

A Multilab Replication of the Induced-Compliance Paradigm of Cognitive Dissonance



Advances in Methods and Practices in Psychological Science
January-March 2024, Vol. 7, No. 1,
pp. 1–26
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/25152459231213375
www.psychologicalscience.org/AMPPS



David C. Vaidis¹, Willem W. A. Sleegers², Florian van Leeuwen²,
Kenneth G. DeMarree³, Bjørn Sætrevik⁴, Robert M. Ross⁵,
Kathleen Schmidt⁶, John Protzko⁷, Coby Morvinski⁸,
Omid Ghasemi⁹, Andrew J. Roberts⁵, Jeff Stone¹⁰, Alexandre Bran¹¹,
Amélie Gourdon-Kanhukamwe¹², Ceren Gunsoy¹³,
Lisa S. Moussaoui¹⁴, Andrew R. Smith¹⁵, Armelle Nugier¹⁶,
Marie-Pierre Fayant¹¹, Ali H. Al-Hoorie¹⁷, Obed K. Appiah¹⁸,
Spencer Arbige¹⁹, Benjamin Aubert-Teillaud¹¹, Olga Bialobrzeska²⁰,
Stéphanie Bordel²¹, Valerian Boudjemadi²², Hilmar Brohmer²³,
Quinn Cabooter²⁴, Mehdi Chahir²⁵, Ianis Chassang²⁶,
Armand Chatard²⁷, Yu Yang Chou²⁸, Sungeun Chung²⁹,
Mioara Cristea³⁰, Joséphine Daga³¹, Gregory J. Depow³²,
Olivier Desrichard³³, Dmitrii Dubrov³⁴, Thomas R. Evans³⁵,
Séverine Falkowicz³¹, Sylvain Ferreira¹⁶, Tim Figureau³⁶,
Valérie Fointiat³¹, Théo Friedrich²⁶, Anastasia Gashkova³⁷,
Fabien Girandola³¹, Marine Granjon³⁸, Dmitry Grigoryev³⁹,
Gul Gunaydin⁴⁰, Şevval Güzel⁴¹, Mahsa Hazrati⁴², Mai Helmy⁴³,
Ayumi Ikeda⁴⁴, Michael Inzlicht³², Sara Jaubert³¹, Dauren Kasanov⁴⁵,
Mohammad Mohsen Khoddami⁴⁶, Taenyun Kim⁴⁷, Kiyoshi Kiyokawa⁴⁸,
Rabia I. Kodapanakkal⁴⁹, Alexandra Kosachenko⁵⁰,
Kortney Maedge⁵¹, John H. Mahaney⁵², Marie-Amélie Martinie²⁷,
Vitor N. Mascheretti⁵³, Yoriko Matsuda⁴⁸, Maxime Mauduy⁵⁴,
Nicolas Mauny⁵⁴, Armand Metzen¹¹, Eva Moreno-Bella⁵⁵,
Miguel Moya⁵⁶, Kévin Nadarajah⁵⁷, Pegah Nejat⁵⁸,
Elisabeth Norman⁵⁹, Irmak Olcaysoy Okten⁶⁰, Asil A. Özdoğru⁶¹,
Ceyda Ozer⁶², Elena Padial-Rojas⁶³, Yuri G. Pavlov⁶⁴,
Monica Perusquia-Hernandez⁶⁵, Dora Proost⁶⁶,
Aleksandra Rabinovitch⁶⁷, Odile Rohmer³⁸, Emre Selcuk⁴⁰,
Cécile Sénémeaud⁵⁴, Yaniv Shani⁶⁸, Elena A. Shmeleva⁶⁹,
Emmelie Simoens⁷⁰, Kaitlin A. Smith⁷, Alain Somat²⁵,
Hayeon Song⁷¹, Fatih Sonmez⁷², Lionel Souchet³¹, John J. Taylor¹⁰,

Corresponding Authors:

David C. Vaidis, Laboratoire Cognition, Langues, Langage, Ergonomie, Université de Toulouse, Centre National de la Recherche Scientifique, Toulouse, France








Email: david.vaidis@univ-tlse2.fr

Willem W. A. Sleegers, Social Psychology, Tilburg University, Tilburg, the Netherlands

Email: w.sleegers@me.com



Creative Commons NonCommercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits noncommercial use, reproduction, and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

**Ilja van Beest², Nicolas Van der Linden⁷³, Steven Verheyen⁷⁴,
Bruno Verschuere⁷⁵, Kevin Vezirian⁷⁶, Luc Vieira¹¹,
Sera Wiechert¹⁸, Guillermo B. Willis⁷⁷, Robin Wollast¹⁶, Ji Xia⁷⁸,
Yuki Yamada⁷⁹, Naoto Yoshimura⁸⁰, and Daniel Priolo³⁶**

¹Laboratoire Cognition, Langues, Langage, Ergonomie, Université de Toulouse, Centre National de la Recherche Scientifique, Toulouse, France; ²Social Psychology, Tilburg University, Tilburg, the Netherlands; ³Department of Psychology, SUNY Buffalo State University, Buffalo, New York; ⁴Department of Psychosocial Science, University of Bergen, Bergen, Norway; ⁵Department of Philosophy, Macquarie University, Sydney, Australia; ⁶Department of Psychology, Ashland University, Ashland, Ohio; ⁷Psychological Science, Central Connecticut State University, New Britain, Connecticut; ⁸Department of Management, Ben-Gurion University of the Negev, Beersheba, Israel; ⁹School of Psychological Sciences, University of New South Wales, Kensington, Australia; ¹⁰Department of Psychology, University of Arizona, Tucson, Arizona; ¹¹Laboratoire de Psychologie Sociale, Université Paris Cité, Paris, France; ¹²Department of Neuroimaging, King's College London, London, England; ¹³Department of Psychology, University of Rhode Island, South Kingstown, Rhode Island; ¹⁴Faculty of Psychology and Education, Health Psychology Research Group, University of Geneva, Geneva, Switzerland; ¹⁵Department of Psychology, Appalachian State University, Boone, North Carolina; ¹⁶Laboratoire Cognition, Langues, Langage, Ergonomie, Laboratoire de Psychologie Sociale et Cognitive, Université Clermont Auvergne, Clermont-Ferrand, France; ¹⁷Royal Commission for Jubail and Yanbu, Jubail English Language and Preparatory Year Institute, Jubail, Saudi Arabia; ¹⁸Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands; ¹⁹Department of Psychology, Macquarie University, Sydney, Australia; ²⁰Faculty of Psychology in Warsaw, Center for Research on Social Relations, Szkoła Wyższa Psychologii Społecznej University of Social Sciences and Humanities, Warsaw, Poland; ²¹PsyCAP, Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement (CEREMA), Saint-Brieuc, France; ²²Department of Psychology, Université de Strasbourg, Strasbourg, France; ²³Institute of Psychology, University of Graz, Graz, Austria; ²⁴Department of Psychology, Ghent University, Ghent, Belgium; ²⁵Department of Psychology, Université de Rennes 2, Rennes, France; ²⁶Department of Psychology, Université de Strasbourg, Strasbourg, France; ²⁷Centre de Recherches sur la Cognition et l'Apprentissage, Université de Poitiers, Centre National de la Recherche Scientifique, Poitiers, France; ²⁸Department of Neuroimaging, King's College London, London, England; ²⁹Media and Communication, Sungkyunkwan University, Seoul, South Korea; ³⁰Department of Psychology, Heriot Watt University, Edinburgh, Scotland; ³¹Laboratoire de Psychologie Sociale, Aix-Marseille Université, Aix-en-Provence, France; ³²Department of Psychology, University of Toronto, Toronto, Ontario, Canada; ³³Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, Geneva, Switzerland; ³⁴Center for Sociocultural Research, Higher School of Economics University, Moscow, Russia; ³⁵School of Human Science, University of Greenwich, London, England; ³⁶Laboratoire Epsilon, Université Paul-Valéry, Montpellier, France; ³⁷Center of Population Research, Ural Federal University, Yekaterinburg, Russia; ³⁸Laboratoire de Psychologie des Cognitions, Université de Strasbourg, Strasbourg, France; ³⁹Center for Sociocultural Research, National Research University Higher School of Economics, Moscow, Russia; ⁴⁰Department of Psychology, Sabanci University, Sabanci, Turkey; ⁴¹Department of Psychology, Üsküdar University, Istanbul, Turkey; ⁴²Department of Psychology, Shahid Beheshti University, Tajrish, Iran; ⁴³Psychology Department, College of Education, Sultan Qaboos University, Seeb, Oman; ⁴⁴Japan Society for the Promotion of Science, Kyushu University, Fukuoka, Japan; ⁴⁵Department of Psychology, Ural Federal University, Yekaterinburg, Russia; ⁴⁶Department of Psychology, Shahid Beheshti University, Tehran, Iran; ⁴⁷Department of Interaction Science, Sungkyunkwan University, Seoul, South Korea; ⁴⁸Cybernetics and Reality Engineering Laboratory, Nara Institute of Science and Technology, Ikoma, Japan; ⁴⁹Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, the Netherlands; ⁵⁰Department of Psychology, Ural Federal University, Yekaterinburg, Russia; ⁵¹Department of Psychology, Southern Illinois University, Carbondale, Illinois; ⁵²Department of Psychology, Appalachian State University, Boone, North Carolina; ⁵³Department of Psychology, Université Paul-Valéry Montpellier, Montpellier, France; ⁵⁴Laboratoire de Psychologie de Caen Normandie, Université de Caen Normandie, Caen, France; ⁵⁵Social and Organizational Psychology, Education Faculty of Psychology, National University of Distance, Madrid, Spain; ⁵⁶Centro de Investigación Mente, Cerebro y Comportamiento, University of Granada, Granada, Spain; ⁵⁷Department of Psychology, Rennes University, Rennes, France; ⁵⁸Faculty of Education and Psychology, Shahid Beheshti University, Tajrish, Iran; ⁵⁹Department of Psychosocial Science, Det Psykologiske Fakultet, Universitetet i Bergen, Bergen, Norway; ⁶⁰Department of Psychology, Florida State University, Tallahassee, Florida; ⁶¹Department of Psychology, Marmara University, TC Üsküdar Üniversitesi, Instabul, Turkey; ⁶²Department of Psychology, Sabanci Üniversitesi, Tuzla, Turkey; ⁶³Department of Social Psychology, University of Granada, Granada, Spain; ⁶⁴Laboratory of Neurotechnology, Ural Federal University, Yekaterinburg, Russia; ⁶⁵Information Science, Nara Institute of Science and Technology, Ikoma, Japan; ⁶⁶Psychology, Ghent University, Ghent, Belgium; ⁶⁷Psychology, Szkoła Wyższa Psychologii Społecznej University of Social Sciences and Humanities, Warsaw, Poland; ⁶⁸Management, Tel Aviv University, Tel Aviv, Israel; ⁶⁹Department of Psychology, Ivanovo State University, Ivanovo, Russia; ⁷⁰Faculteit Psychologie

en Pedagogische Wetenschappen, University of Ghent, Ghent, Belgium; ⁷¹Department of Psychology, Sungkyunkwan University, Seoul, South Korea; ⁷²Business Administration, Mus Alparslan University, Muş, Turkey; ⁷³Center for Social and Cultural Psychology, Université Libre de Bruxelles, Brussels, Belgium; ⁷⁴Department of Psychology, Education and Child Studies, Erasmus University, Rotterdam, the Netherlands; ⁷⁵Department of Clinical Psychology, University of Amsterdam, Amsterdam, the Netherlands; ⁷⁶Department of Psychology, Université Grenoble Alpes, Grenoble, France; ⁷⁷Psicología Social y Metodología de las Ciencias del Comportamiento, University of Granada, Granada, Spain; ⁷⁸Department of Psychology, SUNY Buffalo State University, Buffalo, New York; ⁷⁹Faculty of Arts and Science, Kyushu University, Fukuoka, Japan; and ⁸⁰Research Organization of Open Innovation and Collaboration, Ritsumeikan University, Kyoto, Japan

Abstract

According to cognitive-dissonance theory, performing counterattitudinal behavior produces a state of dissonance that people are motivated to resolve, usually by changing their attitude to be in line with their behavior. One of the most popular experimental paradigms used to produce such attitude change is the induced-compliance paradigm. Despite its popularity, the replication crisis in social psychology and other fields, as well as methodological limitations associated with the paradigm, raise concerns about the robustness of classic studies in this literature. We therefore conducted a multilab constructive replication of the induced-compliance paradigm based on Croyle and Cooper (Experiment 1). In a total of 39 labs from 19 countries and 14 languages, participants ($N = 4,898$) were assigned to one of three conditions: writing a counterattitudinal essay under high choice, writing a counterattitudinal essay under low choice, or writing a neutral essay under high choice. The primary analyses failed to support the core hypothesis: No significant difference in attitude was observed after writing a counterattitudinal essay under high choice compared with low choice. However, we did observe a significant difference in attitude after writing a counterattitudinal essay compared with writing a neutral essay. Secondary analyses revealed the pattern of results to be robust to data exclusions, lab variability, and attitude assessment. Additional exploratory analyses were conducted to test predictions from cognitive-dissonance theory. Overall, the results call into question whether the induced-compliance paradigm provides robust evidence for cognitive dissonance.

Keywords

cognitive dissonance, induced compliance, counterattitudinal essay, attitude change, replication, multilabs

Received 6/5/20; Revision accepted 10/25/23

Cognitive-dissonance theory (CDT; Festinger, 1957) is one of the most influential theories in the psychological sciences (Devine & Brodish, 2003; Gawronski & Strack, 2012; Haggblom et al., 2002). “Cognitive dissonance” refers to the aversive state caused by the co-occurrence of two or more inconsistent cognitions that motivates individuals to find ways to relieve the discomfort (Harmon-Jones, 2019). It has mainly been studied through observing attitude change following an inconsistency between an attitude and a behavior. CDT has frequently been the topic of review articles and books (e.g., Cooper, 2007; Harmon-Jones, 2019; McGrath, 2017), and a discussion of CDT and its applications can be found in most psychology textbooks (Aronson & Aronson, 2018; Griggs & Christopher, 2016). CDT has been applied to a wide variety of situations, including belief disconfirmation, effort justification, hypocrisy, and decision-making (see Freijy & Kothe, 2013).

