Differentiated evaluation of counter-conditioned stimuli as a function of right-wing authoritarianism
Amélie Bret, Brice Beffara Bret, Adrien Mierop, Martial Mermillod

To cite this version:

HAL Id: hal-03312622
https://nantes-universite.hal.science/hal-03312622
Submitted on 2 Aug 2021

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.
Differentiated Evaluation of Counter-Conditioned Stimuli as a Function of Right-Wing Authoritarianism

Amélie Bret¹,², Brice Beffara², Adrien Mierop³, Martial Mermillod¹


Abstract

Right Wing Authoritarianism (i.e., RWA) is associated with enhanced conservatism and social prejudice. Because research linking RWA to attitudes is largely correlational (i.e., it provides control for neither RWA nor attitude learning), it is not clear how RWA relates to attitude learning dynamics. We addressed this question in 11 evaluative conditioning experiments that ensured rigorous control of the affective learning setting. Results from two integrative data analyses suggest that (i) individuals scoring higher in RWA show a stronger acquisition of positive attitudes, and that (ii) the residuals of this stronger acquisition remain even after exposure to counter-attitudinal information. Implications of these findings for research on RWA and its link to social prejudice are discussed.

Keywords

authoritarianism, attitude, change, evaluative counter-conditioning
Two integrative data analyses are reported (N = 1034). We use evaluative counter-conditioning to understand attitude dynamics in right-wing authoritarianism. Contrary to our hypothesis, RWA is likely to be associated with a stronger acquisition of positive attitudes. The residuals of this stronger acquisition remain even after exposure to counter-attitudinal information.

Right Wing Authoritarianism (RWA) is characterized by a covariation of authoritarian aggression, authoritarian submission, and conventionalism (Altemeyer, 1981). Higher levels of RWA are typically associated with more negative attitudes concerning outgroups (Altemeyer, 1998; Duckitt & Sibley, 2007; Duriez & Van Hiel, 2002; Hodson et al., 2017). To date, however, research on RWA and attitudes has been largely observational: both RWA and social attitudes are often studied as observed variables (i.e., without considering causal determinism). This observational strategy has advanced our understanding of how RWA relates to prejudice and social attitudes. However, it does not allow optimal control when investigating attitude learning dynamics as a function of RWA. The aim of the present research was to contribute to the latter question by making use of a carefully controlled attitude learning paradigm. It should be noted, however, that RWA will not be studied as an experimentally manipulated variable here. In this series of experiments, we shall simply try to manipulate the acquisition and change of attitudes. We start by introducing RWA and its relation to attitudes and by considering evaluative conditioning as a controlled attitude learning procedure. We then report and present two integrative data analyses conducted on data from 11 experiments that shed light on our research question, namely how attitudes are acquired and how they resist counter-attitudinal information as a function of RWA.

RWA and Attitudes

RWA is a classic predictor of prejudice (Altemeyer, 1981). It correlates with negative attitudes to a range of social groups (Ekehammar et al., 2004): African Americans (Altemeyer, 1998; Lambert & Chasteen, 1997), homosexuals, (Altemeyer, 1998; Lippa & Arad, 1999), women (Altemeyer, 1998), and immigrants (Zakrisson, 2005). RWA is commonly described as social conservatism because it is positively correlated with conservative beliefs in the socio-political domain (e.g., Van Hiel & Mervielde, 2002). Furthermore, beyond the correlation between RWA and negative attitudes towards outgroups, RWA is associated with more stable attitudes over a period of six months (Asbrock et al., 2010) and with a lower likelihood that these attitudes will change.
This enhanced stability of attitudes in high-RWA individuals is consistent with studies suggesting that RWA is associated with lower cognitive flexibility (Onraet et al., 2015; Sidanius, 1985; Zmigrod, 2020). Several studies have highlighted the importance of adaptation to change in understanding the emergence and maintenance of prejudice (Dhont & Hodson, 2014; Roets & Van Hiel, 2011; Stern & Axt, 2021). In addition, RWA is typically linked with Need for Closure, defined as the motivation to have a definite answer or knowledge instead of uncertainty or doubts about the social environment, and can therefore be linked to a stronger ‘primacy effect’ (i.e., higher levels of RWA are associated with a better memory for the first items on a list as compared to the last items; Jost et al., 2003; Kruglanski et al., 2006).

It should be noted that the conceptualization of right-wing authoritarianism has important implications for its general understanding and even more so in this type of research (Van Hiel, Pandelaere, & Duriez, 2004). Some authors consider right-wing authoritarianism to be a personality trait (Adorno et al., 1950), while others consider it to be a belief (Duckitt, 2001). If authoritarianism is viewed as a personality trait, the associated cognitive differences should be considered a consequence of authoritarian thinking (Van Hiel et al., 2004). In this case, authoritarianism can be seen as a predisposition to cognitive rigidity (as opposed to cognitive flexibility). By contrast, if we consider authoritarianism to be a belief, individuals with high cognitive rigidity would be predisposed to developing authoritarian thinking. Various studies corroborate this last hypothesis, showing an influence of cognitive style on the development of negative exo-group attitudes during adolescence. This suggests that authoritarianism is derived from political socialization (i.e., development) (Alwin & Krosnick, 1991; Duckitt, 2001). Whatever the case may be, this series of experiments is not intended to test these hypotheses or to establish any causal link.

