certainty of expectation and thus increases the likelihood of a defensive regulation strategy when exposed to contradictory information. In the case of an immunizing strategy (e.g., conspiracy theory), it also contributes to the creation of a 'protective belt' likely to transform following contradictory information into confirmatory information. Altogether, this process increases the certainty of the expectation and therefore raises the likelihood of prophylactic strategies (e.g., social isolation), which perpetuates the cycle of radicalization. In sum, the EIM decision rule and the derived typology, by offering new predictions concerning defensive strategies, could help to guide future research into CDT, particularly on important societal phenomena such as radicalization or conspiracy.

CONFLICT OF INTEREST STATEMENT

We confirm that this manuscript is not subject to any conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed.

ETHICS STATEMENT

This theoretical article was conducted in accordance with the Ethical Guidelines of the American Psychological Association (APA).

TRANSPARENCY STATEMENT

This theoretical article aims to provide a comprehensive analysis and conceptual framework without generating empirical data. The content is based on a thorough review of existing literature and theoretical reasoning. While no data collection or statistical analysis was performed, we have strived to ensure clarity and rigor in presenting our theoretical perspectives and arguments.

ORCID

Benjamin Aubert-Teillaud  <https://orcid.org/0000-0003-4470-1856>

Alexandre Bran  <https://orcid.org/0000-0003-2838-3886>

David C. Vaidis  <https://orcid.org/0000-0002-1954-2219>

REFERENCES

- Adams, J.S. (1961). Reduction of cognitive dissonance by seeking consonant information. *The Journal of Abnormal and Social Psychology*, 62(1), 74–78. <https://doi.org/10.1037/h0047029>
- Alamia, A., Van Rullen, R., Pasqualotto, E., Mouraux, A., & Zenon, A. (2019). Pupil-linked arousal responds to unconscious surprisal. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 39(27), 5369–5376. <https://doi.org/10.1523/JNEUROSCI.3010-18.2019>