The recent replication crisis in the psychological sciences has raised several methodological issues,

including low statistical power (Maxwell et al., 2015) and researcher degrees of freedom (Simmons et al., 2011), that call into question the credibility of past findings. These issues also apply to classical studies in the CDT literature and raise concerns about the replicability of prior findings (Vaidis & Bran, 2019). We therefore set out to perform a test of CDT by conducting a preregistered multilab replication of one of the most prevalent cognitive-dissonance paradigms: the induced-compliance paradigm.

The Induced-Compliance Paradigm

Although a large number of studies have been conducted to evaluate CDT, one particular paradigm has become the dominant choice for investigating CDT and its underlying mechanisms (Cooper, 2007; Harmon-Jones, 2019). A procedure borrowed from the persuasion field (Janis & King, 1954), the induced-compliance paradigm consists of inducing participants to perform a behavior that

is opposite to that implied by an existing attitude (i.e., counterattitudinal). According to CDT, this procedure creates an aversive state (cognitive dissonance) that participants are motivated to resolve. This resolution could be achieved when the participants change their attitude to be in line with their behavior. In the first studies using this paradigm,¹ participants were asked to perform a counterattitudinal task in exchange for either a small or sizable reward (e.g., Brehm & Cohen, 1962; Festinger & Carlsmith, 1959). The size of the reward provides either a sufficient or insufficient justification that prevents (large reward) or favors (small reward) attitude change. In later studies, researchers shifted from manipulating reward size to manipulating the choice given to participants as a justification for their counterattitudinal behavior (see Linder et al., 1967).

The most commonly used task in the induced-compliance paradigm is the counterattitudinal-essay task (Kim et al., 2014). In this task, participants are induced to write a short essay consisting of several arguments in favor of a position they themselves do not hold (e.g., Brehm & Cohen, 1962; A. R. Cohen et al., 1958). For example, college students may be asked to argue in favor of a tuition fee increase, usually under the guise that a university committee is considering implementing the change. Dissonance is manipulated by varying the way the request stresses the choice given to the participants to write the arguments: Participants are either reminded that their participation is voluntary (high choice) or merely instructed to perform the task (low choice). According to CDT, participants in the high-choice condition are prevented from attributing their behavior to an external justification, leading them to experience more dissonance than the participants in the low-choice condition. After writing the arguments, participants report their attitude toward the essay topic. If participants experienced dissonance, this can be resolved by expressing an attitude in line with their behavior. The typically reported finding has been that participants show a more behavior-consistent attitude after writing the essay in the high-choice condition compared with the low-choice condition (McGrath, 2017).

The induced-compliance paradigm along with the counterattitudinal-essay task have served as the experimental context to test some of the most fundamental CDT premises. Relying on this paradigm, researchers have (a) examined arousal attributions as an underlying process for attitude change (Croyle & Cooper, 1983; Zanna & Cooper, 1974); (b) revealed that self-affirmation reduces attitude change (Steele & Liu, 1983); (c) shown that dissonance is associated with aversive arousal, which is alleviated following attitude change (Elliot & Devine, 1994); and (d) investigated alternative ways to regulate dissonance, such as trivialization (Simon et al., 1995) and the denial of responsibility (Gosling et al., 2006). In the

same vein, the paradigm has been used to test theoretical moderators, such as public exposure (Baumeister & Tice, 1984), aversive consequences (Harmon-Jones et al., 1996), and self-esteem (Stone & Cooper, 2003). The induced-compliance paradigm has also been found to produce the largest effect size (Cohen's $d = 0.81$) compared with several alternative cognitive-dissonance paradigms (Kenworthy et al., 2011). Finally, the paradigm is still used today and serves as a standard method for testing CDT hypotheses (e.g., Cooper & Feldman, 2019; Forstmann & Sagioglou, 2020; Randles et al., 2015).

Despite the popularity of the induced-compliance paradigm, there are compelling reasons to doubt its replicability. The credibility crisis has increased suspicion about numerous established findings in social psychology (Nelson et al., 2018; Pashler & Wagenmakers, 2012), including CDT findings. Research involving the induced-compliance paradigm has often relied on small sample sizes (i.e., around 20 per cell; e.g., Elliot & Devine, 1994; Linder et al., 1967; Simon et al., 1995; Steele & Liu, 1983; Zanna & Cooper, 1974; but for an exception, see Murray et al., 2012) and has produced large effect sizes in several seminal studies (e.g., $d > 1.5$; Elliot & Devine, 1994; Simon et al., 1995). Although none of these issues necessarily undermine the paradigm or the theory the paradigm seeks to support, they do question the credibility of existing evidence and highlight the importance of conducting a high-powered, preregistered, multilab replication study that addresses these issues.

The Replication Project

An important challenge in setting up a replication of the induced-compliance paradigm is the lack of a single canonical study. Studies in which the induced-compliance paradigm has been used differ in many ways, including the dissonance-inducing task, attitude assessment, and control conditions. In addition, many of these studies suffer from methodological limitations (Vaidis & Bran, 2019). Consequently, we argue that instead of conducting a direct replication, it is better to conduct a constructive replication—a replication study that is close to a seminal study but also includes new elements to address important limitations (Hüffmeier et al., 2016; Nosek & Errington, 2020). After careful consideration, we selected a study performed by Croyle and Cooper (1983, Experiment 1). This study consists of a typical implementation of the induced-compliance paradigm with the addition of several methodological refinements not typically found in this paradigm, such as the inclusion of a pre- and postessay attitude measure and multiple control conditions. We therefore based our constructive replication of the induced-compliance paradigm on this study.

In Croyle and Cooper's (1983) Experiment 1, participants were asked to write an essay on the topic of

implementing an alcohol ban on campus. Thirty participants were randomly assigned to one of three conditions: a high-choice counterattitudinal condition, a high-choice pro-attitudinal condition (i.e., consonant; in which participants wrote arguments against implementing the ban), and a low-choice counterattitudinal condition. The participants' attitude toward the essay topic was first assessed 1 to 3 weeks before the experiment, and only participants who were against the alcohol ban participated in the study. Participants' attitudes were assessed a second time immediately after the essay. Consistent with CDT, participants in the high-choice counterattitudinal condition demonstrated greater attitude change compared with the two remaining conditions. Note, however, that the study consisted of a small sample ($n = 10$ per analysis cell), and the analyses revealed a disproportionately large effect size ($d = 2.40$, 95% confidence interval [CI] = [1.40, 3.37]²) compared with the typical effect size in social psychology ($d = 0.43$, Richard et al., 2003; $d = 0.30$, Schäfer & Schwarz, 2019) and effect sizes found in meta-analyses on the induced-compliance paradigm ($d = 0.81$, Kenworthy et al., 2011; $d = 0.51$,³ Kim et al., 2014). Despite these results, Croyle and Cooper performed a typical implementation of the induced-compliance paradigm that matches those in seminal studies in the CDT literature (e.g., Elliot & Devine, 1994; Gawronski & Strack, 2004; Zanna & Cooper, 1974). We therefore could expect a successful replication but with an effect size smaller than what was found in the original study.

One of the reasons we chose to replicate Experiment 1 of Croyle and Cooper (1983) is that they assessed attitude change rather than relying solely on a postessay attitude assessment. This feature is important because a key hypothesis of Festinger's (1957) theory is that participants will demonstrate a change in attitude following an attitude-relevant discrepancy. As a consequence, the proper assessment of a participant's response to dissonance requires a prior assessment of the attitude in question. However, researchers in the CDT literature have frequently used between-subjects comparisons to infer attitude change from group differences rather than measuring attitude change intrapersonally. Although group differences are informative, they preclude estimates of change within and across groups. Although unlikely, group differences among conditions in a counterattitudinal advocacy study may arise from reactance in the control condition (i.e., even more negative attitudes) rather than dissonance reduction in the experimental conditions. Consequently, a pre-essay attitude assessment is needed to directly test the core hypothesis of CDT. Having this measure provides several other advantages, including a likely increase in statistical power and the opportunity to test moderating factors, such as the size of attitude-behavior discrepancy.

A second reason to replicate Experiment 1 of Croyle and Cooper (1983) is that they included two control conditions. Conceptually, CDT relies on inconsistency to produce a change in the relevant attitude; however, operationally, researchers have mainly used choice to produce a state of dissonance (Vaidis & Bran, 2019). The induced-compliance paradigm typically requires participants to write a dissonance-inducing counterattitudinal essay under high- or low-choice conditions. Researchers assume that participants in the high-choice conditions attribute their inconsistent behavior to their own choice and, therefore, must resolve the inconsistency by changing their attitudes. In contrast, participants in the low-choice conditions can resolve their potential dissonance by attributing the discrepant behavior to their lack of choice (i.e., an external attribution). Therefore, this design employs a control condition (the low-choice condition) in which dissonance may be experienced but is quickly resolved by participants externally justifying their behavior. In other words, the control condition is not likely to be entirely devoid of a dissonance state. This questionable choice of control conditions raised early critiques (e.g., Chapanis & Chapanis, 1964; Kiesler, 1971; Wicklund & Brehm, 1976) that encouraged control conditions that separately control for both inconsistency and choice. Croyle and Cooper followed these suggestions and used three conditions. In the experimental condition, participants wrote a counterattitudinal essay under high choice. In the first control condition, participants wrote the counterattitudinal essay under low choice, and in their second control condition, participants wrote a pro-attitudinal essay (under high choice). The counterattitudinal high-choice condition was compared with both a counterattitudinal low-choice condition and a pro-attitudinal high-choice condition. This design enabled them to control for an effect of choice (low vs. high) and essay position (counterattitudinal vs. pro-attitudinal). Although we believe that the pro-attitudinal control condition has certain limitations, the inclusion of multiple control conditions further justifies this choice of study for the replication.

Deviations From Croyle and Cooper (1983, Experiment 1)

The main goal of this project was to conduct a compelling test of the induced-compliance paradigm based on Experiment 1 of Croyle and Cooper (1983). Rather than performing a direct replication of this study, we conducted a constructive replication that also addresses previously raised limitations of the induced-compliance paradigm (Vaidis & Bran, 2019). In this project, we focused on the following issues: (a) attitude assessment, (b) the operationalization of dissonance, and (c) threats to internal validity.

Attitude assessment

Induced-compliance studies vary substantially in their attitude-assessment methods. These variations involve participant selection, which essay topic is used, and how the pre- and postessay attitudes are measured.

In CDT studies, within-subjects designs are sometimes implemented by assessing participants' attitude several weeks before the study (e.g., Elliot & Devine, 1994; Steele & Liu, 1983), usually at the start of the student semester. For instance, Croyle and Cooper (1983) recruited participants who completed an attitude questionnaire 1 to 3 weeks before the experiment. This pretest enabled them to assess attitude change and to recruit only participants who disagreed with the essay topic. In the current project, we similarly strove to include an attitude assessment at least 1 week before the main study. The main reason for this attitude assessment is to obtain a measure of attitude change. However, unlike Croyle and Cooper (1983), we did not use the initial attitude assessment to recruit only counterattitudinal participants. This deviation was chosen for several reasons. First, not all participating labs can include an attitude assessment several weeks before the study, making it difficult to compare the findings between labs if some labs used a selection criterion and others did not. Second, many induced-compliance studies do not include an initial attitude assessment and cannot be directly compared with studies in which it is included. A design in which some labs have an initial attitude assessment and others do not allows us to compare our findings with a larger range of studies. Third, without a selection criterion, there will be more variation in initial attitudes, allowing tests of the size of the attitude discrepancy as a moderating factor. Overall, our deviation from Croyle and Cooper to not recruit only counterattitudinal participants allows for easier implementation and valuable comparison between studies.

There is considerable variation in which essay topics are used in induced-compliance-paradigm studies. Examples include raising tuition fees, prohibiting alcohol, banning campus speakers, and instituting mandatory comprehensive exams. The suitability of the essay topic greatly depends on the context. For instance, the essay topic used by Croyle and Cooper (1983) was the introduction of an alcohol ban on the university campus. This essay topic is not suitable for all labs because some universities already have similar bans in place. Other topics, such as introducing mandatory comprehensive exams, may also not be applicable to all contexts (e.g., situations in which such exams are already required). It is therefore not surprising that there is much variation between studies regarding the essay topic. A related issue with topic selection is the lack of information concerning additional attitude components (i.e., importance, certainty), making it more difficult to determine suitable essay topics.

To address the issue of which essay topic to use, we conducted a pretest with a subset of the participating labs to determine suitable essay topics. We assessed attitude agreement and other attitude components, such as attitude importance, of 15 different topics. This pretest indicated that an increase in tuition fees seemed to be the most suitable essay topic for many labs. We therefore selected an increase in tuition fees as the default topic to be used, although we also left the option open for some labs to use a different topic if they deemed the essay topic to not be applicable to their lab (for more details, see the Method section).

Croyle and Cooper (1983) used a 31-point scale to assess the participants' attitude toward their essay topic. The recommended number of Likert-scale options is still a controversial issue (Eutsler & Lang, 2015; Krosnick & Presser, 2009; Simms et al., 2019) and highly variable in the cognitive-dissonance literature; the number of points ranges from 5 (e.g., Gawronski & Strack, 2004) to 61 (e.g., Schaffer, 1974; also see Vaidis & Bran, 2019). In the current study, we aimed to strike a balance between validity and sensitivity. A small range of answers could reduce the required variation of the expected effect, and too large a range could reduce the validity of the provided answers (Matell & Jacoby, 1971). We ultimately opted for a 9-point Likert scale to balance these requirements. This scale should permit enough room for variation to obtain differences between the conditions, and each option can be meaningfully labeled to improve the validity of the scale. A 9-point Likert scale has also been used in multiple other induced-compliance studies (e.g., Harmon-Jones et al., 1996; Simon et al., 1995; Starzyk et al., 2009). In addition, Croyle and Cooper assessed the postessay attitude with a single item. To achieve greater reliability in the attitude assessment, we used one main item to assess the attitude but also added three additional items (based on Stalder & Baron, 1998).