The above literature suggests that high-RWA individuals may be less likely to revise their attitudes once they are formed. In other words, they may form evaluations (i.e., possibly acquired by evaluative conditioning, “EC” hereinafter) that are less sensitive to new and counter-attitudinal information (i.e., due to lower flexibility; for a meta-analytical review, see Jost et al., 2003). This could partly explain the link between RWA and prejudice. The current literature links RWA to attitudinal stability over time, but does not allow us to understand this lack of change except in terms of a multitude of social factors (group membership). At present, we do not know precisely how attitudes are formed and even less how they evolve in relation to RWA. What we propose here is to study RWA in a context free of social information. Therefore, if the hypothesis of a lower sensitivity to counter-attitudinal information is confirmed, the net outcome of assessing exposure to attitudinal and then counter-attitudinal information should reflect higher levels of attitudinal information than counter-attitudinal information as RWA increases.
Evaluative Conditioning as a Controlled Attitude Learning Paradigm

To our knowledge, no prior work has examined how RWA modulates this attitude dynamic. The exclusive observational nature of research on RWA and attitudes has allowed for little control of attitude formation and change. These observational studies typically relate RWA levels to attitudes toward existing social groups. As such, they introduce a series of uncontrolled factors, all of which are capable of accounting for the observed relation between RWA and attitudes (i.e., the social nature of the attitudinal objects, previous knowledge and interactions regarding the attitudinal objects, the number of instances of prior attitudinal and counter-attitudinal information regarding the attitudinal objects, the relevance of the context of the investigation for evaluations of the attitudinal objects, the social desirability of and the norms governing the expression of attitudes toward the attitudinal object, etc.). In the current research, we overcame this limitation (i) by employing a carefully controlled evaluative conditioning paradigm that permitted the causal control of attitude formation by means of an experimentally manipulated variable and (ii) by using a set of socially irrelevant and unfamiliar stimuli, for which parameters such as prior knowledge, time and frequency of exposure, social relevance, or group variability were controlled. Here, we consider the RWA variable as a general ideology.

Evaluative conditioning (EC) is a simple and reliable attitude learning paradigm, with an average Cohen’s $d = 0.52$ (Hofmann et al., 2010). It consists of pairing a neutral stimulus (conditioned stimulus, CS) with another positive or negative stimulus (unconditioned stimulus, US) (De Houwer et al., 2001; De Houwer, 2007). Using this pairing procedure, the CS typically acquires the valence of the US it was paired with. For example, if a neutral stimulus such as headphones is presented at the same time as a cute baby, the headphones will take on the baby's valence and become positive. The process is the same for negative images.

This paradigm is efficient for almost any type of stimulus. As a result, this paradigm can be used to manipulate the acquisition and change of an attitude towards a stimulus never previously encountered by the participants. The advantage of this method is that it allows us to examine the dynamics of attitudes in a setting without prior knowledge of or social motivation regarding the stimuli. In other words, it permits a purer examination of how RWA influences attitude dynamics, without being interfered with by uncontrolled variables (e.g., pre-existing attitudes, knowledge and motivations regarding existing social groups), which may and, indeed, are very likely to be at work when forming and adapting attitudes about social groups. At the same time, increasing our knowledge of basic learning mechanisms makes it possible to infer how social knowledge and motivations may influence attitudinal processes with regard to real social groups. However, it seems important (independently of the possible generalization to real social groups) to undertake a causal study of the acquisition of attitudes towards completely...
new stimuli in RWA (by means of experimentally manipulated variables), particularly in the light of the literature on cognitive skills and RWA.

**Method**

**Integrative Data Analyses**

Several EC studies were conducted by our research team between 2015 and 2018 in order to provide information relevant to the question of interest: Is RWA positively related to a resistance toward counter-attitudinal information? In each of the experiments, a set of Greebles (fictive characters, the CSs) were paired with valenced pictures (the USs), with half of the Greebles being paired with positive pictures and the other half being paired with negative pictures. This pairing procedure constitutes the “conditioning phase”. In ten of these experiments, evaluative RWA and CS ratings were collected directly after a “counter-conditioning phase”. To do this, the Greebles that had been paired with a given valence during the conditioning phase were then subsequently paired with USs of the opposite valence prior to the RWA and evaluative measures. In two experiments, RWA and evaluative ratings were collected immediately after the conditioning phase (meaning that there was no counter-conditioning phase in these experiments).

Here, we report and discuss two separate integrative data analyses on these two sets of experiments. These integrative data analyses have two main goals: to test whether RWA is related to more resistant attitudes in a counter-conditioning procedure (integrative data analysis 1), and to test whether this effect is due to differences in RWA present in acquisition attitudes in a conditioning procedure (integrative data analysis 2). First of all, we hypothesized that RWA would be positively related to a resistance to novel counter-attitudinal information. This should result in a final evaluation that is more consistent with the conditioning than the counter-conditioning phase. Let us illustrate this hypothesis with an example. In the conditioning phase, participant 1 sees CS-A with a positive US and CS-B with a negative US. In the counter-conditioning phase, these pairings are then reversed (i.e., CS-A is now presented with a negative US and CS-B with a positive US). In such an example, we would expect CS-A to be evaluated more positively than CS-B due to resistance to the second - counter-attitudinal - item of information. Secondly, we did not hypothesize any difference in the acquisition of information with RWA, i.e., in (simple) conditioning effects. As a methodological precaution, we chose only to evaluate the attitude after counter-conditioning in integrative data analysis 1. This approach makes it possible to reduce the demand effect bias that might result from multiple measures (Charness et al., 2012). However, this choice also reduces the amount of information available for understanding the dynamics of attitudes in RWA. We therefore chose to evaluate the attitude after the conditioning phase in integrative data analysis 2 in order to determine whether the link between RWA and the relative
resistance to counter-attitudinal information is the reflection of a link between RWA and the acquisition of attitudes.

