Journal of Experimental Social Psychology 63 (2016) 1–9

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

# Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation☆

Pieter Van Dessel [a,*], Jan De Houwer [a], Anne Gast [b], Colin Tucker Smith [c], Maarten De Schryver [a]

[a] Ghent University, Belgium
[b] University of Cologne, Germany
[c] University of Florida, USA

## HIGHLIGHTS

- Instructions to approach or avoid fictitious social groups cause changes in implicit evaluations.
- These changes are not fully mediated by changes in explicit evaluations.
- These changes occur even in the absence of changes in explicit evaluations.
- These changes occur even if highly diagnostic incompatible valence information is provided.
- These findings constrain current and future models of implicit evaluation.

## ARTICLE INFO

*Article history:*
Received 1 May 2015
Revised 3 November 2015
Accepted 16 November 2015
Available online 1 December 2015

*Keywords:*
Approach
Avoidance
Training
Instructions
Implicit attitudes
Evaluation

## ABSTRACT

Previous research has shown that linking approach or avoidance actions to novel stimuli through mere instructions causes changes in the implicit evaluation of these stimuli even when the actions are never performed. In two high-powered experiments (total $N = 1147$), we examined whether effects of approach–avoidance instructions on implicit evaluations are mediated by changes in explicit evaluations. Participants first received information about the evaluative properties of two fictitious social groups (e.g., Niffites are good; Luupites are bad) and then received instructions to approach one group and avoid the other group. We observed an effect of approach–avoidance instructions on implicit but not explicit evaluations of the groups, even when these instructions were incompatible with the previously obtained evaluative information. These results indicate that approach–avoidance instructions allow for unintentional changes in implicit evaluations. We discuss implications for current theories of implicit evaluation.



## 1. Introduction

The way in which humans evaluate stimuli as good or bad has long been a central research topic in various sub-disciplines of psychology (Allport, 1935). In contemporary research on evaluations, researchers often contrast deliberate, explicit evaluations and spontaneous, implicit evaluations (see De Houwer, 2009a; Gawronski & Bodenhausen, 2011). Typically, theorists have postulated distinct underlying processes, with explicit evaluations resulting from belief-based processes that involve the validation of propositional information, and implicit evaluations being the product of processes involving the automatic activation of associations in memory (Gawronski & Bodenhausen, 2011).

Given the unique relation between implicit evaluations and behavior (Greenwald, Poehlman, Uhlmann, & Banaji, 2009), it is vital to understand how implicit stimulus evaluations are acquired and can be changed. Because implicit evaluation is traditionally attributed to the activation of associations between representations in memory and because associations are typically thought to develop gradually over many experiences, it is sometimes assumed that implicit evaluations of stimuli arise exclusively as the result of repeated experiences, such as recurrent pairings of physical stimuli (Rydell & McConnell, 2006). Evaluative conditioning (EC) research provides ample evidence that changes in the implicit evaluation of a stimulus (conditioned stimulus; CS) occur when it is paired with a valenced stimulus (unconditioned stimulus; US; for a review see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). Moreover, research on approach and avoidance (AA)

☆ Pieter Van Dessel is supported by a Ph.D. fellowship of the Scientific Research Foundation, Flanders (FWO-Vlaanderen). Jan. De Houwer is supported by the Methusalem Grant BOF09/01M00209 of Ghent University and by the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office (IUAPVII/33). The research in this paper has been supported by Grant FWO12/ASP/275 of FWO-Vlaanderen.

* Corresponding author at: Ghent University, Department of Experimental–Clinical and Health Psychology, Henri Dunantlaan 2, B-9000 Ghent, Belgium.
*E-mail address:* Pieter.vanDessel@UGent.be (P. Van Dessel).

http://dx.doi.org/10.1016/j.jesp.2015.11.002
0022-1031/

training has suggested that changes in implicit evaluations can be obtained by pairing a stimulus with a valenced action (i.e., approach or avoidance). Typically, the repeated approaching of one stimulus and avoiding of another stimulus leads to more positive implicit evaluations for the former stimuli (e.g., Kawakami, Phills, Steele, & Dovidio, 2007; Woud, Maas, Becker & Rinck, 2013; but see Vandenbosch & De Houwer, 2011).

Recent research has, however, shown that implicit evaluations change even when pairings are not experienced directly, but are implied by the verbal presentation of relational information via instructions. For instance, studies on instructed EC have shown that changes in the implicit evaluation of a CS occur when verbal instructions link a CS with a valenced US even when the CS–US pairings are not experienced directly (De Houwer, 2006; Gast & De Houwer, 2012). Similarly, in a recent study we observed typical AA training effects when participants did not actually perform AA actions, but were merely instructed that they would later have to perform these actions (Van Dessel, De Houwer, Gast, & Smith, 2015). That is, participants who received instructions to approach one fictitious social group (e.g., Niffites) and avoid another fictitious social group (e.g., Luupites) showed a preference for the former group both on implicit measures (i.e., the Implicit Association Test, IAT, Greenwald, McGhee, & Schwartz, 1998; and the evaluative priming task, Fazio, Sanbonmatsu, Powell, & Kardes, 1986) and explicit measures of evaluation.