Improved control condition

A strength of Croyle and Cooper's (1983) design is the inclusion of two control conditions. Classically, the counterattitudinal low-choice condition is used as a control condition. As noted before, however, it cannot be considered as a condition in which no cognitive dissonance is experienced. Writing a counterattitudinal essay under low choice still involves an inconsistency; thus, it has to be considered as a control for the effect of choice (high vs. low) on attitude change. An additional control condition is warranted that does not involve an inconsistency at all. Croyle and Cooper included a control condition that indeed is not expected to result in any cognitive dissonance—a consonant high-choice condition. However, because writing in favor of one's attitude may strengthen it (Briñol et al., 2012), we believe that this control condition can be improved by asking participants

to write a neutral essay rather than a pro-attitudinal essay. In this case, no inconsistency is involved, nor do we expect it to change the attitude toward the essay topic of the experimental condition.

Experimenter interaction

A final difference with Croyle and Cooper (1983) lies in the social interaction between the experimenter and participant. In the traditional induced-compliance paradigm, the participant gives the experimenter a verbal agreement to write the essay. This act is assumed to induce the individual to self-attribute the responsibility of writing the essay (Beauvois & Joule, 1996; Kiesler, 1971; McGrath, 2017; Zanna & Cooper, 1974), which may be necessary for the individual to regulate the dissonance state through attitude change (Vaidis & Bran, 2018).

An interaction between the participant and experimenter regarding the participant's choice to write the essay has the unwanted consequence of making the experimenter aware of the participant's experimental condition (i.e., the design is "single blind"). From a methodological point of view, this knowledge allows for experimenter expectations to affect the results (Orne, 1962). It is preferable to use a double-blind procedure instead (for an example in the case of a 2×2 factorial design, see Linder et al., 1967, Experiment 2). One way to achieve a double-blind design is to use a computer-mediated procedure (e.g., Losch & Cacioppo, 1990). However, a fully computer-mediated procedure could result in high numbers of participants refusing to write the essay (about 25% attrition in the case of Losch & Cacioppo, 1990). Thus, including a social interaction seems important to induce compliance in high-choice-condition participants.

A constructive replication therefore requires a double-blind procedure to avoid any influence of the experimenter but should also include a social interaction between the participant and experimenter regarding the participant's decision to complete the essay task. In our design, we aimed to respect both requirements. We relied on a computer-mediated procedure that minimizes the experimenter-participant interaction, thereby reducing possible demand effects, while still including an interaction to promote compliance. In addition, the experimenter was kept blind to a subset of the conditions. The experimenter was inevitably aware of whether the participant was in the low-choice or high-choice conditions (because of the social interaction) but remained unaware of whether the participant had to write a neutral or counterattitudinal essay. This procedure was a compromise between reducing the risk of attrition in high-choice conditions and limiting experimenter demands.

Summary of the Registered Replication Report

In the previous sections, we have shown that there is considerable variation in methodological practices across induced-compliance studies and that a proper assessment of the attitude change in the most used cognitive-dissonance paradigm is limited by methodological shortcomings. As a result, a direct replication of any individual previous study cannot provide a strong test for the induced-compliance paradigm. We therefore developed a constructive replication (Hüffmeier et al., 2016) based on Experiment 1 of Croyle and Cooper (1983) that addressed its methodological limitations. With this study, we aimed to assess the replicability of the induced-compliance paradigm and its findings.

Disclosures

The project was a Registered Report. The design, measures, manipulations, sample size, and data exclusions were approved before data collection. Analyses indicated as primary and secondary were preregistered and can be found in the Stage 1 accepted manuscript. All relevant files, including the Stage 1 manuscript, materials, data, ethical approval documents, and code, are publicly available on OSF at <https://osf.io/9xsmj/>.

Method

Design, sample size, and participating laboratories

Design. The induced-compliance paradigm with an essay-writing task was used to examine whether a state of cognitive dissonance can produce a change in attitude toward a new university policy. The experience of dissonance was manipulated with three conditions, two of which serve as control conditions. These control groups were set to separately assess the effect of choice (high choice vs. low choice in writing the essay) and inconsistency (writing a counterattitudinal essay vs. a neutral essay). For most labs, the attitude of the participants toward the university policy was assessed twice: several weeks before the study and again during the main study. Some labs were unable to recruit the same participants twice for the study and included only a postessay attitude assessment. The study therefore mainly consists of a 2 (within subjects: Pre-essay Attitude Assessment vs. Postessay Attitude Assessment) \times 3 (between-subjects: High-Choice Counterattitudinal Essay [HC-CE], Low-Choice Counterattitudinal Essay [LC-CE], High-Choice Neutral Essay [HC-NE]) mixed design; 37 labs included a pre-essay attitude assessment, and two did not.

Table 1. Overview of the Study's Design, Including Sample Size per Condition

No.	Choice	Essay	Label	Purpose	<i>N</i> (pre-post)	<i>N</i> (total)
1	High	Counterattitudinal	HC-CE	Experimental condition	1,329	1,891
2	Low	Counterattitudinal	LC-CE	Choice control	1,185	1,725
3	High	Neutral	HC-NE	Inconsistency control	889	1,282

Note: The *N* (pre-post) column indicates the sample size of participants who completed both the pre-essay and postessay attitude assessment. The *N* (total) column indicates the total sample size, including participants who were unable to complete the pre-essay attitude assessment.

Sample size. We determined the minimum required sample size based on simulated data. Our goal was to obtain an overall power of 95% for the primary analyses (for details, see the Power Analysis section below). In short, the power analysis revealed that a minimum of 1,760 participants was needed. Because no laboratory would be able to reach this requirement alone and because the goal of the registered replication report is not to determine whether each individual lab obtains a statistically significant result, we did not set an identical required sample size per lab. Instead, each lab was free to recruit as many participants as possible based on available resources, which resulted in a total of 4,898 participants (see Table 1).

Participating labs and lab recruitment. A total of 39 labs from 19 countries and 14 languages participated in the data collection. Groups of labs joined the replication effort in several phases. The first group of researchers was composed of researchers experienced with CDT studies. A second group was composed of interested researchers via the Society for the Improvement of Psychological Science conference (Sleegers & Vaidis, 2019). A third group was composed following calls for contribution spread through social networks and academic mailing lists. A final fourth group of researchers joined the project following the Stage 1 acceptance of the current article. There were no exclusion criteria for labs to join the project. Initially, 23 additional labs including seven other countries joined the project but withdrew because of unforeseen challenges in data collection (in particular, recruiting participants for in-person data collection during the COVID-19 pandemic).

When translations were required, labs were asked to translate the materials and back-translate them to ensure accuracy of the translations. The lead authors (D. Vaidis and W. Sleegers) coordinated this effort to ensure all labs used similar materials. Labs were also instructed to remain blind to the results until data collection was completed. The expectations of lab members regarding the replication results and effect sizes were measured before data collection (see Additional Exploratory Analyses Plan section). Each lab received a training video of the procedure. They were also required to take photos or record a video of their lab and to make their study materials available on OSF.

The principal investigator from each lab signed a co-laborator agreement indicating the ethical guidelines and the preregistered procedure would be followed. For an overview of the lab characteristics, see Table 2.

Population and participant recruitment. As in most studies using the induced-compliance paradigm, participants were university students. Labs were encouraged to recruit first-year students to reduce the likelihood that participants were familiar with CDT. The study was advertised to students as a series of tasks, some concerning students' opinions. Varying incentives were used to recruit participants: They took part in the study in exchange for course credits, monetary incentives, or no incentive. The labs documented their use of incentives because this could affect the attitude change (e.g., Festinger & Carlsmith, 1959; Kim et al., 2014). Lab characteristics (e.g., type of room, incentive) are available in the supplemental material on OSF for use in exploratory analyses.

Selection of the essay topic

Many different essay topics have been used in the CDT literature, although a popular topic is an increase in tuition fees (e.g., Elliot & Devine, 1994; Steele & Liu, 1983). We used a pretest to determine suitable essay topics from a subset of data-collection sites.⁴ The pretest assessed 15 hypothetical educational-policy changes (e.g., [University name] should raise the tuition fee for the upcoming academic year) on four dimensions: agreement (1 = *strongly disagree*, 9 = *strongly agree*), importance (1 = *not important at all*, 9 = *very important*), preference to write arguments against or in favor of the topic (*in favor/against/indifferent*), and plausibility of implementation (*likely/unlikely*).

The pretest provided participating labs with a set of potential essay topics to choose from. On the basis of the results of the pretest, we found that the increase in tuition fees essay topic was a suitable topic in most of the labs and also showed the least amount of variance in terms of attitude agreement. We therefore set the increase in tuition fees as the default topic to be used. However, because the topic would not be suitable for all participating labs (e.g., because of the university not

Table 2. Sample Size, Language, and Topic Used per Lab University

University	Country, language	Topic used	Pre-post design
Université Paris Cité	FR, French	Tuition fee	Yes
Tilburg University	NL, Dutch, English	Tuition fee	Yes
Université Paul Valéry Montpellier 3	FR, French	Tuition fee	Yes
Appalachian State University	USA, English	Tuition fee	No
University of Amsterdam	NL, Dutch	Tuition fee	Yes
Université Caen Normandie	FR, French	Tuition fee	Yes
Aix-Marseille Université	FR, French	Tuition fee	Yes
University of Arizona	USA, English	Tuition fee	Yes
University at Buffalo	USA, English	Tuition fee	Yes
King's College London	UK, English	Tuition fee	No
University of Bergen	NOR, Norwegian	Harder exam	Yes
Higher School of Economics University	RU, Russian	Tuition fee	Yes
Central Connecticut State University	USA, English	Tuition fee	Yes
University of Granada	ES, Spanish	Tuition fee	Yes
Southern Illinois University	USA, English	Tuition fee	Yes
University of Toronto	CAN, English	Tuition fee	Yes
Université Grenoble Alpes	FR, French	Tuition fee	Yes
Université de Poitiers	FR, French	Tuition fee	Yes
Université de Genève	CH, French	Tuition fee	Yes
Clemson University	USA, English	Tuition fee	Yes
Université Clermont Auvergne	FR, French	Tuition fee	Yes
Üsküdar University	TUR, Turkish	Tuition fee	Yes
Université Rennes 2	FR, French	Tuition fee	Yes
SWPS University of Social Sciences and Humanities	PL, Polish	Tuition fee	Yes
Université de Strasbourg	FR, French	Tuition fee	Yes
Macquarie University	AUS, English	Tuition fee	Yes
Sungkyunkwan University	KOR, Korean	Tuition fee	Yes
Kyushu University	JAP, Japanese	Tuition fee	Yes
Muş Alparslan University	TUR, Turkish	Tuition fee	Yes
Université Libre de Bruxelles	BE, French	Tuition fee	Yes
University of Graz	AUT, German	Tuition fee	Yes
Ben-Gurion University	ISR, Hebrew	Tuition fee	Yes
Ural Federal University	RU, Russian	Tuition fee	Yes
Sultan Qaboos University	OMN, Arabic	Tuition fee	Yes
Florida State University	USA, English	Tuition fee	Yes
Sabancı University	TUR, Turkish	Tuition fee	Yes
Shahid Beheshti University	IRN, Persian	6 a.m. classes	Yes
Erasmus University Rotterdam	NL, English	Tuition fee	Yes
Nara Institute of Science and Technology	JAP, Japanese	Tuition fee	Yes

Note: Labs are chronologically ordered based on when they joined the project.

having any tuition fees), two labs used a different topic (see Table 2). We defined a topic as suitable when (a) pretest participants reported a median attitude score lower than 5 on the agreement dimension, (b) pretest participants reported a median attitude score greater than 5 on the importance dimension, and (c) at least 80% of the sample preferred writing against this topic. The pretest showed that plausibility ratings were often low, so rather than using a cutoff score, we used the plausibility ratings to select the most suitable topic (i.e., with the highest plausibility rating). In the following sections, we use the tuition-fees topic as an example to

describe the materials and refer to the OSF page for the exact materials that each lab used.

For labs that included the pre-essay attitude assessment ($k = 37$), participants completed a survey at least 1 week before the main study ($M = 2.28$ weeks, $Mdn = 1.85$).⁵ The pre-essay measure was chiefly performed online, with an anonymous code to link the data between the survey and the main study. Among other filler tasks, participants responded to a short questionnaire about student topics, including the essay topic from the main study. For each topic, attitudes were assessed on both an agreement and an importance dimension using a

9-point-scale from 1 (*strongly disagree or not important at all*) to 9 (*strongly agree or very important*).

Procedure

The main study was performed in lab rooms in isolated or semi-isolated cubicles. The study was conducted on a computer using the Qualtrics survey platform. The study was presented as a set of smaller studies, starting with a neutral filler task that was alleged to be the main study. After the filler task, participants completed the essay task.

Experimenters were trained to follow a predefined script, including a business-casual dress code (for details, see OSF). The experimenter guided each participant to the computer but remained at a distance while participants completed the study, except for a specific moment during the study during which participants contacted the experimenter to start the essay task. All participants within a particular session were randomly assigned to one of the three conditions. In the low-choice condition, the experimenter instructed participants to continue the essay task. In the high-choice conditions, the experimenter asked participants to give their consent to write the essay and proceed to the essay task. This interaction was designed to keep the experimenter blind to whether the participant had been assigned a counterattitudinal or neutral topic. This procedure guaranteed that (a) participants had to engage in a social interaction, (b) experimenters remained blind to the participant's essay content in the high-choice conditions, and (c) experimenters had a limited opportunity to influence participants.

Essay instructions

All participants were introduced to the essay task with a cover story that claimed this part of the study was about a survey concerning student life that had been requested by the university. The essay instructions were based on Croyle and Cooper (1983) while also adapting aspects from other CDT studies that have used a similar procedure (e.g., Elliot & Devine, 1994; Gosling et al., 2006; Losch & Cacioppo, 1990). The task started the same way for all participants (here illustrated with the increase in tuition topic):

[University name] is currently discussing making changes to several university policies. Our department has been asked to assist a committee by collecting information on a number of issues. Because the changes to the policies could impact students currently studying at [University name], we would like your participation in the discussion of these issues.