The experiments included in these two integrative data analyses were designed to test a variety of research questions, which involved running both control and experimental conditions. However, the main goal of this line of research was the same: to understand the roots of resistance to new information in RWA. Stimuli, numbers of trials, measures of attitudes, time of presentation and counter-balanced design were constant across these various control conditions which are the focus of the analyses presented here (as they are directly relevant to our questions). The additional experimental conditions are not relevant to our current questions and will not be discussed. In most cases, they provided mixed or non-replicated evidence (see Table 1 for an overview of the experimental settings of the different experiments).

Table 1

<table>
<thead>
<tr>
<th>Experiment</th>
<th>IDA</th>
<th>N</th>
<th>Data collection</th>
<th>Outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IDA 1</td>
<td>31</td>
<td>Laboratory</td>
<td>N = 0</td>
</tr>
<tr>
<td>2</td>
<td>IDA 1</td>
<td>57</td>
<td>Laboratory</td>
<td>N = 0</td>
</tr>
<tr>
<td>3</td>
<td>IDA 1 and 2</td>
<td>202 and 197</td>
<td>Online</td>
<td>N = 2</td>
</tr>
<tr>
<td>4</td>
<td>IDA 1</td>
<td>37</td>
<td>Laboratory</td>
<td>N = 0</td>
</tr>
<tr>
<td>5</td>
<td>IDA 1</td>
<td>54</td>
<td>Online</td>
<td>N = 0</td>
</tr>
<tr>
<td>6</td>
<td>IDA 1</td>
<td>58</td>
<td>Online</td>
<td>N = 0</td>
</tr>
<tr>
<td>7</td>
<td>IDA 1</td>
<td>75</td>
<td>Online</td>
<td>N = 0</td>
</tr>
<tr>
<td>8</td>
<td>IDA 1</td>
<td>61</td>
<td>Online</td>
<td>N = 0</td>
</tr>
<tr>
<td>9</td>
<td>IDA 1</td>
<td>59</td>
<td>Laboratory</td>
<td>N = 0</td>
</tr>
<tr>
<td>10</td>
<td>IDA 1</td>
<td>91</td>
<td>Online</td>
<td>N = 0</td>
</tr>
<tr>
<td>11</td>
<td>IDA 2</td>
<td>115</td>
<td>Online</td>
<td>N = 0</td>
</tr>
</tbody>
</table>

Sample Size

We initially performed a priori power analyses to determine the required sample size for the individual studies taken separately. However, after data collection, we realized that these power analyses were derived from models which were not suitable for the type of data we had collected. In addition, the estimated expected sample size was inadequate because the initially specified parameters were inaccurate. What is more, the final sample sizes for the individual studies are highly heterogeneous as they depended on the availability of resources at the time we were running each study. Because the present article focuses on the integrative data analysis of several studies, we explain how we subsequently considered its a posteriori sensitivity. This sensitivity results from the analysis described in the data analysis section and makes it possible to determine
our smallest detectable effect size (Lakens, 2021). More details can be found in the Supplementary Materials section.

The smallest effect size we were able to detect was determined based on the precision of the estimation of our parameter of interest (Lakens, 2017; Lakens et al., 2018). We considered that a null effect corresponded to a parameter value lower than $|0.13|$ (see the Supplementary Materials for more information on how we determined this value). In other words, we would consider a 95% HDI (Highest Density Interval) lying between -0.13 and 0.13 as a null effect. Parameter values $< |0.13|$ are considered practically equivalent to 0 (Kruschke, 2018).

**Integrative Data Analysis 1: Net Outcome of a Conditioning-Then-Counter-Conditioning Procedure**

Ten studies were run examining the association between RWA and the evaluation of stimuli using a conditioning-then-counter-conditioning procedure. We hypothesized that the net outcome of a conditioning-then-counter-conditioning procedure would indicate that more evaluative information is acquired in the conditioning than in the counter-conditioning phase at higher levels of RWA. In other words we expected negative evaluations of negatively conditioned (and positively counter-conditioned) stimuli in higher-RWA individuals. We expected positive evaluations of positively conditioned (and negatively counter-conditioned) stimuli in higher-RWA individuals.

**Participants and Design**

A total of 720 participants were recruited in exchange for credits (for laboratory studies) or money (for online studies). The participants in the laboratory studies were French-speaking (French and Belgian students) and those in the online studies were English-speaking (from the general population). Participants were randomly assigned to one of the two counter-balanced groups in which half of the CSs (i.e., group A) were paired either with positive or negative USs and the other half (i.e., group B) were paired either with negative or positive USs, respectively. This counter-balanced design did not moderate the effects and so will not be discussed further.