These findings pose a challenge to associative and dual-process models of evaluation which assume that implicit evaluations result from the gradual formation of associations in memory as the result of actual pairings (Rydell & McConnell, 2006; Smith & DeCoster, 2000). In contrast, contemporary dual-process models in which association formation processes can interact with propositional learning processes allow one to explain the effects of instructions on implicit evaluations. For instance, the Associative-propositional Evaluation (APE) model (Gawronski & Bodenhausen, 2006, 2011, 2014) postulates that associations may sometimes arise as the result of the generation and validation of propositions. More specifically, when people determine in a propositional manner that a stimulus is either positive or negative this may instigate the proactive construction of new associations between representations of the stimulus and representations of positivity or negativity. As a result, any information that allows participants to consciously entertain the proposition that a stimulus is positive or negative may influence implicit evaluations. In line with this idea, changes in implicit evaluations have been observed when participants are provided with information about the valenced properties of a stimulus (Castelli, Zogmaister, Smith, & Arcuri, 2004; Cone & Ferguson, 2015; Gregg, Seibt, & Banaji, 2006; Whitfield & Jordan, 2009).

Importantly, these models predict a specific pattern of mediation such that instruction effects on explicit evaluation should mediate effects on implicit evaluation (see Gawronski & Bodenhausen, 2006; Case 4). That is, instructions should first influence whether participants consider a stimulus positive or negative (which is reflected in explicit evaluations) before this may lead to the formation of novel associations (which is reflected in implicit evaluations). Support for this idea was found by Whitfield and Jordan (2009), who observed that receiving information about the behavior of unknown individuals caused changes in implicit evaluations of these individuals that were fully mediated by changes in explicit evaluations.

Contrasting this result, our previous study on AA instruction effects provided evidence that changes in explicit evaluations do not fully mediate the effects of AA instructions on implicit evaluations. Statistical mediation analyses indicated that the impact of AA instructions on implicit evaluations was partly mediated by changes in explicit evaluations, but an effect remained after controlling for changes in explicit evaluation (Van Dessel et al., 2015). This is an intriguing finding because it suggests that mere (AA) instructions may sometimes cause unintentional changes in (implicit) stimulus evaluations. Instructions may have a direct effect on implicit evaluation (i.e., unmediated by changes in explicit evaluation) and may therefore cause changes in implicit evaluations even when participants do not consider the instructions as a valid basis for their (explicit) evaluation.

However, on the basis of the available evidence it is premature to conclude that AA instructions can influence implicit evaluation without any mediation by changes in explicit evaluation. Most importantly, our earlier AA instruction study (Van Dessel et al., 2015) included only statistical analyses of mediation. This measurement-of-mediation approach, however, is ultimately correlational in nature, and is thus problematic for establishing a causal chain (Spencer, Zanna, & Fong, 2005). This is especially the case when examining patterns of mediation between implicit and explicit evaluations. When a manipulation affects both implicit and explicit measures of evaluation, the particular direction of the obtained mediation pattern is strongly influenced by the internal consistency of the employed measure (Gawronski & Bodenhausen, 2011). Moreover, when implicit and explicit evaluations are strongly correlated (as was the case in our previous study), this creates multicollinearity which inflates the standard error of all variables in the mediation model and compromises the estimation of the indirect effect (Alin, 2010). Hence, when examining mediation of implicit and explicit evaluations, it is strongly recommended to supplement statistical mediation analyses with experimental manipulations (De Houwer, Gawronski, & Barnes-Holmes, 2013). This is particularly true if, as in our case, a theoretical debate requires the precise understanding of the causal relation.

In the current studies, we used both a statistical and an experimental approach to test the extent to which the impact of AA instructions on implicit evaluation is mediated by changes in explicit evaluation. We manipulated the proposed mediating variable (i.e., changes in explicit evaluation) by providing participants with 'trait instructions' that should prevent an impact of AA instructions on explicit evaluation. In line with Gregg et al. (2006), we asked participants to imagine that the members of one fictitious social group had very positive traits and the members of another fictitious social group had very negative traits (e.g., Niffites are peaceful, civilized, benevolent, and law-abiding; Luupites are violent, savage, malicious, and lawless). Subsequently, participants received instructions to approach or avoid these social groups. Whereas trait instructions directly specify the evaluative properties of the social group, AA instructions only provide evaluative information if participants infer that the task to approach or avoid members of a group tells something about the evaluative properties of that group. Participants might rely on this inference when they have no other information about the evaluative properties of the group, but even then they will probably be aware that this inference rests on shaky grounds. Prior research indeed suggests that participants are likely to refrain from using information that has a low diagnostic validity (such as AA instructions) when more valid information (such as instructions about evaluative traits) is available (Cone & Ferguson, 2015; Lynch, 2005). For these reasons, we expected that participants who received trait instructions would not take the AA instructions into account when explicitly evaluating the stimuli. We examined whether, under these circumstances, AA instructions would still cause changes in implicit evaluation. That is, we examined whether an AA instruction effect on implicit evaluation would be observed not only in the absence of mediation by changes in explicit evaluation, but even when there is no impact on explicit evaluation. The latter result would not only confirm that AA instructions can have a direct effect on implicit evaluation (because mediation via changes in explicit evaluation can occur only if there are changes in explicit evaluation) but would also support the novel conclusion that this direct effect can arise even when participants do not have the intention to use the AA instructions for evaluating the stimuli.

If we would find that AA instructions influence implicit evaluation in the absence of (mediation by) changes in explicit evaluation, this is bound to have important theoretical implications. First, it would strongly constrain current and future models of (implicit) evaluation. For