One of the evaluated policies concerns [University name]'s tuition. More specifically, [University name] has set up a committee on campus to investigate the possibility of a substantial tuition increase. After reviewing what they find, the committee will make a recommendation to the administration regarding the tuition increase that could be implemented at [University name] in the next academic year.

In the HC-CE and LC-CE conditions, participants were then instructed to write arguments counter to their attitude (e.g., in favor of a tuition-fee increase). For them, the complementary instruction was as follows:

Before making a recommendation, however, the committee would like to gather arguments on both sides of the issue. We have found that a good way of doing this is simply to ask people to list all the arguments they can think of that support a particular side of the issue. In other words, whatever one's personal opinion, we want them to write arguments either only in favor or only against the increase of tuition.

Because we have already finished gathering arguments against a tuition increase, we are now ready to gather arguments in favor of a tuition increase. We thus want you to write **ONLY** arguments **IN FAVOR** of the increase in tuition.

This was followed by a summary of the instructions on the next page:

To summarize, in the next task, you will have to write **ONLY** arguments **IN FAVOR** of the tuition increase. Your arguments will be sent directly to the committee, who will make a decision on this issue based on the arguments they receive from you and other students.

In the HC-NE control condition, participants were instructed to suggest arguments about additional topics that they think should be investigated by a future committee that is currently being set up:

Beside this topic, the committee would also like to gather suggestions for additional educational policy changes that require the attention of future committees. We would like you to write arguments in favor of specific topics that should be investigated by future committees.

This was followed by a summary of the instructions on the next page:

To summarize, in the next task, you will have to write arguments for additional topics that should be investigated by future committees. Your arguments will be sent directly to the current committee, who will make a decision on this issue based on the arguments they receive from you and other students.

After the essay instructions, participants were instructed to contact the experimenter. When called, the experimenter approached the participant, keeping a neutral face (no smiling), and asked the participant (with a neutral tone), "Did you understand the instructions?" The experimenter waited for the participant's agreement (verbal or nonverbal) before continuing. To ensure that the experimenter remained blind to the essay content, this interaction did not reveal what essay the participant had been instructed to write.

In the low-choice condition, the experimenter simply instructed the participant to begin the task. In the high-choice conditions, the experimenter reminded the participant about the voluntary nature of the task:

Because this next task is part of a research project, we want to remind you that your participation is completely voluntary. We would appreciate your help, but we do want to let you know that it's completely up to you.⁶

The experimenter then asked participants whether they were willing to perform the task. After the experimenter received verbal agreement, the participant received a form that served the purpose of reinforcing the choice manipulation. The form stated,

I understand the nature of the task I am being asked to perform. I am aware that my list of arguments is intended for use by the University committee. I further understand that I will receive [compensation/my participation credit, if applicable] regardless of whether or not I write, and allow release of, my list of arguments.

Participants were asked to sign and date the form and check one of two boxes to indicate whether they allowed their arguments to be released to the committee.

If the participant asked any clarification questions, the experimenter answered them following a predefined script. If participants questioned the cover story, experimenters explained that they are only in charge of the session and are unable to give further details. Participants were then asked to follow the instructions as best as they can. If the participant expressed doubts about willingness to perform the task, the experimenter

repeated the instructions once more (respectively, in high-choice and low-choice conditions: "Your participation is voluntary. It's completely up to you." vs. "We would like you to complete the task."). Finally, if participants refused to perform the essay task, they were asked to skip the essay task and continue the study. If they refused this as well, the study was stopped.

Next, participants were presented with instructions for writing their arguments. They were reminded of the previous instruction to write in favor of the policy or to provide additional topics to be investigated and were instructed to write three short paragraphs including their argumentation. They were informed they had a maximum of 5 min to complete the task. On the next page, three different text areas required keyboard input from the participant. A reminder of the remaining time was displayed after 3 min. A second message was displayed when 30 s remained, which informed participants that their responses would soon be automatically submitted. The submit button was active only after 3 min, and the page was automatically submitted after 5 min.

Dependent measures

Postessay attitude assessment. The postessay attitude was assessed directly after the completion of the essay task.⁷ Regardless of the condition, participants indicated their attitude toward the main policy on four items. The first attitude measurement was identical to the pre-essay attitude measure (e.g., "[University name] should raise the tuition fee for the upcoming academic year") and was used to assess attitude change. Three additional attitude items were adapted from Stalder and Baron (1998) and were used to assess a broader attitude: "How would you describe your overall opinion toward [raising tuition]?" (1 = *extremely unfavorable*, 9 = *extremely favorable*); "To what extent do you think there are disadvantages to [raising tuition]?" (1 = *no disadvantages*, 9 = *a great many*; reverse-scored); and "To what extent do you think that [raising tuition] is good for [University name]?" (1 = *not at all*, 9 = *a great deal*).

At the end of the procedure, after the attitude assessment, participants were asked to complete a measure of perceived choice, a measure of affect, and a self-construal scale. These questions were presented under the cover story that the previously mentioned committee was interested in these questions.

Perceived choice. Because perceived choice has often been used as a manipulation check in dissonance studies, we first included this perception of choice measure. Participants were asked to indicate their perceived degree of choice to perform the essay task on two 9-point scales (1 = *no choice at all*, 9 = *totally free to choose*): "How free

did you feel to decline to write the essay?" (Croyle & Cooper, 1983), and "How much choice did you feel you had to participate in this study?" (Cooper & Mackie, 1983). The first item refers directly to the freedom to decline writing the essay and was therefore used as a manipulation check. The second item was included for exploratory purposes only and is not analyzed in this article.

Affect and self-construal. In addition to assessing perceived choice, we also included additional measures for exploratory purposes. We included an adapted version of the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) to assess the participant's affective experience while performing the essay task. We chose the PANAS because it has already been translated into several languages that are used in this study. Participants were instructed to indicate how they felt while writing the essay task on a 5-point scale (1 = *very slightly or not at all*, 5 = *extremely*). Two items, uncomfortable and conflicted, were added to specifically assess a dissonance state (also see Devine et al., 2019).

Finally, we added a questionnaire to assess self-construal (20 items; Park & Kitayama, 2014) because some perspectives on dissonance consider the self to be an important construct related to the experience of dissonance (Aronson, 1992, 2019; Steele & Liu, 1983; Stone & Cooper, 2003). The research on self-construal suggests that the dissonance effect could be experienced differently depending on one's culture (Markus & Kitayama, 1991). More specifically, in the free-choice paradigm (Brehm, 1956), results suggest that interdependent self-construal participants rationalize their choices less than independent self-construal participants (e.g., Heine & Lehman, 1997; Kitayama et al., 2004). To our knowledge, no study has evaluated the role of self-construal in the induced-compliance paradigm. Because this multilab study involves participants from a large number of cultures, we considered it an opportune situation to examine the impact of self-construal in this paradigm.

Demographics and funnel debriefing

After completing the steps of the experiment listed above, participants answered several demographic questions: age, sex, academic degree and major, fluent languages, and native country. Two questions concerning the participant's university affiliation were used as exclusion criteria: "What is your current university?" and "Do you plan to study at the same university next year?"

At the end of the study, participants answered several questions, presented one by one, that served as a funnel debriefing: "Do you have any comments?"; "What do you think is the purpose of the study?"; "Do you have any psychological phenomena or theories in mind that you

think this study is about? If so, which one(s)?" "If the goal of the study was not the one we presented, would you have an idea about the reason why we asked you to write down arguments?"; "Have you ever heard about cognitive dissonance theory? If so, please briefly describe the theory."; and "Have you ever participated in a similar study where you were asked to provide arguments for a particular position?" After answering these questions, participants were fully debriefed by the experimenter and thanked for their participation.

Exclusion criteria

Participant characteristics. Our registered plan was to exclude participants that were not students at the time of the study or that did not plan to be students at the same university in the following year. Participants were also required to be unaware of the study's purpose. We therefore registered a plan for three raters, blind to the conditions, to categorize each participant as being aware or unaware of the study's purpose based on their responses to the funnel-debriefing questions. Participants were excluded from the analysis if at least two of the raters categorized them as being aware of the study's purpose.

Essay content. We had no registered plan to exclude participants based on their essay response. Traditionally, participants who did not comply with the essay instructions are excluded from the analyses; their data are removed if they gave at least one pro-attitudinal or no counterattitudinal argument. This attrition is strongly conditional and can comprise around 20% of a counterattitudinal condition (e.g., Azdia & Joule, 2001; Losch & Cacioppo, 1990; Simon et al., 1995, Experiment 1; see also Chapanis & Chapanis, 1964). Although we agree that this exclusion is conceptually in accordance with the idea of manipulating dissonance, this is a potential methodological limitation that introduces an important bias into the experimental design, resulting in an increased likelihood of a false positive (Ranganathan et al., 2016; Zhou & Fishbach, 2016). Participants with the strongest attitudes are more likely to be the ones to defy the instructions and also less likely to change their attitude. Excluding these participants leaves participants who are more likely to display attitude change in the expected direction. We thus decided to keep all the participants in the primary analyses (i.e., an intention-to-treat analysis). However, to be consistent with prior studies, we also performed secondary analyses in which participants who refused to write the essay or who were rated to be nonconforming were excluded (i.e., per protocol analysis). Each argument was evaluated by three raters blind to the choice instructions. They categorized participants as not compliant when they wrote at least one pro-attitude argument or no

counterattitudinal argument in the counterattitudinal conditions or no arguments in the neutral-essay condition. Participants were excluded when at least two raters categorized the participant as noncompliant. In addition, the raters evaluated the written arguments in HC-CE and LC-CE conditions and rated their quality (reported variable persuasiveness; 1 = *poorly persuasive*, 4 = *very persuasive*). These results were recorded to allow them to be used for potential exploratory analyses not included here.

Session characteristics. Each lab kept a log with anonymized participant numbers and reported any unusual events that occurred during the sessions. Participants who were disturbed during the data-collection session for any specific reason (e.g., phone interruption, alarm, computer crash) were excluded.

Lab characteristics. Out of concerns for data quality, labs were excluded from data analysis if they experienced 25% or more attrition because of session characteristics or participant noncompliance. The number of participants tested and attrition rates are reported in Table 3.

Data analysis

The analyses were divided into a manipulation check, primary analyses, and secondary analyses. The main analyses were performed on the aggregated data across all samples and involved four one-tailed Welch's *t* tests (Delacre et al., 2017) comparing the HC-CE condition with the two control conditions in both postessay attitude only and attitude change. The secondary analyses repeated the primary analyses on a data set in which participants who defied the essay instructions had been excluded. We also inspected lab variability in terms of the main analyses and on a multi-item assessment of the postessay attitude. Additional exploratory analyses regarding two dissonance-related affect items are also provided in the present article to further interpret the results.

Power analysis

Our goal was to obtain an overall power of 95% for the four primary analyses (see Maxwell, 2004). To determine the required sample size, we ran a simulation-based power analysis in which we repeatedly simulated the data necessary to conduct the primary (one-tailed) analyses. We then counted how frequently we observed a significant effect in all four analyses and divided it by the total number of simulations to assess the overall power. Despite large effect sizes in the CDT literature (Kenworthy et al., 2011), we assumed two small effect sizes (J. Cohen, 1988) to be conservative: an effect size

of $d = 0.20$ for the difference between the HC-CE and LC-CE conditions and a slightly larger effect size of $d = 0.25$ for the difference between HC-CE and HC-NE. An additional required parameter for the power analysis is the correlation between the pre- and postessay attitudes. Because no information on this correlation was available, we calculated the power assuming a correlation of .9 in the HC-NE condition, .8 in the LC-CE condition, and .7 in the HC-CE condition, based on the expectation that an effect of the manipulation reduces the correlation between the pre- and postessay attitudes. Using the *GenOrd* package (Ferrari & Barbiero, 2012), we simulated Likert responses for each condition with the specified correlations and effect sizes across a range of different sample sizes per condition. The results showed that a sample size of 1,760 provided 95% power, with 660 participants in the HC-CE and LC-CE conditions and 440 participants in the HC-NE condition.

Results

Sample size and exclusions

A total of 39 laboratories contributed to the project, with a grand total of 4,898 recruited participants. Table 3 presents the sample size and exclusions for each laboratory. Following our exclusion criteria, 902 were excluded based on participant characteristics, and 112 were excluded based on session characteristics (with some overlap). We also excluded the data from one lab because its attrition rate was greater than the preregistered attrition limit. One of the control conditions for one lab was also discarded because of an error in the study materials. This resulted in a total of 38 labs and total sample size of 3,822 for the primary analyses involving the postessay attitude and a total of 37 labs and 2,724 participants for the primary analyses involving attitude change. For one of the secondary analyses, regarding essay noncompliance, we excluded an additional 28 participants who refused to complete the essay task and 362 nonconforming participants who failed to follow the essay-task instructions.