**Stimuli and Measures**

These studies were conducted in the laboratory or online. We used Testable and Prolific Academic for the online studies and E-prime software 2.0 for the laboratory studies (Schneider et al., 2002). The US set consisted of eight unpleasant and eight pleasant pictures from the International Affective Picture System (IAPS; Lang et al., 2008; see Supplementary Materials). Pleasant pictures were rated more positively than unpleasant pictures, $t(14) = 21.14$, $p < .001$, but it was not possible to conclude that there was a difference in arousal, $t(14) = .76$, $p = .46$. The CSs were 16 pictures from two different
groups of neutral characters that were equalized on contrast and luminance (i.e., “Greebles”, Sheinberg & Tarr, 2009). RWA was assessed using 10 items (Cronbach’s $\alpha = .837$) from the French translation of Altemeyer’s 20-item (1988; Bougie & Perreault, 2006) RWA Scale (from 1 = totally disagree to 7 = totally agree).

**Procedure**

For a visualization of the experimental design see Figure 1.

*Categorization stage:* In the first step, participants were asked to carefully examine, for an unlimited duration, one slide displaying the two groups of CSs (i.e., Greebles groups A and B). They were then randomly presented with individual CSs and had to sequentially categorize them into one of the two groups by pressing a corresponding key. After each trial, participants received feedback about the accuracy of the answer.

*Figure 1*

Description of the Experimental Procedure: From the First Phase (Left) “Categorization Stage” to the Last One “RWA Scale Completion” (Right)

*Conditioning and counter-conditioning stage:* In the second step, participants took part in a conditioning procedure in which Greebles were positively or negatively conditioned (depending on their group and the counter-balancing condition). More specifically, par-

---

1) Across all the experiments, accuracy was high (above 98%), thus indicating participants’ high capacity to distinguish between the two groups.
Participants were presented with individual CSs (i.e., Greeble exemplars) that were paired either with individual positive or negative USs (i.e., a positive or negative picture). Each individual Greeble was paired with a given US four times, resulting in 64 trials. Any given group (A or B, composed of several Greebles) was therefore paired 32 times with either positive or negative USs (pairing of groups with positive or negative USs was counterbalanced among participants). A trial consisted in the simultaneous presentation of a CS and a US for 1000ms, directly followed by a black screen displayed for 1500ms (inter-trial interval, ITI). The conditioning phase was immediately followed by the counter-conditioning. CSs were paired with USs whose valence was opposite to that used in the conditioning phase for the same number of trials and using the same pairing procedure. Nothing indicated to the participants that a new phase had started.

**Evaluative stage and scale completion:** After the counter-conditioning phase, participants were asked to state their attitude towards each individual Greeble (from 1 = "very negative" to 9 = "very positive"). CSs were presented in a random order. Finally, the participants completed the RWA scale, were thanked, debriefed and dismissed.

**Statistics and Results**

We report posterior means and 95% Highest Density Intervals (HDIs) for each parameter (β) based on a multilevel Bayesian ordinal regression (Bürkner & Vuorre, 2018; Liddell & Kruschke, 2018) computed with the “brms” package (Rstudio; Bürkner, 2017a, 2017b, RStudio Team, 2021). HDIs correspond to the probability density of “credible” values of β. We interpret these values in the light of the 95% highest probability mass (“confidence interval” in the Bayesian perspective). For an accessible introduction to Bayesian data analysis and interpretation, see Kruschke and Liddell (2018). More information about this model can be found in the Supplementary Materials.

We included the initial valence (i.e., valence corresponding to the US valence initially attached to a CS family in the evaluative conditioning stage; within-participant dichotomous variable: negative = -0.5 or positive = +0.5) and RWA (observed variable initially bounded between 7 and 70 and with precision 1 and then transformed into a z-score), as well as their interaction as fixed factors and varying intercepts for participants and stimuli in the model. The dependent variable is the evaluation of Greebles on the ordinal scale from 1 (very negative) to 9 (very positive) (Liddell & Kruschke, 2018). Integrative data analysis was performed on the data of all experiments: Experiments 1 and 11 for integrative data analysis 1 and Experiments 1 through 10 for integrative data analysis 2 (see Table 1 for an overview). The same parameters were kept for the integrative data analyses and a varying intercept for experiments was added. Before running the models, and before centering and scaling variables, we checked for the presence of possible extreme RWA scores in our sample (possibly due to measurement errors including -for instance- misunderstanding the questions or disengaging from the task). This step was based on quartile statistics, i.e., first quartile (Q1 = 25% of the sample), third quartile
(Q3 = 75% of the sample) and interquartile range (IQR = Q3 - Q1). RWA scores were considered extreme when > Q3 + 3*IQR or < Q1 - 3*IQR. We excluded two participants over these 10 experiments (see Table 1 for an overview of all the experiments). The results are reported in accordance with the guidelines proposed by Makowski et al. (2019), as adapted for our ordinal models. For easier reading, we only report the results concerning our main parameters of interest in the manuscript. However, two tables detail the results concerning all parameters in the Supplementary Materials.