- Alexander, W. H., & Brown, J. W. (2015). Hierarchical error representation: A computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, 27(11), 2354–2410. <https://doi.org/10.1162/NECOa00779>
- Aronson, E., & Carlsmith, J. M. (1963). Effect of the severity of threat on the devaluation of forbidden behavior. *The Journal of Abnormal and Social Psychology*, 66(6), 584–588. <https://doi.org/10.1037/h0039901>
- Aronson, E., Fried, C., & Stone, J. (1991). Overcoming denial and increasing the intention to use condoms through the induction of hypocrisy. *American Journal of Public Health*, 81(12), 1636–1638. <https://doi.org/10.2105/AJPH.81.12.1636>
- Aronson, E., Turner, J. A., & Carlsmith, J. M. (1963). Communicator credibility and communication discrepancy as determinants of opinion change. *The Journal of Abnormal and Social Psychology*, 67(1), 31–36. <https://doi.org/10.1037/h0045513>
- Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. Introduction. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521), 1181–1182. <https://doi.org/10.1098/rstb.2008.0321>
- Batson, C. D. (1975). Rational processing or rationalization? The effect of disconfirming information on a stated religious belief. *Journal of Personality and Social Psychology*, 32(1), 176–184. <https://doi.org/10.1037/h076771>
- Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221. <https://doi.org/10.1038/nn1954>
- Bergin, A. E. (1962). The effect of dissonant persuasive communications upon changes in a self-referring attitude. *Journal of Personality*, 30(3), 423–438. <https://doi.org/10.1111/j.1467-6494.1962.tb02314.x>
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill Book Company. <https://doi.org/10.1037/11164-000>
- Bland, A. R., & Schaefer, A. (2012). Different varieties of uncertainty in human decision-making. *Frontiers in Neuroscience*, 6, 85. <https://doi.org/10.3389/fnins.2012.00085>
- Bochner, S., & Insko, C. A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology*, 4(6), 614–621. <https://doi.org/10.1037/h0021192>
- Boudry, M. (2013). The hypothesis that saves the day. Ad hoc reasoning in pseudoscience. *Logique et Analyse*, 56, 245–258.
- Boudry, M., & Braeckman, J. (2011). Immunizing strategies and epistemic defense mechanisms. *Philosophia*, 39(1), 145–161. <https://doi.org/10.1007/s11406-010-9254-9>
- Braem, S., Coenen, E., Bombeke, K., van Bochove, M. E., & Notebaert, W. (2015). Open your eyes for prediction errors. *Cognitive, Affective & Behavioral Neuroscience*, 15(2), 374–380. <https://doi.org/10.3758/s13415-014-0333-4>
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), 384–389. <https://doi.org/10.1037/h0041006>
- Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. John Wiley & Sons Inc. <https://doi.org/10.1037/11622-000>
- Burris, C. T., Harmon-Jones, E., & Tarpley, W. R. (1997). "By faith alone": Religious agitation and cognitive dissonance. *Basic and Applied Social Psychology*, 19(1), 17–31. https://doi.org/10.1207/s15324834basps1901_2
- Chaiken, S., Pomerantz, E. M., & Giner-Sorolla, R. (1995). Structural consistency and attitude strength. In R. E. Petty, & J. A. Krosnick (Eds.), *Ohio State University series on attitudes and persuasion, Vol. 4. Attitude strength: Antecedents and consequences* (pp. 387–412). Lawrence Erlbaum Associates, Inc.
- Chalmers, A. F. (2013). *What is this thing called science?* Hackett Publishing. (Original work published 1976).
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, 99(4), 573–594. <https://doi.org/10.1037/a002021>
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrocioni, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Clark, A. (2016). *Surfing uncertainty*. Oxford University Press.
- Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, 17(3), 521–534. <https://doi.org/10.1007/s11097-017-9525-z>
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1481), 933–942. <https://doi.org/10.1098/rstb.2007.2098>
- Cooper, J., & Mackie, D. (1983). Cognitive dissonance in an intergroup context. *Journal of Personality and Social Psychology*, 44(3), 536–544. <https://doi.org/10.1037/0022-3514.44.3.536>
- Cooper, J., & Pearce, L. (2023). The arc of dissonance. In J. P. Forgas, W. D. Crano, & K. Fiedler (Eds.), *The psychology of insecurity* (pp. 95–107). Routledge.
- Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R., Shanks, D. R., Robbins, T. W., Bullmore, E. T., Dickinson, A., & Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain: A Journal of Neurology*, 130(9), 2387–2400. <https://doi.org/10.1093/brain/awm173>
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10(7), 294–300. <https://doi.org/10.1016/j.tics.2006.05.004>
- Croyle, R. T., & Cooper, J. (1983). Dissonance arousal: Physiological evidence. *Journal of Personality and Social Psychology*, 45(4), 782–791. <https://doi.org/10.1037/0022-3514.45.4.782>
- D'Alessio, D., & Allen, M. (2002). Selective exposure and dissonance after decisions. *Psychological Reports*, 91(2), 527–532. <https://doi.org/10.2466/pr0.2002.91.2.527>
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876–879. <https://doi.org/10.1038/nature04766>
- Dein, S., & Dawson, L. L. (2008). The 'Scandal' of the Lubavitch Rebbe: Messianism as a response to failed prophecy. *Journal of Contemporary Religion*, 23(2), 163–180. <https://doi.org/10.1080/13537900802024550>
- den Ouden, H. E., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, 3, 548. <https://doi.org/10.3389/fpsyg.2012.00548>
- Devine, P. G., & Brodish, A. B. (2003). Modern classics in social psychology. *Psychological Inquiry*, 14(3-4), 196–202. https://doi.org/10.1207/S15327965PL1403&4_3
- Devine, P. G., Tauer, J. M., Barron, K. E., Elliot, A. J., & Vance, K. M. (1999). Moving beyond attitude change in the study of dissonance-related processes. In E. Harmon-Jones, & J. Mills (Eds.), *Science conference series. Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 297–323). American Psychological Association.
- Dickerson, C. A., Thibodeau, R., Aronson, E., & Miller, D. (1992). Using cognitive dissonance to encourage water conservation. *Journal of Applied Social Psychology*, 22(11), 841–854. <https://doi.org/10.1111/j.1559-1816.1992.tb00928.x>
- Diederen, K. M., Spencer, T., Vestergaard, M. D., Fletcher, P. C., & Schultz, W. (2016). Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. *Neuron*, 90(5), 1127–1138. <https://doi.org/10.1016/j.neuron.2016.04.019>
- Doosje, B., Moghaddam, F. M., Kruglanski, A. W., De Wolf, A., Mann, L., & Feddes, A. R. (2016). Terrorism, radicalisation and de-radicalisation.