Manipulation check

To see whether participants indeed experienced greater choice freedom in the high-choice conditions, we conducted two one-tailed Welch's *t* tests comparing the HC-CE condition and the HC-NE condition with the LC-CE condition. The two *t* tests showed that the choice manipulation was successful. Compared with the LC-CE condition ($M = 4.44$, $SD = 2.78$), participants felt more free to decline writing the essay in both the HC-CE condition ($M = 6.50$, $SD = 2.62$), $t(2,740.40) = 20.18$, $p < .001$,

Table 3. Exclusions and Attrition Rate per Lab University

University	Recruited	Participant characteristic exclusions			Session characteristic exclusions	Essay refusals	Attrition rate	Essay content exclusions
		Not a student next year	Aware of study purpose	N after initial exclusions				
Université Paris Cité	237	16	23	198	8	1	5%	14
Tilburg University	176	3	24	149	2	3	3%	6
Université Paul Valéry Montpellier 3	105	13	4	89	2	4	7%	13
Appalachian State University	213	20	6	187	1	5	3%	16
University of Amsterdam	117	4	29	88	13	1	15%	8
Université Caen Normandie	169	24	6	141	7	5	6%	26
Aix-Marseille University	211	61	6	144	7	3	6%	49
University of Arizona	225	20	25	182	8	1	5%	24
University of Buffalo	242	19	20	209	11	6	7%	34
King's College London	29	7	1	22	0	0	0%	3
University of Bergen	101	16	5	81	0	0	0%	3
Higher School of Economics University	86	5	2	79	0	0	0%	17
Central Connecticut State University	27	3	2	22	0	0	0%	2
University of Granada	194	6	8	180	8	1	4%	8
Southern Illinois University	117	6	5	106	3	0	3%	13
University of Toronto	131	13	3	115	4	7	9%	19
Université Grenoble Alpes	157	14	9	135	3	2	4%	8
Université de Poitiers	106	12	9	85	2	0	2%	0
Université de Genève	211	9	48	157	4	5	5%	19
Clemson University	127	6	10	112	0	2	2%	4
Université Clermont Auvergne	177	11	20	146	1	1	1%	19
Üsküdar University	107	18	4	86	4	1	5%	15
Université Rennes 2	105	10	33	65	1	1	2%	11
SWPS University of Social Sciences and Humanities	189	7	13	169	3	2	3%	19
Université de Strasbourg	81	3	19	60	1	0	2%	16
Macquarie University	63	2	9	52	0	0	0%	5
Sungkyunkwan University	100	14	4	82	1	7	10%	9
Kyushu University	70	3	2	65	2	0	3%	17
Muş Alparslan University	89	18	0	71	0	0	0%	48
Université Libre de Bruxelles	93	1	4	88	2	1	2%	13
University of Graz	66	10	10	47	0	0	0%	10
Ben-Gurion University	146	31	11	108	0	0	0%	2
Ural Federal University	84	4	2	78	0	1	1%	17
Sultan Qaboos University ^a	153	77	0	76	8	40	55%	58
Florida State University	76	4	4	69	0	0	0%	4
Sabancı University	117	29	4	85	2	2	5%	12
Shahid Beheshti University ^b	89	3	10	76	1	6	9%	37
Erasmus University Rotterdam	50	0	7	43	1	2	7%	8
Nara Institute of Science and Technology	62	3	10	49	2	1	6%	11
	4,898	525	411	3,996	112	111	5%	617

Note. The attrition rate is the percentage of participants who were excluded based on the session characteristics and the participants' refusal to continue the study after the essay task introduction, which may overlap.

^aOne lab was excluded from the analyses because its attrition rate was greater than 25%.

^bData from the high-choice neutral essay condition were not included because of an error in the study materials.

Table 4. Descriptive Statistics by Condition

	High-Choice Counterattitudinal Essay	Low-Choice Counterattitudinal Essay	High-Choice Neutral Essay
Primary analyses			
Perceived choice	6.50 (2.62) <i>n</i> = 1,448	4.44 (2.78) <i>n</i> = 1,344	6.91 (2.36) <i>n</i> = 1,032
Postessay attitude	2.60 (1.91) <i>n</i> = 1,448	2.66 (1.99) <i>n</i> = 1,343	2.14 (1.61) <i>n</i> = 1,031
Pre-essay attitude	1.60 (1.23) <i>n</i> = 1,062	1.66 (1.23) <i>n</i> = 937	1.62 (1.28) <i>n</i> = 727
Attitude change	1.00 (1.81) <i>n</i> = 1,060	0.97 (1.80) <i>n</i> = 937	0.47 (1.49) <i>n</i> = 726
Secondary analyses			
Postessay attitude	2.66 (1.89) <i>n</i> = 1,290	2.63 (1.92) <i>n</i> = 1,263	2.10 (1.56) <i>n</i> = 882
Attitude change	1.05 (1.75) <i>n</i> = 961	0.96 (1.77) <i>n</i> = 892	0.50 (1.47) <i>n</i> = 617
Four-item attitude	3.35 (1.54) <i>n</i> = 1,447	3.40 (1.57) <i>n</i> = 1,343	2.98 (1.35) <i>n</i> = 1,031
Exploratory analyses			
PANAS, uncomfortable	2.13 (1.22) <i>n</i> = 1,449	2.16 (1.26) <i>n</i> = 1,344	1.87 (1.12) <i>n</i> = 1,032
PANAS, conflicted	2.61 (1.36) <i>n</i> = 1,449	2.66 (1.38) <i>n</i> = 1,344	1.89 (1.07) <i>n</i> = 1,032

Note. Higher attitude numbers indicate an attitude more in favor of the policy. Attitude change was calculated by subtracting the pre-essay attitude from the postessay attitude. PANAS = Positive and Negative Affect Schedule.

$d = 0.77$, 95% CI = [0.69, 0.84], and the HC-NE condition ($M = 6.91$, $SD = 2.36$), $t(2,349.50) = 23.45$, $p < .001$, $d = 0.95$, 95% CI = [0.86, 1.04].

Primary analyses

Our main analyses for determining the replicability of the induced-compliance paradigm consisted of four one-tailed Welch's t tests on the aggregated samples. We examined whether the experimental condition (HC-CE) differed significantly from each of both control conditions (LC-CE and HC-NE). First, we analyzed the postessay attitude assessment. Second, we analyzed attitude change by subtracting the pre-essay attitude from the postessay attitude (see Table 4).

Postessay attitude. In our test of the classic cognitive-dissonance effect, we did not find a significant difference in postessay attitude between the HC-CE condition ($M = 2.60$, $SD = 1.91$) and the LC-CE condition ($M = 2.66$, $SD = 1.99$), $t(2,751.46) = -0.79$, $p = .79$, $d = -0.03$, 95% CI = [-0.10, 0.04]. However, we did find a significant difference comparing the postessay attitude between the HC-CE condition and the HC-NE condition, $t(2,408.92) = 6.51$, $p < .001$, $d = 0.26$, 95% CI = [0.18, 0.34]. Participants in the HC-CE condition reported a more positive attitude

than participants in HC-NE condition ($M = 2.14$, $SD = 1.61$).

Attitude change. Because of the inclusion of a pre-essay attitude assessment in an earlier session, we could perform an analysis testing whether the manipulation changed participant attitudes from baseline. As in the postessay attitude analysis, we did not observe a significant difference between the HC-CE condition ($M = 1.00$, $SD = 1.81$) and LC-CE condition ($M = 0.97$, $SD = 1.80$), $t(1,969.56) = 0.38$, $p = .35$, $d = 0.017$, 95% CI = [-0.071, 0.11]. We did again observe a significant difference between the HC-CE condition and HC-NE condition, $t(1,725.20) = 6.72$, $p < .001$, $d = 0.31$, 95% CI = [0.22, 0.41]. Participants in the HC-CE condition reported a more positive attitude change than participants in the HC-NE condition ($M = 0.47$, $SD = 1.49$).⁸

Secondary analyses

Essay-noncompliance exclusion. In the main analyses, we included participants who did not comply with the essay instructions (e.g., refused to write the essay). To allow consistency with standard analysis in the cognitive-dissonance literature, we repeated the main analyses while excluding participants who did not comply with the essay instructions. This led to the exclusion of 159 participants in

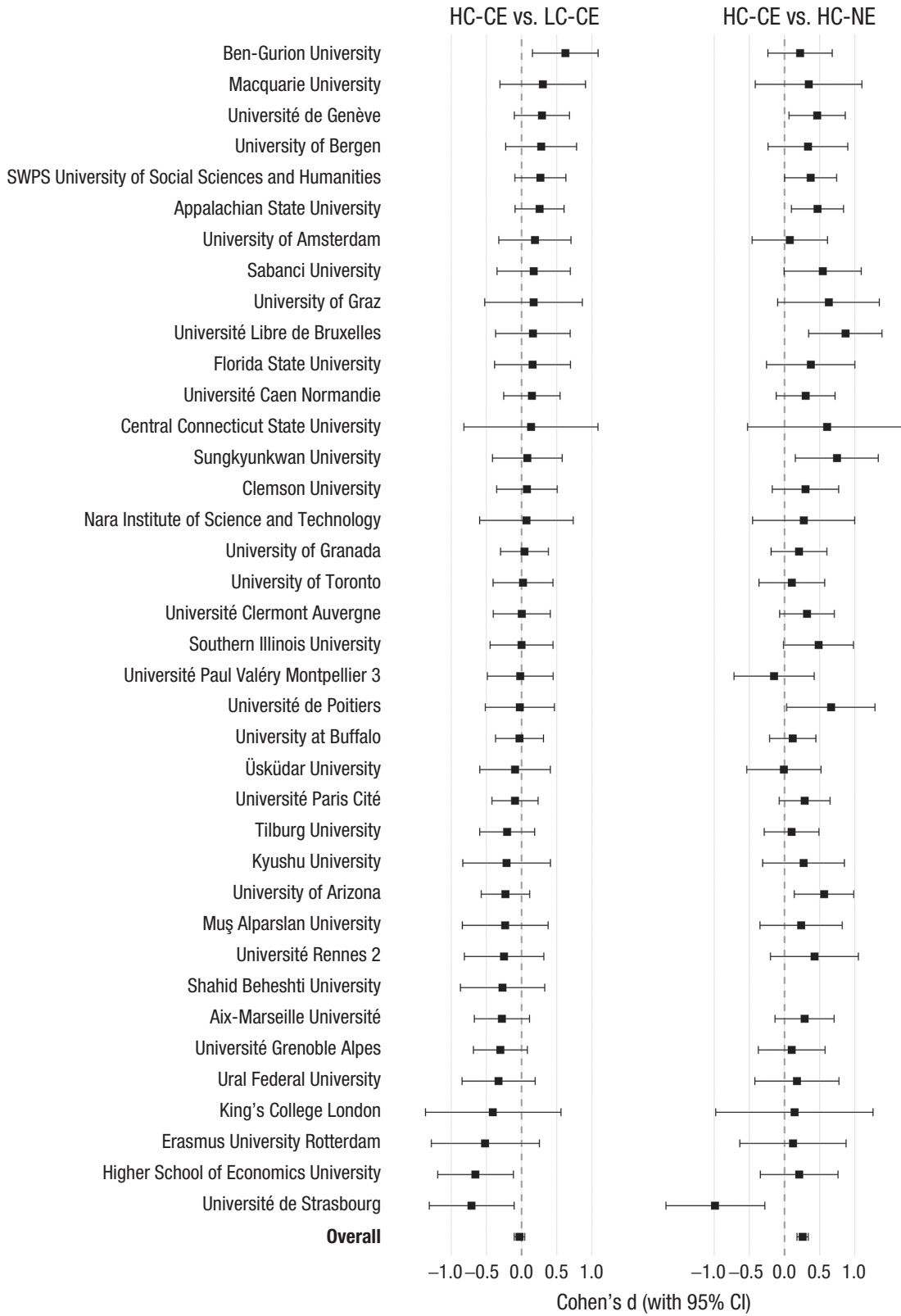


Fig. 1. Overall comparisons for each lab effect sizes (Cohen's *d*) on postessay attitude between high choice and low choice and between counterattitudinal essay and neutral essay. A positive Cohen's *d* represents a more positive attitude mean in the high-choice counterattitudinal essay condition, thus supporting the hypotheses.

the HC-CE condition (11%), 80 in the LC-CE condition (6%), and 151 in the HC-NE condition (15%).

As in the primary analysis, the HC-CE condition ($M = 2.66$, $SD = 1.89$) did not significantly differ from the LC-CE condition ($M = 2.63$, $SD = 1.92$), $t(2,547.56) = 0.38$, $p = .35$, $d = 0.015$, 95% CI = [-0.063, 0.093]. The HC-CE condition again differed significantly from the HC-NE condition, $t(2,095.19) = 7.43$, $p < .001$, $d = 0.31$, 95% CI = [0.23, 0.40]. Participants in the HC-CE condition reported a more positive attitude compared with participants in the HC-NE condition ($M = 2.10$, $SD = 1.56$).

Regarding attitude change, we replicated the same pattern of results. The HC-CE condition ($M = 1.05$, $SD = 1.75$) did not significantly differ from the LC-CE condition ($M = 0.96$, $SD = 1.77$), $t(1,837.21) = 1.04$, $p = .15$, $d = 0.048$, 95% CI = [-0.043, 0.14], but did differ significantly from the HC-NE condition ($M = 0.50$, $SD = 1.47$), $t(1,467.82) = 6.72$, $p < .001$, $d = 0.33$, 95% CI = [0.23, 0.44].

Lab variability. Lab variability was investigated by comparing a linear mixed model with by-lab random intercepts to a linear mixed model with both by-lab random intercepts and by-lab random slopes for the effects. In each model, the first postessay attitude measure was regressed on the experimental conditions, which were included as dummy-coded fixed effects with the HC-CE condition as the reference group. Using analysis of variance, we found no significant difference in model fit when the two models were compared, $\chi^2(5) = 10.19$, $p = .070$. Thus, we did not find statistically significant heterogeneity of the effect of the experimental conditions across the different laboratories (also see Fig. 1).

Four-item attitude assessment. In addition to the single-item measure of attitude analyzed above, attitude was also assessed with three additional items after the initial postessay attitude item for increased reliability. We computed a new attitude score by averaging four items (Cronbach's $\alpha = .87$, McDonald's $\omega = .88$) and repeated the postessay attitude analyses. Again, these analyses showed the same pattern of results. There was no significant difference in attitude between participants in the HC-CE condition ($M = 3.35$, $SD = 1.54$) and participants in the LC-CE condition ($M = 3.40$, $SD = 1.57$), $t(2,761.20) = -0.86$, $p = .81$, $d = -0.033$, 95% CI = [-0.11, 0.042]. There was, again, a significant difference in attitude between participants in the HC-CE condition and HC-NE condition, $t(2,367.28) = 6.37$, $p < .001$, $d = 0.25$, 95% CI = [0.17, 0.33]. Participants in the HC-CE condition reported a more positive attitude compared with participants in the HC-NE condition ($M = 2.98$, $SD = 1.35$). See the supplemental material on OSF for additional analyses.