The interaction between the initial valence and RWA has a probability of 100% of being positive (Median = 0.21, 95% HDI [0.15, 0.27], see Figure 2 and Figure 3.1) and can be considered to be very small (Sawilowsky, 2009) according to standard benchmarks but significant (0.50% in the region of practical equivalence, $R_{naz}^2 = .003 = d = 0.12^2$). Higher RWA scores were associated with higher evaluations of the positively conditioned CSs (Median = 0.16, 95% HDI [0.11, 0.22], probability of 100% of being positive, see Figure 2 and Figure 3.2). It is uncertain whether this effect is practically significant (11.68% in the region of practical equivalence). However, for the negatively conditioned CSs, despite the fact that the association between RWA and evaluation has a 93.75% probability of being negative, (Median = -0.04, 95% HDI [-0.10, 0.01], see Figure 2 and Figure 3.3), the effect seems to be practically equivalent to 0 (99.72% in the region of practical equivalence).

2) The effect sizes reported here should be interpreted in the context of future similar studies rather than compared to benchmarks. Comparisons to benchmarks are simply given in order to permit familiarity. So far, it is complicated to determine the smallest practically significant effect in the context of this study. The reader is free to interpret the data. The most important part of the analysis is the ongoing estimation of parameters that can be used for meta-analyses or integrative analyses.

3) See the Supplementary Materials for more details about the computation of effect sizes.
Figure 2

*Evaluation of Stimuli After Counter-Conditioning as a Function of RWA and Conditioning Valence*

![Evaluation of stimuli after counter-conditioning as a function of RWA and conditioning valence: negative (left) vs. positive (right). Based on the estimation given by the ordinal model. Correct graphical representation of the output of the model.](image1)

**Note.** Top: Evaluation of stimuli after counter-conditioning as a function of RWA and conditioning valence: negative (left) vs. positive (right). Based on the estimation given by the ordinal model. Correct graphical representation of the output of the model. *Bottom left:* Evaluation of stimuli after counter-conditioning as a function of RWA and conditioning valence. Based on the estimation given by the ordinal model but adapted as a linear representation. This representation is simply given in order to permit familiarity but does not rigorously represent the ordinal nature of the model. The lines correspond to randomly sampled credible regression lines. *Bottom right:* Distribution of the data points.

Figure 3.1

*Forest Plot Showing Interaction Parameter Estimates for the Interaction Effect Between the Initial Valence and RWA in the Evaluation of the Greebles*

![Forest Plot Showing Interaction Parameter Estimates for the Interaction Effect Between the Initial Valence and RWA in the Evaluation of the Greebles](image2)

**Note.** The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the Supplementary Materials.
Figure 3.2

*Forest Plot Showing Simple Slope Parameter Estimates for RWA in the Evaluation of Positively Conditioned Greebles (Following Negative Counter-Conditioning)*

Note. The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the Supplementary Materials.

Figure 3.3

*Forest Plot Showing Simple Slope Parameter Estimates for RWA in the Evaluation of Negatively Conditioned Greebles (Following Positive Counter-Conditioning)*

Note. The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the Supplementary Materials.

Discussion of Integrative Data Analysis 1

The outcomes of integrative data analysis 1 are partially consistent with our predictions. As in the case of social attitudes, evaluations become more closely aligned with attitu-
dinal than counter-attitudinal information as the RWA level increases. However, this interaction between valence and RWA does not seem to rely significantly on negative conditioning and is probably mainly driven by positive conditioning, although more data are probably needed to (in)validate this point. We did not initially predict this possible asymmetry. As mentioned above, an interesting question that remains unanswered at this stage is whether the link between RWA and the relative resistance to counter-attitudinal information reflects a link between RWA and the acquisition of attitudes. Indeed, we hypothesized that RWA was associated with a lack of flexibility, namely a difficulty in processing and assimilating new (counter-attitudinal) information. However, while this hypothesis is compatible with the results of this IDA, at least one other explanation remains possible. Indeed, these findings could instead result from a stronger effect of initial conditioning for higher RWA individuals. One possible way to address this issue is to compare the evaluations of Greebles in two experimental conditions (across the RWA continuum): one condition where the CSs are counter-conditioned and another condition where they are only conditioned.

Since we would still have (only) one CS evaluation per participant using this design, we would not increase the demand effect. Thus, if the link between RWA and the relative resistance to counter-attitudinal information found in integrative data analysis 1 simply reflects attitude acquisition differences, then the higher the RWA score, the higher the ratings should be when a positive attitude is acquired.

**Integrative Data Analysis 2: Acquisition of Positive and Negative Attitudes**

Two studies examined the acquisition of positive and negative attitudes as a function of RWA in an evaluative conditioning procedure. A recent meta-analysis on the impact of conservatism on cognitive functions did not report an effect of RWA in (non-evaluative) learning (Van Hiel, Onraet, & De Pauw, 2010). In line with this meta-analysis, we hypothesized that attitude acquisition would not vary as a function of RWA.

**Participants and Design**

A total of 317 participants were recruited in exchange for credits (for laboratory studies) or money (for online studies). The participants in the laboratory studies were French-speaking (French and Belgian students) and those in the online studies were English-speaking (from the general population). Participants were randomly assigned to one of the two counter-balanced groups in which half of the CSs (i.e., group A) were paired with either positive or negative USs and the other half (i.e., group B) were paired with either negative or positive USs, respectively. This counter-balanced design did not moderate effects and so will not be discussed further.
Stimuli and Measures

These two experiments were conducted online using Testable and Prolific Academic. Materials were the same as for integrative data analysis 1.