- Current Opinion in Psychology*, 11, 79–84. <https://doi.org/10.1016/j.copsyc.2016.06.008>
- Eagly, A. H. (1969). Responses to attitude-discrepant information as a function of intolerance of inconsistency and category width. *Journal of Personality*, 37(4), 601–617. <https://doi.org/10.1111/j.1467-6494.1969.tb01768.x>
- Eaton, A. A., & Visser, P. S. (2008). Attitude importance: Understanding the causes and consequences of passionately held views. *Social and Personality Psychology Compass*, 2(4), 1719–1736. <https://doi.org/10.1111/j.1751-9004.2008.00125.x>
- Eisenstadt, D., & Leippe, M. R. (2005). Dissonance and importance: Attitude change effects of personal relevance and race of the beneficiary of a counterattitudinal advocacy. *The Journal of Social Psychology*, 145(4), 447–467. <https://doi.org/10.3200/SOCP.145.4.447-468>
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 215. <https://doi.org/10.3389/fnhum.2010.00215>
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2), 203–210. <https://doi.org/10.1037/h0041593>
- Festinger, L., Riecken, H., & Schachter, S. (1956). *When prophecy fails*. University of Minnesota Press.
- Feyerabend, P. (1993). *Against method*. Verso. (Original work published 1975).
- Fointiat, V., Priolo, D., Saint-Bauzel, R., & Milhabet, I. (2013). Justifier nos transgressions pour réduire notre hypocrisie? Hypocrisie induite et identification des transgressions [Justifying our own counter-normative behaviors as a route for hypocrisy reduction? Dissonance and transgressions identification]. *Revue Internationale de Psychologie Sociale*, 26(4), 49–78.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Friston, K. (2018). Active inference and cognitive consistency. *Psychological Inquiry*, 29(2), 67–73. <https://doi.org/10.1080/1047840X.2018.1480693>
- Friston, K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, 3, 130. <https://doi.org/10.3389/fpsyg.2012.00130>
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417–458. <https://doi.org/10.1007/s11229-007-9237-y>
- Gawronski, B., & Brannon, S. M. (2019). What is cognitive consistency, and why does it matter? In E. Harmon-Jones (Ed.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (pp. 91–116). American Psychological Association. <https://doi.org/10.1037/0000135-005>
- Gawronski, B., & Strack, F. (2012). *Cognitive consistency: A fundamental principle in social cognition*. Guilford Press.
- Gendolla, G. H. E., & Koller, M. (2001). Surprise and motivation of causal search: How are they affected by outcome valence and importance? *Motivation and Emotion*, 25(4), 327–349. <https://doi.org/10.1023/A:1014867700547>
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. Free Press.
- Gosling, P., Denizeau, M., & Oberlé, D. (2006). Denial of responsibility: A new mode of dissonance reduction. *Journal of Personality and Social Psychology*, 90(5), 722–733. <https://doi.org/10.1037/0022-3514.90.5.722>
- Haarsma, J., Fletcher, P. C., Griffin, J. D., Taverne, H. J., Ziauddeen, H., Spencer, T. J., Miller, C., Katthagen, T., Goodyer, I., Diederens, K., & Murray, G. K. (2020). Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Molecular Psychiatry*, 26(9), 5320–5333. Advance online publication. <https://doi.org/10.1038/s41380-020-0803-8>
- Haggbloom, S. J., Warnick, R., Warnick, J. E., Jones, V. K., Yarbrough, G. L., Russell, T. M., Borecky, C. M., McGahhey, R., Powell, J. L., Beavers, J., & Monte, E. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology*, 6(2), 139–152. <https://doi.org/10.1037/1089-2680.6.2.139>
- Hardyck, J. A., & Braden, M. (1962). Prophecy fails again: A report of a failure to replicate. *The Journal of Abnormal and Social Psychology*, 65(2), 136–141. <https://doi.org/10.1037/h0048636>
- Hardyck, J. A., & Kardush, M. (1968). A modest modish model for dissonance reduction. In R. P. Abelson, E. Aronson, W. T. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 684–692). Rand McNally.
- Harmon-Jones, E. (2019). *Cognitive dissonance: Reexamining a pivotal theory in psychology*. American Psychological Association.
- Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology*, 70(1), 5–16. <https://doi.org/10.1037/0022-3514.70.1.5>
- Harmon-Jones, E., & Harmon-Jones, C. (2019). Understanding the motivation underlying dissonance effects: The action-based model. In E. Harmon-Jones (Ed.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (pp. 63–89). American Psychological Association. <https://doi.org/10.1037/0000135-004>
- Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Current Directions in Psychological Science*, 24(3), 184–189. <https://doi.org/10.1177/0963721414566449>
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91. <https://doi.org/10.1037/0022-3514.78.1.81>
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, 10(1), 47–66. https://doi.org/10.1207/s15327957pspr1001_3
- Haselton, M. G., Nettle, D., & Murray, D. R. (2016). The evolution of cognitive bias. In D. M. Buss (Ed.), *The handbook of evolutionary psychology: Integrations* (pp. 968–987). John Wiley & Sons, Inc.
- Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review: An Official Journal of the SOCIETY for Personality and Social Psychology, Inc.*, 10(2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1
- Howe, L. C., & Krosnick, J. A. (2017). Attitude strength. *Annual Review of Psychology*, 68(1), 327–351. <https://doi.org/10.1146/annurev-psych-122414-033600>
- Hume, D. (2007). *Enquiry concerning human understanding*. Cambridge University Press. (original work published in 1748).
- Izuma, K., & Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: A critical methodological review. *Frontiers in Psychology*, 4, 41. <https://doi.org/10.3389/fpsyg.2013.00041>
- Jacks, J. Z., & Devine, P. G. (2000). Attitude importance, forwarning of message content, and resistance to persuasion. *Basic and Applied Social Psychology*, 22(1), 19–29. https://doi.org/10.1207/S15324834BASP2201_3
- Johnson, D. D., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends In Ecology & Evolution*, 28(8), 474–481. <https://doi.org/10.1016/j.tree.2013.05.014>

- Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., Nash, K., Proulx, T., & Quirin, M. (2014). Threat and defense: From anxiety to approach. In J. M. Olson, & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 49, pp. 219–286). Elsevier Academic Press.
- Joule, R. V., & Martinie, M.-A. (2008). Forced compliance, misattribution and trivialization. *Social Behavior and Personality: An International Journal*, 36(9), 1205–1212. <https://doi.org/10.2224/sbp.2008.36.9.1205>
- Kaaronen, R. O. (2018). A theory of predictive dissonance: Predictive processing presents a new take on cognitive dissonance. *Frontiers in Psychology*, 9, 2218. <https://doi.org/10.3389/fpsyg.2018.02218>
- Kastner, L., Kube, J., Villringer, A., & Neumann, J. (2017). Cardiac concomitants of feedback and prediction error processing in reinforcement learning. *Frontiers in Neuroscience*, 11, 598. <https://doi.org/10.3389/fnins.2017.00598>
- Kennerly, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7), 940–947. <https://doi.org/10.1038/nn1724>
- Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, 84(6), 1329–1342. <https://doi.org/10.1016/j.neuron.2014.12.015>
- Krosnick, J. A. (1988). Attitude importance and attitude change. *Journal of Experimental Social Psychology*, 24(3), 240–255. [https://doi.org/10.1016/0022-1031\(88\)90038-8](https://doi.org/10.1016/0022-1031(88)90038-8)
- Kruglanski, A. W., Jasko, K., Milyavsky, M., Chernikova, M., Webber, D., Pierro, A., & di Santo, D. (2018). Cognitive consistency theory in social psychology: A paradigm reconsidered. *Psychological Inquiry*, 29(2), 45–59. <https://doi.org/10.1080/1047840X.2018.1480619>
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. University of Chicago Press. (Original work published 1962).
- Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. In *Proceedings of the Aristotelian society* (Vol. 69, pp. 149–186). Aristotelian Society, Wiley.
- Lakatos, I., & Musgrave, A. (Eds.). (1970). *Criticism and the growth of knowledge: Proceedings of the International Colloquium in the Philosophy of Science*, London, 1965. Cambridge University Press. <https://doi.org/10.1017/CBO9781139171434>
- Lakatos, I. (1976). Falsification and the methodology of scientific research programmes. In *Can theories be refuted?* Springer.
- Leibniz, G. W. (2010). *Theodicy*. Cosimo, Inc.
- Leippe, M. R., & Eisenstadt, D. (1994). Generalization of dissonance reduction: Decreasing prejudice through induced compliance. *Journal of Personality and Social Psychology*, 67(3), 395–413. <https://doi.org/10.1037/0022-3514.67.3.395>
- Levy, N., Harmon-Jones, C., & Harmon-Jones, E. (2018). Dissonance and discomfort: Does a simple cognitive inconsistency evoke a negative affective state? *Motivation Science*, 4(2), 95–108. <https://doi.org/10.1037/mot0000079>
- Martinie, M.-A., Milland, L., & Olive, T. (2013). Some theoretical considerations on attitude, arousal and affect during cognitive dissonance. *Social and Personality Psychology Compass*, 7(9), 680–688. <https://doi.org/10.1111/spc3.12051>
- Martinie, M.-A., Olive, T., & Milland, L. (2010). Cognitive dissonance induced by writing a counterattitudinal essay facilitates performance on simple tasks but not on complex tasks that involve working memory. *Journal of Experimental Social Psychology*, 46(4), 587–594. <https://doi.org/10.1016/j.jesp.2009.10.018>
- Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers In Human Neuroscience*, 5, 39. <https://doi.org/10.3389/fnhum.2011.00039>