Exploratory analyses of affect

Neither the primary nor the secondary analyses showed the predicted effect of choice on attitude (i.e., no significant difference between HC-CE and LC-CE conditions) but showed only an effect of writing an attitude-inconsistent essay (i.e., HC-CE condition was significantly greater than HC-NE condition). We therefore decided to run additional analyses to help interpret these results. One important question was whether we could find evidence of subjective conflict or discomfort that would suggest our manipulation successfully induced a cognitive-dissonance state in the counterattitudinal-essay conditions, particularly in the high-choice condition. Participants completed a PANAS, which included two additional items (i.e., uncomfortable and conflicted) to specifically assess a cognitive-dissonance state. We performed two separate Welch's t tests to compare differences on these items between the conditions.

As in the primary analyses, we did not find a significant difference on the uncomfortable item between the HC-CE condition ($M = 2.13$, $SD = 1.22$) and the LC-CE condition ($M = 2.16$, $SD = 1.26$), $t(2,761.87) = -0.72$, $p = .47$, $d = -0.027$, 95% CI = [-0.10, 0.047]. Comparing the uncomfortable feeling in the HC-CE condition and the HC-NE condition, we found a significant effect, $t(2,325.89) = 5.44$, $p < .001$, $d = 0.22$, 95% CI = [0.14, 0.30]. Participants in the HC-CE condition reported being more uncomfortable than participants in HC-NE condition ($M = 1.87$, $SD = 1.12$). The difference between the LC-CE condition and the HC-NE condition was also significant, $t(2,320) = 5.97$, $p < .001$, $d = 0.24$, 95% CI = [0.16, 0.32]. These results suggest that participants in the conditions containing an inconsistency (HC-CE and LC-CE) felt more uncomfortable while writing the essay compared with the neutral-essay condition.

The same pattern was observed on the item about feeling conflicted. There was no significant difference between the HC-CE condition ($M = 2.61$, $SD = 1.36$) and LC-CE condition ($M = 2.66$, $SD = 1.38$), $t(2,765.24) = -1.01$, $p = .31$, $d = 0.038$, 95% CI = [0.11, 0.036]. However, there was a significant difference between the HC-CE condition and the HC-NE condition, $t(2,454.28) = 14.72$, $p < .001$, $d = 0.58$, 95% CI = [0.49, 0.66]. Participants in the HC-CE condition reported feeling more conflicted than participants in the HC-NE condition ($M = 1.89$, $SD = 1.07$). The difference between the LC-CE condition and the HC-NE condition was also significant, $t(2,373.96) = 15.29$, $p < .001$, $d = 0.61$, 95% CI = [0.53, 0.71]. These results suggest that the conditions with an inconsistency (HC-CE and LC-CE) generated more conflict. Together, these two analyses show that writing a counterattitudinal essay elicits feelings akin to a dissonance state.

We conducted additional analyses to test whether the affective experience mediated the relationship between writing a counterattitudinal essay versus a neutral essay on the postessay attitude. We analyzed only the difference in attitude between the HC-CE condition and the HC-NE condition to focus on the effect of inconsistency. A crucial test of the mediation process is that the effect of condition diminishes when the affective measure is included in the model. Without including the affective measure, we observed a significant difference between the two conditions ($b = -0.46$, 95% CI = $[-0.61, -0.32]$). Including the affective measures in separate models increased the difference between conditions (uncomfortable: $b = -0.51$, 95% CI = $[-0.65, -0.37]$; conflicted: $b = -0.63$, 95% CI = $[-0.77, -0.48]$). Relatedly, we observed negative relationships between the affective state and the postessay attitude in the HC-CE condition, uncomfortable: $b = -0.28$, $SE = 0.040$, $t(1,445) = -6.80$, $p < .001$; conflicted: $b = -0.32$, $SE = 0.036$, $t(1,445) = -8.95$, $p < .001$. In the HC-NE condition, the relationships between affect and attitude were not significant, uncomfortable: $b = -0.029$, $SE = 0.045$, $t(1,029) = -0.65$, $p = .52$; conflicted: $b = -0.012$, $SE = 0.047$, $t(1,029) = -0.26$, $p = .80$. These results do not support a mediation process consisting of inconsistency producing feelings of discomfort and conflict that motivate positive attitude change.

Additional exploratory analyses plan

In this article, we mainly presented the results of the primary and secondary analyses. The data and the code needed to reproduce the reported analyses are available on OSF. One third of the full data set, including all recorded measures from the study, is currently available for researchers to explore the data. The full data set will be made available within 4 years after article publication. We encourage researchers to preregister an analysis plan (possibly based on exploratory analysis of the first third of the data set) before the full data set is made public (for a similar procedure, see Klein et al., 2014; or Klein et al., 2018).

Discussion

In the present study, we aimed to corroborate the induced-compliance paradigm of CDT by replicating a modified version of a carefully chosen instance of this paradigm (Croyle & Cooper, 1983, Experiment 1). Almost 4,900 participants from 39 laboratories across 19 countries were recruited to participate in an in-person lab study to perform a counterattitudinal-essay task, making this the largest CDT study to date. In each lab, participants were informed that an unpopular policy change

was being considered at their university (in most labs, raising the tuition). After being informed about this policy change, they were instructed to provide arguments (i.e., an essay). One third of the participants were directly instructed to write a counterattitudinal essay in favor of the policy (LC-CE condition). One third of participants received the same instructions but were reminded that they were free to refuse to write the essay (HC-CE condition). A final third was asked to provide arguments for other policies to be examined, under high choice (HC-NE condition). The attitude toward the essay topic was assessed just after writing the essay, and for a subset of the participants, the same attitude was also assessed in a previous session. According to CDT, we predicted that participants in the high-choice counterattitudinal condition would report a more positive attitude than participants in the low-choice condition. The neutral-essay condition constituted an additional control condition with an improved operationalization of inconsistency. We predicted participants would likewise report a more positive attitude in the high-choice counterattitudinal condition compared with the neutral-essay condition.

A vital component of the induced-compliance paradigm is that participants in the high-choice conditions experience more choice than participants in the low-choice condition so that the counterattitudinal action cannot readily be ascribed to the experimental context, which would otherwise prevent, reduce, or eliminate any cognitive dissonance. Our manipulation of choice was successful—participants in the high-choice conditions reported feeling more free to decline writing the essay than participants in the low-choice condition.

Yet despite various checks that the experiment was conducted successfully, we failed to replicate the classic dissonance effect. Across multiple analyses, we found that writing a counterattitudinal essay produced a more favorable attitude toward the essay topic than writing a neutral essay, but this occurred regardless of choice freedom. In other words, the amount of choice that participants experienced when writing the essay did not affect their subsequent attitude to the essay topic. Instead, differences in attitudes were observed only between writing counterattitudinal arguments and writing arguments for other policies (i.e., noncounterattitudinal). Participants became more in favor of the essay topic after writing counterattitudinal arguments. This pattern of results was robust across analyses involving a single attitude item, an average across multiple attitude items, and a change in attitude.

The lack of an effect of choice on attitude in our multilab replication was unexpected. This result is inconsistent with the long-standing proposition in cognitive-dissonance studies that choice is a key factor for attitude

change (see McGrath, 2017; Vaidis & Bran, 2018). The current observations are contrary to the expectations drawn from this paradigm and contradict prominent theoretical perspectives on CDT (e.g., Beauvois & Joule, 1982, 1996; Cooper, 2007; Cooper & Fazio, 1984). According to these perspectives, choice freedom was considered as the essential component to produce attitude change (e.g., Brehm & Cohen, 1962; Linder et al., 1967). Our current replication results suggest that choice may not be necessary—or might be necessary but not sufficient—to produce attitude change when using the induced-compliance paradigm.

We could also argue that the results should not be entirely surprising. Several seminal induced-compliance studies have produced effect sizes that could be argued to be unrealistically large (e.g., $d_s > 1.5$; Elliot & Devine, 1994; Simon et al., 1995) while having small samples (i.e., around 20 participants per cell). In addition, classic cognitive-dissonance designs were extensively criticized in the 1960s (see Chapanis & Chapanis, 1964), in which important methodological limitations and inadequacies in the statistical analyses were pointed out. Finally, the replication crisis in psychology has shown that even established findings in the psychological literature can fail to be reliably demonstrated. It is thus possible that typical cognitive-dissonance findings, as they were and are commonly studied, may fall in this category as well.

Despite failing to replicate the classic effect that cognitive dissonance requires experiencing a freedom of choice, we did find that participants who wrote a counterattitudinal essay reported a more positive attitude than participants who wrote a neutral essay. A straightforward interpretation could be that this is caused by the inconsistency manipulation, which would support Festinger's (1957, 2019) original argument that cognitive dissonance is first and foremost about inconsistency. This also fits with the most recent perspective on CDT (e.g., Gawronski, 2012; Harmon-Jones, 1999, 2019), which gives a more central role to inconsistency than to choice freedom. Our findings could be interpreted to support this account of cognitive dissonance, although there are several potential alternative explanations.

The finding that participants reported a more positive attitude toward the topic after giving arguments in favor could be due to a self-perception process (Bem, 1965, 1967). This explanation has, historically, been widely debated in the literature on cognitive dissonance (e.g., Beauvois & Joule, 1982; Fazio et al., 1977; Greenwald, 1975) and has profoundly influenced the development of CDT (Cooper, 2007; Vaidis & Bran, 2019). However, self-perception theory does not involve negative affect or cognitive conflict. We found that participants in the counterattitudinal conditions experienced more discomfort and more conflict than participants in the

neutral-essay condition. This finding contradicts an explanation in terms of self-perception processes.

Another alternative explanation for the observed effects is that they stem from self-persuasion processes. Generating one's own arguments, even when experimentally directed (Killeya & Johnson, 1998), can produce attitude change (e.g., Baldwin et al., 2013), particularly if people have an easy time generating the arguments (Xu & Wegener, 2023). It is plausible that participants in the counterattitudinal-essay conditions simply persuaded themselves after spending time generating reasons in favor of the essay topic. This explanation has also been a subject of long-standing debate in the dissonance literature (e.g., Cohen et al., 1959; Brehm & Cohen, 1962). Although our experiment may not be able to definitively rule out this alternative explanation, we note that self-persuasion processes do not require the emotional arousal or a sense of conflict posited in CDT. Therefore, even though both CDT and self-persuasion could potentially yield the same effect on attitude, CDT appears to account for the motivational component.

Reconsidering the induced-compliance paradigm

The induced-compliance paradigm of CDT posits that providing individuals with the choice to produce a counterattitudinal writing would create a state of cognitive dissonance, motivating them to change their attitude to align with their behavior (Cooper, 2007). According to our manipulation check, the procedure was correctly implemented: Participants did perceive more choice freedom and wrote arguments about a counterfactual topic. However, the results did not align with the expected outcomes. This raises doubts about the validity of the induced-compliance paradigm to elicit cognitive dissonance, assuming there are no substantial critiques of our study.

One critique of our study could be directed at the low control over how each lab executed the procedure (Ellefson & Oppenheimer, 2022). This possibility has only limited merit given that we standardized many of the materials, including procedural scripts of the experimenter-participant interaction, to reduce this possibility and maximize the validity of the procedure. Nevertheless, there could have been some variation in the exact wording or expression of the experimenters while interacting with the participant, although we note that multiple labs involved in this project have a significant track record of conducting dissonance studies and have successfully shown the effect in the past. Note also that we did not observe significant heterogeneity between labs in finding the effect. The inclusion of experienced labs and lack of heterogeneity speak against explaining the failure to

replicate the classic dissonance effect in terms of low control over how each lab executed the procedure.

One could also note that this study was performed in the context of a global pandemic, during which most laboratories performed the experiment with health-safety precautions (e.g., wearing masks, increased distance). In response, we point out that these precautions had already become a norm for almost a year when the data collection began in most labs. Mask wearing specifically might even have served to some extent in neutralizing facial-expression variations during interactions. Thus, it is challenging to explain the results and the absence of effect as due to any of the above circumstances.

Reexploring the role of inconsistency

Aside from results related to the effect of choice freedom on attitude change in the induced-compliance paradigm, there are several other results of this replication study that are worth highlighting. Our study showed that counterattitudinal essay writing had an effect on attitude—participants who wrote a counterattitudinal essay reported a more positive attitude compared with participants who wrote a neutral essay. Although the effect size was small, it appeared to be robust despite various methodological constraints we maintained in this replication procedure. For instance, our study had a double-blind procedure regarding the high-choice conditions to ensure that experimenters' expectations could not explain the difference between these two conditions. The effect of counterattitudinal writing on attitude change reaffirms the role of inconsistency in the dissonance process.

We conducted exploratory analyses to further understand the effects of counterattitudinal writing and found that participants in the counterattitudinal conditions reported feeling more discomfort and conflict while writing the essay. We also observed that the degree of discomfort and conflict was negatively related to the postessay attitude in the high-choice counterattitudinal condition, and we did not find evidence of a mediating influence of reported discomfort or conflict in the inconsistency-attitude relationship.

Some of the results of the exploratory analyses may be seen as incompatible with CDT. We exercise caution in drawing premature conclusions about whether these results support or contradict predictions from CDT for two reasons. First, our study was not specifically designed to capture the role of affect in the dissonance process. Second, the exact relationship between affect and attitude change is subject to debate in the literature, ranging from a positive relationship stemming from a mediation process (e.g., Devine et al., 2019) to a potentially negative relationship because of emotional reappraisal (e.g., Cancino-Montecinos et al., 2018), or when affect is

assessed following attitude change (e.g., Elliot & Devine, 1994, Study 2), or even to no relationship (e.g., Proulx, 2018). Disentangling these different processes can prove challenging, and even among dissonance theorists, consensus may be elusive regarding the appropriate methodology and ensuing interpretations.

Conclusion

The induced-compliance paradigm with a counterattitudinal-essay task was used to examine whether a state of cognitive dissonance can produce a change in attitude. This paradigm has been fundamental for CDT. Despite a large sample and an improved methodology, we did not observe the expected effect of choice on attitude. We did observe that performing counterattitudinal behavior, regardless of perceived choice, affected attitudes.