Procedure

Categorization stage: This step was similar to the one presented in integrative data analysis 1.

Conditioning stage: The conditioning phase, similar to that reported for integrative data analysis 1, was immediately followed by the CS evaluation.

Evaluative stage and scale completion: Participants were asked to state their attitude towards each individual Greeble (from 1 = "very negative" to 9 = "very positive"). CSs were presented in a random order. Finally, the participants completed the RWA scale, were thanked, debriefed and dismissed.

Statistics and Results

The rationale underlying the data analysis was the same as for integrative data analysis 1. We also checked for the presence of possible extreme RWA scores in our sample. We excluded two participants over these two experiments (see Table 1 for an overview of all the experiments). The interaction between the initial valence and RWA has a probability of 100% of being positive (Median = 0.22, 95% HDI [0.12, 0.33], see Figure 4 and Figure 5.1). It is uncertain whether this effect is statistically significant. Even though only 4.05% of the parameter values lie in the region of practical equivalence, the 95% HDI slightly overlaps with it. This effect can be considered to be very small (Sawilowsky, 2009) according to standard benchmarks (pseudo-\(R^2 = .003 \approx d = 0.12\)). Higher RWA scores were associated with higher evaluations of the positively conditioned CSs (Median = 0.25, 95% HDI [0.14, 0.34], probability of 100% of being positive, see Figure 4 and Figure 5.2). This effect can be considered to be significant and small (0.95% in the region of practical equivalence). However, for the negatively conditioned CSs, the association between RWA and evaluation has a 70.85% probability of being negative, (Median = -0.03, 95% HDI [-0.07, 0.13], see Figure 4 and Figure 5.3), but the effect seems to be practically equivalent to 0 (97.67% in the region of practical equivalence).

4) The effect sizes reported here should be interpreted in the context of future similar studies rather than compared to benchmarks. Comparisons to benchmarks are simply given in order to permit familiarity. So far, it is complicated to determine the smallest practically significant effect in the context of this study. The reader is free to interpret the data. The most important part of the analysis is the ongoing estimation of parameters that can be used for meta-analyses or integrative analyses.

5) See the Supplementary Materials for more details about the computation of effect sizes.
**Figure 4**

*Evaluation of Stimuli After Conditioning as a Function of RWA and Conditioning Valence*

![Figure 4](image)

**Note.** Top: Evaluation of stimuli after conditioning as a function of RWA and conditioning valence: negative (left) vs. positive (right). Based on the estimation given by the ordinal model. Correct graphical representation of the output of the model. Bottom left: Evaluation of stimuli after conditioning as a function of RWA and conditioning valence. Based on the estimation given by the ordinal model but adapted as a linear representation. This representation is simply given to permit familiarity but does not rigorously represent the ordinal nature of the model. The lines correspond to randomly sampled credible regression lines. Bottom right: Distribution of the data points.

**Figure 5.1**

*Forest Plot Showing Parameter Estimates for the Interaction Effect Between RWA and the Initial Valence in the Evaluation of Conditioned Greebles*

![Figure 5.1](image)

**Note.** The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the Supplementary Materials.
**Figure 5.2**

*Forest Plot Showing Simple Slope Parameter Estimates for RWA in the Evaluation of Positively Conditioned Greebles*

Note. The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the [Supplementary Materials](#).

**Figure 5.3**

*Forest Plot Showing Simple Slope Parameter Estimates for RWA in the Evaluation of Negatively Conditioned Greebles*

Note. The purple area corresponds to a region of practical equivalence where the values are considered equivalent to 0 (Kruschke, 2018). More information can be found in the [Supplementary Materials](#).

**Discussion of Integrative Data Analysis 2**

Even though some uncertainty remains at the statistical level, it is important to note that the results of the second IDA tend to run counter to our initial hypothesis. Indeed,
contrary to our hypotheses, higher RWA levels were associated with stronger attitude acquisition. In line with the results of the first IDA, this effect is reliably observed only in the acquisition of positive attitudes, but not in the acquisition of negative attitudes. It is noteworthy that this effect was replicated in a third conditioning experiment that could not be included in integrative data analysis 1 because it used difference scale anchors for the evaluative measure. Hence, the RWA effect in the acquisition of positive attitudes seems to be reliable and robust, at least for the materials and procedures considered here. However, a larger sample size would enhance the precision and sensitivity of the test of the interaction between RWA and valence. These findings suggest that the results obtained in the first IDA are likely to be driven by a positivity effect in attitude acquisition (but not attitude change) in higher-RWA individuals. We elaborate on this point in the general discussion.

**General Discussion**

The two integrative data analyses reported here permit a first examination of how RWA relates to attitude learning dynamics in a controlled evaluative learning setting. Because RWA is associated with more stable attitudes over time and is negatively associated with cognitive flexibility (Sidanius, 1985; Zmigrod, 2020), we hypothesized that individuals with higher levels of RWA would exhibit lower sensitivity to information encountered in the second stage (counter-attitudinal information) than to the information encountered first (attitudinal information, resulting from conditioning).

Together, the data collected here suggest that high RWA scorers (here, RWA is understood as a general ideology) exhibit a preferential acquisition of positive attitudes. We observe the residuals of this higher sensitivity to initial (positive) conditioning after the exposure to counter-attitudinal information. When we commenced the present research project, we initially predicted that RWA would be associated with more resistance to counter-attitudinal information but not with facilitated acquisition of evaluations. This prediction was not supported by the results. More specifically, and against expectations, we observed initial evidence suggesting a positivity bias in attitude acquisition as a function of RWA.