- McCauley, C., & Moskaleiko, S. (2008). Mechanisms of political radicalisation: Pathways toward terrorism. *Terrorism and Political Violence*, 20(3), 415–433. <https://doi.org/10.1080/09546550802073367>
- McGrath, A. (2017). Dealing with dissonance: A review of cognitive dissonance reduction. *Social and Personality Psychology Compass*, 11(12), 1–17. <https://doi.org/10.1111/spc3.12362>
- Meyer, W.-U. (1988). Die Rolle von Überraschung im Attributionsprozeß [The role of surprise in the attribution process]. *Psychologische Rundschau*, 39, 136–147.
- Moghaddam, F. M. (2005). The staircase to terrorism: A psychological exploration. *American Psychologist*, 60(2), 161–169. <https://doi.org/10.1037/0003-066X.60.2.161>
- Nassar, M. R., Wilson, R. C., Heasley, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(37), 12366–12378. <https://doi.org/10.1523/JNEUROSCI.0822-10.2010>
- Noordewier, M. K., Topolinski, S., & Van Dijk, E. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass*, 10(3), 136–149. <https://doi.org/10.1111/spc3.12242>
- Piaget, J. (1954). *The construction of reality in the child*. Basic Books. (Original work published 1937).
- Popper, K. (2005). *The logic of scientific discovery*. Routledge. (Original work published 1934).
- Popper, K. R. (2002). *Conjectures and refutations: The growth of scientific knowledge*. Routledge. (Original work published 1963).
- Proulx, T., & Inzlicht, M. (2012). The five “A”s of meaning maintenance: Finding meaning in the theories of sense-making. *Psychological Inquiry*, 23(4), 317–335. <https://doi.org/10.1080/1047840X.2012.702372>
- Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences*, 16(5), 285–291. <https://doi.org/10.1016/j.tics.2012.04.002>
- Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, 11(1), 50–74. <https://doi.org/10.1111/tops.12292>
- Reiss, S., Leen-Thomele, E., Klackl, J., & Jonas, E. (2021). Exploring the landscape of psychological threat: A cartography of threats and threat responses. *Social and Personality Psychology Compass*, 15(4), 1–17. <https://doi.org/10.1111/spc3.12588>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (Vol. 2, pp. 64–99). Appleton Century Crofts.
- Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology*, 16(4), 388–403. [https://doi.org/10.1016/0022-1031\(80\)90030-X](https://doi.org/10.1016/0022-1031(80)90030-X)
- Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, 68(2), 247–260. <https://doi.org/10.1037/0022-3514.68.2.247>
- Starzyk, K. B., Fabrigar, L. R., Soryal, A. S., & Fanning, J. J. (2009). A painful reminder: The role of level and salience of attitude importance in cognitive dissonance. *Personality & Social Psychology Bulletin*, 35(1), 126–137. <https://doi.org/10.1177/0146167208325613>
- Stiensmeier-Pelster, J., Martini, A., & Reisenzein, R. (1995). The role of surprise in the attribution process. *Cognition and Emotion*, 9(1), 5–31. <https://doi.org/10.1080/02699939508408963>
- Stone, J., Aronson, E., Crain, A. L., Winslow, M. P., & Fried, C. B. (1994). Inducing hypocrisy as a means of encouraging young adults to use condoms. *Personality and Social Psychology Bulletin*, 20(1), 116–128. <https://doi.org/10.1177/0146167294201012>
- Stone, J., & Fernandez, N. C. (2008). To practice what we preach: The use of hypocrisy and cognitive dissonance to motivate behavior change. *Social and Personality Psychology Compass*, 2(2), 1024–1051. <https://doi.org/10.1111/j.1751-9004.2008.00088.x>
- Sun, Z., & Firestone, C. (2020). The dark room problem. *Trends in Cognitive Sciences*, 24(5), 346–348. <https://doi.org/10.1016/j.tics.2020.02.006>
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. MIT Press.

- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, 20(4), 410–433. <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>
- Vaghi, M. M., Luyckx, F., Sule, A., Fineberg, N. A., Robbins, T. W., & De Martino, B. (2017). Compulsivity reveals a novel dissociation between action and confidence. *Neuron*, 96(2), 348–354.e4. <https://doi.org/10.1016/j.neuron.2017.09.006>
- Vaidis, D. C., & Bran, A. (2018). Some prior considerations about dissonance to understand its reduction: Comment on McGrath (2017). *Social and Personality Psychology Compass*, 12(9), e12411. <https://doi.org/10.1111/spc3.12411>
- Vaidis, D. C., & Bran, A. (2019). Respectable challenges to respectable theory: Cognitive dissonance theory requires conceptualization clarification and operational tools. *Frontiers in Psychology*, 10, 1189. <https://doi.org/10.3389/fpsyg.2019.01189>
- Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, 1464(1), 242–268. <https://doi.org/10.1111/nyas.14321>
- Weiser, N. (1974). The effects of prophetic disconfirmation of the committed. *Review of Religious Research*, 16(1), 19–30. <https://doi.org/10.2307/3510194>
- Yousaf, O., & Gobet, F. (2013). The emotional and attitudinal consequences of religious hypocrisy: Experimental evidence using a cognitive dissonance paradigm. *The Journal of Social Psychology*, 153(6), 667–686. <https://doi.org/10.1080/00224545.2013.814620>
- Zanna, M. P., & Cooper, J. (1974). Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, 29(5), 703–709. <https://doi.org/10.1037/h0036651>
- Zhou, Y., & Shen, L. (2022). Confirmation bias and the persistence of misinformation on climate change. *Communication Research*, 49(4), 500–523. <https://doi.org/10.1177/00936502211028049>
- Zuwerink, J. R., & Devine, P. G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology*, 70(5), 931–944. <https://doi.org/10.1037/0022-3514.70.5.931>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aubert-Teillaud, B., Bran, A., & Vaidis, D. C. (2023). Expectation violation and cognitive dissonance theory: Proposal for an Epistemic Inconsistency Management Model. *European Journal of Social Psychology*, 1–17. <https://doi.org/10.1002/ejsp.2981>