These findings prompted us to draw several conclusions. Freedom of choice, typically considered to be a vital component of the induced-compliance paradigm, does not reliably produce attitude change. Consequently, this raises serious doubts about previous findings that stem from this paradigm and certain interpretations of CDT, specifically those that have emphasized the importance of choice (e.g., Beauvois & Joule, 1996; Cooper, 2007). Consistent with the core principles of CDT, writing a counterattitudinal essay did lead to attitude change. Additional exploratory analyses revealed that writing a counterattitudinal essay also produced more discomfort and conflict. These results are consistent with CDT, although our results regarding the relationship between affect and attitude change were less clearly in support of the theory. All in all, further theoretical and empirical work is necessary to determine whether these findings align with versions of CDT focused on inconsistency (e.g., Gawronski, 2012; Harmon-Jones, 2019) or are better explained by alternative theories.

Transparency

Action Editor: Alexa M. Tullet

Editor: David A. Sbarra

Author Contribution(s)

D. C. Vaidis and W. W. A. Sleegers shared first authorship.

David C. Vaidis: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Supervision; Validation; Writing – original draft; Writing – review & editing.

Willem W. A. Sleegers: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Florian van Leeuwen: Investigation; Methodology; Writing – review & editing.

Kenneth G. DeMarree: Methodology; Supervision; Validation; Writing – review & editing.

- Bjørn Sætrevik:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- Robert M. Ross:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- Kathleen Schmidt:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- John Protzko:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- Coby Morvinski:** Investigation; Supervision; Validation; Writing – review & editing.
- Omid Ghasemi:** Investigation; Writing – review & editing.
- Andrew J. Roberts:** Investigation; Methodology; Writing – review & editing.
- Jeff Stone:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- Alexandre Bran:** Methodology; Writing – review & editing.
- Amélie Gourdon-Kanhukamwe:** Investigation; Methodology; Supervision; Validation; Writing – review & editing.
- Ceren Gunsoy:** Investigation; Writing – review & editing.
- Lisa S. Moussaoui:** Investigation; Writing – review & editing.
- Andrew R. Smith:** Investigation; Supervision; Validation; Writing – review & editing.
- Armelle Nugier:** Investigation; Supervision; Validation; Writing – review & editing.
- Marie-Pierre Fayant:** Methodology; Writing – review & editing.
- Ali H. Al-Hoorie:** Investigation; Writing – review & editing.
- Obed K. Appiah:** Investigation.
- Spencer Arbige:** Investigation.
- Benjamin Aubert-Teillaud:** Investigation.
- Olga Bialobrzeska:** Investigation; Supervision; Validation; Writing – review & editing.
- Stéphanie Bordel:** Investigation.
- Valerian Boudjemadi:** Investigation.
- Hilmar Brohmer:** Investigation.
- Quinn Cabooter:** Investigation.
- Mehdi Chahir:** Investigation.
- Ianis Chassang:** Investigation.
- Armand Chatard:** Investigation; Writing – review & editing.
- Yu Yang Chou:** Investigation; Supervision; Validation.
- Sungeun Chung:** Investigation; Supervision; Validation.
- Mioara Cristea:** Investigation; Supervision; Validation; Writing – review & editing.
- Joséphine Daga:** Investigation.
- Gregory J. Depow:** Investigation; Supervision; Validation; Writing – review & editing.
- Olivier Desrichard:** Supervision.
- Dmitrii Dubrov:** Investigation; Supervision; Validation; Writing – review & editing.
- Thomas R. Evans:** Data curation; Writing – review & editing.
- Séverine Falkowicz:** Investigation.
- Sylvain Ferreira:** Investigation.
- Tim Figureau:** Investigation.
- Valérie Fointiat:** Investigation.
- Théo Friedrich:** Investigation.
- Anastasia Gashkova:** Investigation.
- Fabien Girandola:** Investigation; Supervision; Validation.
- Marine Granjon:** Investigation; Supervision; Validation.
- Dmitry Grigoryev:** Investigation; Supervision; Validation; Writing – review & editing.
- Gul Gunaydin:** Investigation; Supervision; Validation; Writing – review & editing.
- Şevval Güzel:** Investigation.
- Mahsa Hazrati:** Investigation; Supervision; Validation.
- Mai Helmy:** Investigation; Writing – review & editing.
- Ayumi Ikeda:** Investigation.
- Michael Inzlicht:** Investigation; Supervision; Validation.
- Sara Jaubert:** Investigation; Supervision; Validation; Writing – review & editing.
- Dauren Kasanov:** Investigation.
- Mohammad Mohsen Khoddami:** Investigation; Supervision; Validation.
- Taenyun Kim:** Investigation.
- Kiyoshi Kiyokawa:** Investigation.
- Rabia I. Kodapanakkal:** Investigation.
- Alexandra Kosachenko:** Investigation; Validation; Visualization.
- Kortney Maedge:** Investigation.
- John H. Mahaney:** Investigation.
- Marie-Amélie Martinie:** Investigation; Writing – review & editing.
- Vitor N. Mascheretti:** Investigation.
- Yoriko Matsuda:** Investigation.
- Maxime Mauduy:** Investigation; Supervision; Validation; Writing – review & editing.
- Nicolas Mauny:** Investigation; Supervision; Validation.
- Armand Metzen:** Investigation.
- Eva Moreno-Bella:** Investigation.
- Miguel Moya:** Investigation; Supervision; Validation.
- Kévin Nadarajah:** Investigation; Supervision; Validation.
- Pegah Nejat:** Investigation; Supervision; Validation.
- Elisabeth Norman:** Investigation; Supervision; Validation; Writing – review & editing.
- Irmak Olcaysoy Okten:** Investigation; Supervision; Validation; Writing – review & editing.
- Asil A. Özdoğru:** Investigation; Supervision; Validation; Writing – review & editing.
- Ceyda Ozer:** Investigation; Supervision; Validation.
- Elena Padial-Rojas:** Investigation.
- Yuri G. Pavlov:** Investigation; Software; Supervision; Writing – review & editing.
- Monica Perusquia-Hernandez:** Investigation; Supervision; Validation; Writing – review & editing.
- Dora Proost:** Investigation.
- Aleksandra Rabinovitch:** Investigation.
- Odile Rohmer:** Investigation.
- Emre Selcuk:** Investigation; Supervision; Validation.
- Cécile Sénémeaud:** Investigation.
- Yaniv Shani:** Investigation.
- Elena A. Shmeleva:** Investigation.
- Emmelie Simoens:** Investigation.

Kaitlin A. Smith: Investigation; Supervision; Validation.
Alain Somat: Investigation.
Hayeon Song: Investigation; Supervision; Validation.
Fatih Sonmez: Investigation.
Lionel Souchet: Investigation.
John J. Taylor: Investigation; Supervision; Validation; Writing – review & editing.
Ilja van Beest: Investigation; Supervision; Validation; Writing – review & editing.
Nicolas Van der Linden: Investigation; Writing – review & editing.
Steven Verheyen: Investigation; Supervision; Validation; Writing – review & editing.
Bruno Verschuere: Supervision; Validation; Writing – review & editing.
Kevin Vezirian: Investigation; Supervision; Validation; Writing – review & editing.
Luc Vieira: Investigation; Supervision; Validation; Writing – review & editing.
Sera Wiechert: Investigation; Supervision; Validation.
Guillermo B. Willis: Investigation; Supervision; Validation.
Robin Wollast: Investigation.
Ji Xia: Investigation; Supervision; Validation; Writing – review & editing.
Yuki Yamada: Investigation; Supervision; Validation; Writing – review & editing.
Naoto Yoshimura: Investigation.
Daniel Priolo: Investigation; Methodology; Project administration; Supervision; Validation; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declare that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding




R. M. Ross was supported by a Macquarie University Research Fellowship and the John Templeton Foundation (Grant ID: 62631). E. A. Shmeleva was supported by a grant from the Russian Science Foundation (Grant ID: 22-18-00678, IvSU). D. Dubrov and D. Grigoryev were supported within the Basic Research Program at the National Research University Higher School of Economics. Y. Yamada was supported by JSPS KAKENHI (JP18K12015, JP20H04581, JP21H03784, and JP22K18263).

Open Practices

This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



ORCID iDs

David C. Vaidis  <https://orcid.org/0000-0002-1954-2219>
 Florian van Leeuwen  <https://orcid.org/0000-0002-9694-8300>
 Kenneth G. DeMarree  <https://orcid.org/0000-0001-5815-2646>

Bjørn Sætrevik  <https://orcid.org/0000-0002-9367-6987>
 Robert M. Ross  <https://orcid.org/0000-0001-8711-1675>
 Kathleen Schmidt  <https://orcid.org/0000-0002-9946-5953>
 John Protzko  <https://orcid.org/0000-0001-5710-8635>
 Coby Morvinski  <https://orcid.org/0000-0001-8081-7085>
 Omid Ghasemi  <https://orcid.org/0000-0001-7511-5580>
 Amélie Gourdon-Kanhukamwe  <https://orcid.org/0000-0002-3060-1320>
 Lisa S. Moussaoui  <https://orcid.org/0000-0003-0392-7402>
 Andrew R. Smith  <https://orcid.org/0000-0001-5302-3343>
 Ali H. Al-Hoorie  <https://orcid.org/0000-0003-3810-5978>
 Stéphanie Bordel  <https://orcid.org/0000-0001-7784-0217>
 Ianis Chassang  <https://orcid.org/0000-0002-1094-3331>
 Armand Chatard  <https://orcid.org/0000-0003-4823-2903>
 Yu Yang Chou  <https://orcid.org/0009-0002-3988-0917>
 Sungeun Chung  <https://orcid.org/0000-0002-9337-134X>
 Gregory J. Depow  <https://orcid.org/0000-0001-9995-4143>
 Olivier Desrichard  <https://orcid.org/0000-0003-3269-8813>
 Dmitrii Dubrov  <https://orcid.org/0000-0001-8146-4197>
 Thomas R. Evans  <https://orcid.org/0000-0002-6670-0718>
 Sylvain Ferreira  <https://orcid.org/0009-0007-3999-016X>
 Tim Figureau  <https://orcid.org/0009-0001-7535-284X>
 Dmitry Grigoryev  <https://orcid.org/0000-0003-4511-7942>
 Gul Gunaydin  <https://orcid.org/0000-0003-0490-4528>
 Ayumi Ikeda  <https://orcid.org/0000-0002-1688-2875>
 Alexandra Kosachenko  <https://orcid.org/0000-0001-8896-3837>
 Kévin Nadarajah  <https://orcid.org/0000-0003-0338-3732>
 Pegah Nejat  <https://orcid.org/0000-0003-1410-9720>
 Irmak Olcaysoy Okten  <https://orcid.org/0000-0001-6369-0879>
 Asil A. Özdoğru  <https://orcid.org/0000-0002-4273-9394>
 Ceyda Ozer  <https://orcid.org/0009-0009-5846-0492>
 Yuri G. Pavlov  <https://orcid.org/0000-0002-3896-5145>
 Monica Perusquia-Hernandez  <https://orcid.org/0000-0002-0486-1743>
 Aleksandra Rabinovitch  <https://orcid.org/0000-0002-4773-7814>
 Emre Selcuk  <https://orcid.org/0000-0002-2955-4221>
 Yaniv Shani  <https://orcid.org/0000-0001-8777-8162>
 Kaitlin A. Smith  <https://orcid.org/0000-0002-2449-2009>
 Alain Somat  <https://orcid.org/0000-0003-1415-1777>
 Fatih Sonmez  <https://orcid.org/0000-0002-4054-0269>
 Lionel Souchet  <https://orcid.org/0000-0002-5620-2790>
 Nicolas Van der Linden  <https://orcid.org/0000-0002-1288-948X>
 Bruno Verschuere  <https://orcid.org/0000-0002-6161-4415>
 Kevin Vezirian  <https://orcid.org/0000-0003-4013-7725>
 Sera Wiechert  <https://orcid.org/0000-0002-0260-1702>
 Robin Wollast  <https://orcid.org/0000-0001-5395-9969>
 Yuki Yamada  <https://orcid.org/0000-0003-1431-568X>
 Daniel Priolo  <https://orcid.org/0000-0003-3430-3459>

Acknowledgments

We would like to thank Jasmin Koglek for their assistance in correcting an error in the data analysis script.

Notes

1. The procedure called “forced compliance” was initially used to challenge the reinforcement studies (Kelman, 1953) and to show that the degree of reward for counterattitudinal advocacy has an inverse effect on attitude change. However, because of alternative explanations, the paradigm shifted to the use of choice instead of rewards to manipulate dissonance (e.g., Linder et al., 1967).
2. Converted from a reported F value of 38.86 with 1 and 27 degrees of freedom using the *effectsize* R package (Ben-Shachar et al., 2020).
3. Converted from a reported effect size of $r = .248$ using the *effectsize* R package (Ben-Shachar et al., 2020).
4. The topics, procedure, and results of the pretest are available on OSF (<https://osf.io/mgjh8/>).
5. This calculation is based on the Qualtrics time-stamp function and could be inaccurate depending on the final validation of the survey by the participant (with an automatic closure of a survey after 1 week). All the data are available for exploratory purposes.
6. From Croyle and Cooper (1983). Adapted phrasings were also proposed to fit with sessions with multiple participants. For details of the material in each language, see OSF.
7. A part of the literature (e.g., Elliot & Devine, 1994, Experiment 1; Jonas et al., 2014; Proulx & Inzlicht, 2012) suggests that a delay could have an impact on the regulation process of CDT. Most of the dissonance literature has not included a delay. The time for option selection, page submission, and the subsequent measures were recorded for exploratory purposes.
8. All three conditions displayed a significant positive attitude change from the first session to the second session: HC-CE: $t(1,060) = 17.97, p < .001$; LC-CE: $t(936) = 16.50, p < .001$; HC-NE: $t(725) = 8.53, p < .001$. We also reanalyzed the primary analyses and included an interaction test between condition and whether the participant completed the premeasure. A significant interaction would indicate that the presence or absence of the premeasure affected the difference in attitudes between conditions. No significant interactions were found (HC-CE vs. LC-CE: $b = -0.098, SE = 0.16, t(2,787) = -0.60, p = .55$; HC-CE vs. HC-NE: $b = -0.13, SE = 0.16, t(2,475) = -0.79, p = .43$), meaning there is no evidence to suggest that the premeasure affected the impact of the manipulation.