At first sight, this positivity effect is difficult to reconcile with the literature pointing to a negativity bias in high RWA scorers (Castelli & Carraro, 2011; Shook & Fazio, 2009), which is defined as a higher psychological sensitivity to negative information (Baumeister et al., 2001). In addition, RWA has been linked to a greater allocation of resources to negative information (Hibbing et al., 2014). More surprisingly, when participants are allowed to look at images (positive and negative) without any time limitation and without any particular task to perform, there is a positive correlation between the level of conservatism (which is an important dimension of RWA) and fixation time for
negative images (Dodd et al., 2012). Such an attentional bias should have favored rather than lowered negative conditioning effects in high-RWA individuals.

One first possible explanation for this discrepancy would be that individuals who are more sensitive to negative information are also those who are more likely to try to stop negative experiences from arising when initially given a chance to form positive evaluations (Hibbing et al., 2014). From a functional perspective, RWA is considered an ideological response designed to reduce a high level of anxiety and threat perception (Katz, 1960). The direct consequence of this threat perception is the implementation of safety strategies to prevent negative events (Doty et al., 1991; Lavine et al., 2005). Based on this literature and our current results, we can hypothesize (for future studies) that an increased sensitivity to positive information could serve this preventive purpose. In addition, the greater conditioning for positive stimuli that we observed in the conditioning phase may be consistent with an enhanced Need for Closure in high RWA scorers (Napier & Jost, 2008). Indeed, more conservative participants tend to prefer very simple things that quickly satisfy their needs with a minimum of ambiguity (Kruglanski et al., 2006).

A second important point to consider is the social context. Previous studies indicating a negativity bias in high-RWA individuals have focused on effects that occur after an initial learning phase (e.g., during memory consolidation, retrieval, decision making, or behavioral enactment) and this is not optimally taken into account in the deliberately simplified paradigm used here. Intergroup processes, prior knowledge, and social motivations can undoubtedly modulate attitudinal effects when dealing with real social groups that involve complex social knowledge and motivations. In the light of this observation, it should be noted that the effect sizes of the statistically significant effects described in our study are very small. This is not surprising given the absence of social context in our protocols, as we discussed above (see also our discussion about the ecological aspects of our study below). This should be considered in combination with the nature of our task, which is not a self-reported measure of cognitive style. Indeed, the meta-analysis carried-out by Van Hiel et al. (2016) indicates that behavioral measures of cognitive style usually yield smaller effect sizes than self-reported measures when addressing the question of right-wing attitudes (see also Zmigrod, 2020, for a more recent overview). Further studies comparing social vs. non-social tasks and laboratory vs. field experiments could inform us about the minimum practically significant effect size in this field of research. Our integrative data analysis shows the number of observations needed to detect effect sizes of $d \approx 0.1$ (and null effects corresponding to $d < 0.07$) in this context (Lakens, 2017, 2021; Lakens et al., 2018).

A third (complementary and not exclusive) possible way of accounting for asymmetrical valence effect in both the conditioning and counter-conditioning designs is to consider that some CSs were evaluated more positively even before the beginning of the

---

6) The same database we used in the experiments presented here.
experiment. Thus, we could have added pre-conditioning ratings in order to address this issue. However, even if we can hypothesize that ratings are more positive with increasing RWA (as suggested by the results presented in the two integrative data analyses), we also hypothesize that all the stimuli (CS-A and CS-B) should be rated the same. In addition, CSs were counterbalanced between participants, that is to say some participants got to see CS-A with positive conditioning while others saw CS-B with positive conditioning. Thus, the valence asymmetry cannot be explained by a more positive evaluation of one specific group of CSs before conditioning.

Even if the evidence clearly suggests that RWA is associated with a certain approach to negative stimuli and is consequently not associated with particularly positive evaluations (Dodd et al., 2012), we still do not know with certainty whether RWA is associated with a more positive evaluation of neutral or mildly positive stimuli. It would be useful to evaluate the attitude towards neutral stimuli in the light of RWA scores without any acquisition of attitude phase. This would allow us to identify the net impact of conditioning or counter-conditioning on the CS.

Related to the previous point, the evaluation of the resistance to change in these studies can be criticized. Indeed, there are no before-and-after measurements that permit a real assessment of resistance. Conditioning has certainly taken place (whatever the level of RWA and the associated valence). Indeed, in the second IDA, we observed that all participants (regardless of RWA level) were receptive to the evaluative conditioning (but with different magnitudes associated with RWA). Based on these experiments, we can assume that the evaluation of attitudes (after counter-conditioning) reflects a combination of the conditioning and counter-conditioning effects. There are probably no substantial differences between lower and higher RWA levels in the assimilation of counter-attitudinal information. Instead, RWA is likely to be associated with stronger initial conditioning (at least in the positive valence). The observations when attitudes were evaluated after counter-conditioning therefore probably reflect a residual effect of this stronger conditioning. Seen in this light, we provide (still incomplete) evidence that RWA is not specifically associated with cognitive rigidity but rather with higher sensitivity to initial (positive) information.