References

- Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry, 3*(4), 303–311. https://doi.org/10.1207/s15327965pli0304_1
- Aronson, E. (2019). Dissonance, hypocrisy, and the self-concept. In E. Harmon-Jones (Ed.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (pp. 141–157). American Psychological Association. <https://doi.org/10.1037/0000135-007>
- Aronson, E., & Aronson, J. (2018). *The social animal* (12th ed.). Worth Publishers.
- Azdia, T., & Joule, R. V. (2001). Double soumission forcée et engagement: le cas des comportements inconsistants [Double forced-compliance and commitment: the case of inconsistent behavior]. *Revue Internationale de Psychologie Sociale, 14*(1), 31–55.
- Baldwin, A. S., Rothman, A. J., Vander Weg, M. W., & Christensen, A. J. (2013). Examining causal components and a mediating process underlying self-generated health arguments for exercise and smoking cessation. *Health Psychology, 32*, 1209–1217. <https://doi.org/10.1037/a0029937>
- Baumeister, R. F., & Tice, D. M. (1984). Role of self-presentation and choice in cognitive dissonance under forced compliance: Necessary or sufficient causes? *Journal of Personality and Social Psychology, 46*(1), 5–13. <https://doi.org/10.1037/0022-3514.46.1.5>
- Beauvois, J. L., & Joule, R. V. (1982). Dissonance versus self-perception theories: A radical conception of Festinger's theory. *The Journal of Social Psychology, 117*(1), 99–113.
- Beauvois, J. L., & Joule, R. V. (1996). *A radical dissonance theory*. Taylor & Francis. <https://doi.org/10.1037/10318-003>
- Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology, 1*, 199–218.
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review, 74*, 183–200.
- Ben-Shachar, M. S., Makowski, D., & Lüdtke, D. (2020). *Compute and interpret indices of effect size*. <https://github.com/easystats/effectsize>
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52*, 384–389. <https://doi.org/10.1037/h0041006>
- Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. John Wiley & Sons. <https://doi.org/10.1037/11622-000>
- Briñol, P., McCaslin, M. J., & Petty, R. E. (2012). Self-generated persuasion: Effects of the target and direction of arguments. *Journal of Personality and Social Psychology, 102*(5), 925–940. <https://doi.org/10.1037/a0027231>
- Cancino-Montecinos, S., Björklund, F., & Lindholm, T. (2018). Dissonance reduction as emotion regulation: Attitude change is related to positive emotions in the induced compliance paradigm. *PLOS ONE, 13*(12), Article e0209012. <https://doi.org/10.1371/journal.pone.0209012>
- Chapanis, N. P., & Chapanis, A. (1964). Cognitive dissonance. *Psychological Bulletin, 61*(1), 1–22. <https://doi.org/10.1037/h0043457>
- Cohen, A. R., Brehm, J. W., & Fleming, W. H. (1958). Attitude change and justification for compliance. *The Journal of Abnormal and Social Psychology, 56*(2), 276–278. <https://doi.org/10.1037/h0047070>
- Cohen, A. R., Terry, H. I., & Jones, C. B. (1959). Attitudinal effects of choice in exposure to counterpropaganda. *The Journal of Abnormal and Social Psychology, 58*(3), 388–391. <https://doi.org/10.1037/h0046951>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum.
- Cooper, J. (2007). *Cognitive dissonance: 50 years of a classic theory*. Sage. <https://doi.org/10.4135/9781446214282>
- Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In *Advances in experimental social psychology* (Vol. 17, pp. 229–266). Academic Press.
- Cooper, J., & Feldman, L. A. (2019). Does cognitive dissonance occur in older age? A study of induced compliance in a

- healthy elderly population. *Psychology and Aging*, 34(5), 709–713. <https://doi.org/10.1037/pag0000338>
- 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>
- 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>
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, 30(1), 92–101. <https://doi.org/10.5334/irsp.82>
- Devine, P. G., & Brodish, A. B. (2003). Modern classics in social psychology. *Psychological Inquiry*, 14(3–4), 196–202. <https://doi.org/10.1080/1047840x.2003.9682879>
- Devine, P. G., Tauer, J. M., Barron, K. E., Elliot, A. J., Vance, K. M., & Harmon-Jones, E. (2019). Moving beyond attitude change in the study of dissonance-related processes: An update on the role of discomfort. In E. Harmon-Jones (Ed.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (pp. 247–269). American Psychological Association. <https://doi.org/10.1037/0000135-012>
- Ellefsen, M. R., & Oppenheimer, D. M. (2022). Is replication possible without fidelity? *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000473>
- Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, 67(3), 382–394. <https://doi.org/10.1037/0022-3514.67.3.382>
- Eutsler, J., & Lang, B. (2015). Rating scales in accounting research: The impact of scale points and labels. *Behavioral Research in Accounting*, 27(2), 35–51. <https://doi.org/10.2308/bria-51219>
- Fazio, R. H., Zanna, M. P., & Cooper, J. (1977). Dissonance and self-perception: An integrative view of each theory's proper domain of application. *Journal of Experimental Social Psychology*, 13(5), 464–479.
- Ferrari, P. A., & Barbiero, A. (2012). Simulating ordinal data. *Multivariate Behavioral Research*, 47(4), 566–589. <https://doi.org/10.1080/00273171.2012.692630>
- Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson.
- Festinger, L. (2019). Reflections on cognitive dissonance: 30 years later. In E. Harmon-Jones (Ed.) *Cognitive dissonance: Progress on a pivotal theory in social* (pp. 381–385). American Psychological Association.
- 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>
- Forstmann, M., & Sagioglou, C. (2020). Religious concept activation attenuates cognitive dissonance reduction in free-choice and induced compliance paradigms. *The Journal of Social Psychology*, 160(1), 75–91. <https://doi.org/10.1080/00224545.2019.1609400>
- Freijy, T., & Kothe, E. J. (2013). Dissonance-based interventions for health behaviour change: A systematic review. *British Journal of Health Psychology*, 18(2), 310–337. <https://doi.org/10.1111/bjhp.12035>
- Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition*, 30(6), 652–668. <https://doi.org/10.1521/soco.2012.30.6.652>
- Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, 40(4), 535–542. <https://doi.org/10.1016/j.jesp.2003.10.005>
- Gawronski, B., & Strack, F. (2012). *Cognitive consistency: A fundamental principle in social cognition*. The Guilford 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>
- Greenwald, A. G. (1975). On the inconclusiveness of “crucial” cognitive tests of dissonance versus self-perception theories. *Journal of Experimental Social Psychology*, 11(5), 490–499. [https://doi.org/10.1016/0022-1031\(75\)90051-7](https://doi.org/10.1016/0022-1031(75)90051-7)
- Griggs, R. A., & Christopher, A. N. (2016). Who's who in introductory psychology textbooks: A citation analysis redux. *Teaching of Psychology*, 43(2), 108–119. <https://doi.org/10.1177/00986283166636276>
- Haggblom, 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>
- Harmon-Jones, E. (1999). Toward an understanding of the motivation underlying dissonance effects: Is the production of aversive consequences necessary. In E. Harmon-Jones, & J. Mills (Eds.), *Cognitive dissonance: Progress on a Pivotal Theory in Social Psychology* (pp. 71–99). American Psychological Association.
- 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>
- Heine, S. J., & Lehman, D. R. (1997). Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin*, 23(4), 389–400. <https://doi.org/10.1177/0146167297234005>
- Hüffmeier, J., Mazei, J., & Schultze, T. (2016). Reconceptualizing replication as a sequence of different studies: A replication typology. *Journal of Experimental Social Psychology*, 66, 81–92. <https://doi.org/10.1016/j.jesp.2015.09.009>
- Janis, I. L., & King, B. T. (1954). The influence of role playing on opinion change. *The Journal of Abnormal and Social Psychology*, 49(2), 211–218. <https://doi.org/10.1037/h0056957>
- 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). Academic Press. <https://doi.org/10.1016/b978-0-12-800052-6.00004-4>
- Kelman, H. (1953). Attitude change as a function of response restriction. *Human Relations*, 6, 185–214. <https://doi.org/10.1177/001872675300600301>
- Kenworthy, J. B., Miller, N., Collins, B. E., Read, S. J., & Earleywine, M. (2011). A trans-paradigm theoretical synthesis of cognitive dissonance theory: Illuminating the nature of discomfort. *European Review of Social Psychology*, 22(1), 36–113. <https://doi.org/10.1080/10463283.2011.580155>
- Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief*. Academic Press.
- Killea, L. A., & Johnson, B. T. (1998). Experimental induction of biased systematic processing: The directed-thought technique. *Personality and Social Psychology Bulletin*, 24(1), 17–33. <https://doi.org/10.1177/0146167298241002>
- Kim, S. Y., Allen, M., Preiss, R. W., & Peterson, B. (2014). Meta-analysis of counterattitudinal advocacy data: Evidence for an additive cues model. *Communication Quarterly*, 62(5), 607–620. <https://doi.org/10.1080/01463373.2014.949385>
- Kitayama, S., Snibbe, A. C., Markus, H. R., & Suzuki, T. (2004). Is there any “free” choice? Self and dissonance in two cultures. *Psychological Science*, 15(8), 527–533. <https://doi.org/10.1111/j.0956-7976.2004.00714.x>
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemačić, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., . . . Nosek, B. A. (2014). Investigating variation in replicability: A “many labs” replication project. *Social Psychology*, 45(3), 142–152. <https://doi.org/10.1027/1864-9335/a000178>
- Klein R. A., Vianello M., Hasselman F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., . . . Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 1(4), 443–490. doi:10.1177/2515245918810225
- Krosnick, J. A., & Presser, S. (2009). Question and questionnaire design. In P. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (pp. 263–314). Emerald.
- Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, 6(3), 245–254. <https://doi.org/10.1037/h0021220>
- Losch, M. E., & Cacioppo, J. T. (1990). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology*, 26(4), 289–304. [https://doi.org/10.1016/0022-1031\(90\)90040-s](https://doi.org/10.1016/0022-1031(90)90040-s)
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253. <https://doi.org/10.1037/0033-295x.98.2.224>
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31(3), 657–674. <https://doi.org/10.1177/001316447103100307>
- Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods*, 9(2), 147–163. <https://doi.org/10.1037/1082-989x.9.2.147>
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, 70(6), 487–498. <https://doi.org/10.1037/a0039400>
- McGrath, A. (2017). Dealing with dissonance: A review of cognitive dissonance reduction. *Social and Personality Psychology Compass*, 11(12), Article e12362. <https://doi.org/10.1111/spc3.12362>
- Murray, A. A., Wood, J. M., & Lilienfeld, S. O. (2012). Psychopathic personality traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality*, 46(5), 525–536. <https://doi.org/10.1016/j.jrp.2012.05.011>
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology’s renaissance. *Annual Review of Psychology*, 69, 511–534. <https://doi.org/10.1146/annurev-psych-122216-011836>
- Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLOS Biology*, 18(3), Article e3000691. <https://doi.org/10.1371/journal.pbio.3000691>
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776–783. <https://doi.org/10.1037/h0043424>
- Park, J., & Kitayama, S. (2014). Interdependent selves show face-induced facilitation of error processing: Cultural neuroscience of self-threat. *Social Cognitive and Affective Neuroscience*, 9, 201–208. <https://doi.org/10.1093/scan/nss125>
- Pashler, H., & Wagenmakers, E. (2012). Editors’ introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7(6), 528–530. <https://doi.org/10.1177/1745691612465253>
- Proulx, T. (2018). Lumping the affective and behavioral responses to inconsistency: A lump too far? *Psychological Inquiry*, 29(2), 82–85. <https://doi.org/10.1080/1047840X.2018.1480588>
- 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>
- Randles, D., Inzlicht, M., Proulx, T., Tullett, A. M., & Heine, S. J. (2015). Is dissonance reduction a special case of fluid compensation? Evidence that dissonant cognitions cause compensatory affirmation and abstraction. *Journal of Personality and Social Psychology*, 108(5), 697–710. <https://doi.org/10.1037/a0038933>
- Ranganathan, P., Pramesh, C. S., & Buyse, M. (2016). Common pitfalls in statistical analysis: The perils of multiple testing. *Perspectives in Clinical Research*, 7(2), 106–107.
- Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331–363. <https://doi.org/10.1037/1089-2680.7.4.331>
- Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers*

- in *Psychology*, 10, Article 813. <https://doi.org/10.3389/fpsyg.2019.00813>
- Schaffer, W. M. (1974). Optimal reproductive effort in fluctuating environments. *The American Naturalist*, 108(964), 783–790. <https://doi.org/10.1086/282954>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, 31(4), 557–566. <https://doi.org/10.1037/pas0000648>
- 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>
- Sleegers, W. W. A., & Vaidis, D. C. (2019, July). *Setting up a large scale replication project of cognitive dissonance theory* [Paper presentation]. Hackathon presented at the 4th meeting of the Society for Improvement of Psychological Science, Rotterdam.
- Stalder, D. R., & Baron, R. S. (1998). Attributional complexity as a moderator of dissonance-produced attitude change. *Journal of Personality and Social Psychology*, 75(2), 449–455. <https://doi.org/10.1037/0022-3514.75.2.449>
- 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 and Social Psychology Bulletin*, 35(1), 126–137. <https://doi.org/10.1177/0146167208325613>
- Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, 45(1), 5–19. <https://doi.org/10.1037/0022-3514.45.1.5>
- Stone, J., & Cooper, J. (2003). The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes. *Journal of Experimental Social Psychology*, 39(5), 508–515. [https://doi.org/10.1016/s0022-1031\(03\)00018-0](https://doi.org/10.1016/s0022-1031(03)00018-0)
- 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), Article 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, Article 1189. <https://doi.org/10.3389/fpsyg.2019.01189>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Erlbaum.
- Xu, M., & Wegener, D. T. (2023). Persuasive benefits of self-generated arguments: Moderation and mechanism. *Social Psychological and Personality Science*. Advance online publication. <https://doi.org/10.1177/1948550622146612>
- 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, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, 111(4), 493–504. <https://doi.org/10.1037/pspa0000056>