We accept that in order to permit a clear-cut test of attitude change in these experimental designs, it would be useful to conduct new studies which include two new measures of attitudes: one before the first conditioning phase and one between the conditioning and counter-conditioning phases (while also continuing to measure attitudes after counter-conditioning). This might allow us to observe the detailed evolution of attitudes across the different conditioning phases (and thus clearly measure the potential resistance to change in relation to RWA). However, it is important to be aware of other biases that such a procedure introduces. Measuring attitudes toward the same stimuli three times could lead to 1) a considerable demand bias, 2) an artificial focus on the
change of valence between the two conditioning phases, and 3) a certain weariness in performing the experiment.

We want to emphasize that this research examined a restricted kind of stimuli, thus limiting the generalizability to different types of stimuli. Unpublished data from our team\(^7\) show that sensitivity to initial positive information compared to counter-attitudinal information is also observed with even less social stimuli (other types of Greebles were used in this experiment). Nevertheless, we believe that this effect should be tested with other types of stimuli, for example, in the form of more ecologically valid stimuli, or with additional measures involving, for example, the control or manipulation of the social distance between the stimuli and the participants. Indeed, in many studies, RWA has been found to be correlated with social distance (see for example: Koc & Anderson, 2018; Sayilan et al., 2020). However, this is a more complex question because its investigation requires us to reduce the level of control of attitude acquisition. These limitations reveal the need for much further research on attitudinal and counter-attitudinal acquisition.

Another dimension relating to the ecological nature of the research concerns the context in which the participants respond. Future research could address the issue of attitudinal change while taking participants' political context into account. Indeed, recent studies have shown a link between national identity and RWA (Vargas-Salfate et al., 2020). More specifically, RWA positively predicts national identification. It would therefore be very relevant to look at the interaction between RWA and the political context of the home country in the evolution of attitudes. More specifically, it would be very interesting to investigate this effect in authoritarian countries in order to (1) obtain a larger panel of high-RWA participants and (2) investigate the cross-cultural components of this effect.

The ecological problem leads to a critical question: from an attitude-learning perspective, why are higher RWA levels related to greater prejudice? One intriguing possibility relates to the positivity bias found here. More specifically, it might be assumed that people’s environments are such that individuals are mostly exposed to positive information, on the one hand, and in-group information, on the other (Unkelbach et al., 2010). Hence, high-RWA individuals may form positive in-group attitudes which emerge from the initial information (normally) available in the environment. This, in turn, may contribute to out-group derogation effects and out-group avoidance in these individuals. More generally, more positive in-group views may fuel the resistance of high-RWA individuals to social changes and motivate various confirmatory biases in these individuals. Because the current study revealed residuals of positively acquired attitudes after exposure to counter-attitudinal information, one way to proceed would be to promote less biased exposure to social information when social attitudes form in these individuals.

---

\(^7\) A second condition of Experiment 2. This condition was not included in the integrative data analysis because the stimuli used in this condition were not equivalent to the stimuli used in the other condition and other experiments.
Conclusion

Here, we presented data showing that RWA is linked to a higher sensitivity to initial positive information in a controlled attitude acquisition paradigm. This finding contradicted our hypothesis about cognitive rigidity in RWA. As mentioned above, from a motivated social cognition perspective, specific information processing in RWA would develop in such a way as to meet the need for closure and the need for structure (Jost et al., 2003; Napier & Jost, 2008) and would be found particularly in right-wing ideologies (Jost, 2017). However, other authors defend the idea that these motives are not specific to right-wing ideologies but derive rather from a general ideological rigidity due to in-group favoritism (Greenberg & Jonas, 2003; Van Prooijen & Krouwel, 2019). Seen from this this perspective, the entire continuum of ideologies, from right wing to left wing, should be affected by the impact of ideology on information processing, at least when group identities are engaged. Further research should examine the association between ideology and cognitive style within such a paradigm in order to understand their underlying mechanisms and their causal relationship.

Funding: This work benefited from the Alpes Grenoble Innovation Recherche Grant from the University Grenoble Alpes and a “CNRS-Attentats” grant provided by the CNRS. We thank the Université de Nantes, the Université Grenoble Alpes, and the Université Catholique de Louvain for supporting our research. The authors received no specific financial support for the research, authorship, and/or publication of this article.

Acknowledgments: We thank the reviewers for insightful comments and suggestions during the peer-review process.

Competing Interests: The authors declare no potential conflicts of interest with respect to research, authorship, and/or publication of this article.

Ethics Statement: This research project was approved by the IPSY ethics committee in Louvain la Neuve (Belgium).

Data Availability: For this article, a dataset is freely available (Bret, Beffara, Mierop, & Mermillod, 2021).

Supplementary Materials

Materials (basic scripts and information about/links to the stimuli), data, and analysis scripts are available on OSF (for access see Index of Supplementary Materials below).

Index of Supplementary Materials

References


Dodd, M. D., Balzer, A., Jacobs, C. M., Gruszczynski, M. W., Smith, K. B., & Hibbing, J. R. (2012). The political left rolls with the good and the political right confronts the bad: Connecting


https://doi.org/10.1177/0146167204264333

https://doi.org/10.1177/0963721418817755


https://doi.org/10.1016/j.paid.2005.02.026

https://doi.org/10.1016/j.cobeha.2019.10.016