

Systematic Review

Thus We Are Agreed! A Systematic Mapping Review of Adversarial Collaborations in Science

Annalena Seyr, Ulrich S. Tran  and Martin Voracek 

Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, A-1010 Wien, Austria; 98annalena@gmail.com (A.S.); ulrich.tran@univie.ac.at (U.S.T.)

Abstract

In a time where the methods of science are frequently called into question, open science practices are gaining increased importance. A particularly cooperative form of transparent science is the approach of adversarial collaboration (AC), where research groups with different explanations of the same phenomenon work closely together to develop a study and publish their results jointly, regardless of the outcome. The aim of this work, a systematic mapping review, is to provide a comprehensive overview of the formal aspects of AC and identify the relevant factors that need to be considered when applying this method. For this purpose, both AC studies and publications related to this topic were analyzed. The total of 89 articles retrieved, analyzed, and integrated indicates that accounts on this topic have slightly increased (mostly in psychology) in recent years. However, only a handful of authors have delved more deeply into ACs, and there currently is no journal specifically promoting ACs. The analyzed sources provide a solid foundation for best-practice recommendations and a comprehensive overview of the pros and cons of AC, in the form of organized lists of AC-related known advantages and challenges (alongside solutions to the latter ones). Specifically, some AC studies show weaknesses in study quality, mainly due to a lack of information on the implementation steps. This systematic mapping review serves not only as an introduction to the topic, but also as a guide for future ACs.

Keywords: adversarial collaboration; systematic mapping review; replication crisis; research methods; science reform; team science; open science; metascience



Academic Editor: Andrew Kirby

Received: 17 October 2024

Revised: 15 July 2025

Accepted: 13 July 2026

Published: 24 July 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Science evidently is in crisis ([Ellemers et al., 2020](#)). Since the emergence of the replicability debate and confidence crisis across many empirical disciplines during the early 2010s, including, but not confined to, psychological science (e.g., irreproducibility of research findings, the prevalence and extent of questionable research practices and underpowered studies, and related phenomena), it has become clear that results from individual studies must be viewed with skepticism ([Makel et al., 2019](#)). The traditional study design invites selection and confirmation bias. Study results can be exaggerated, and far-reaching implications might be drawn from them even though they lack statistical power ([Bateman](#)

*Thus we are agreed.
I crave our composition may be written,
And seal'd between us.*
Shakespeare, Antony and Cleopatra, II.6

et al., 2005). Low-impact studies often are less interesting for journals and their readers and thus are less frequently published (Marks-Anglin & Chen, 2020). Different explanatory approaches of different research groups lead to the dissection of contradictory results and authors being criticized, sometimes on a personal level (Tetlock & Mitchell, 2009b).

The scientific validity of an empirical study and its execution and interpretation are evaluated through peer review, which is carried out by scientists on a purely voluntary basis (Buranyi, 2017). In a lot of cases, this system apparently does not suffice to guarantee scientific quality. This raises the question as to how much trust can be placed in the results of studies and the conclusions drawn from them. The problems not only affect empirical methods or statistics but already begin during the formation of theories. Instead of gaining new knowledge through scientific curiosity, the focus is on proving that one's own hypotheses are better than those of others (Ellemers et al., 2020).

Research in the social sciences follows the model of publication–critique–response–rebuttal: a scientist publishes an article in which they advocate a particular hypothesis. Next, another scientist publishes a critique of that study. The first researcher then publishes a rebuttal to the critique, which is followed by another response (Berzon, 2012; Tetlock & Mitchell, 2009b). These days, a theory does not stand on its own; it is assigned to certain groups of researchers, which creates competition. Using the theories of other researchers is equated with using someone else's toothbrush. This undermines the constructive exchange between researchers and hampers the scientific process. Instead of rigidly adhering to one's own theories, the merging of knowledge and the synthesis of theories is much more fruitful.

Currently, there exist no clear guidelines on how research groups with contradictory theories can collaborate with each other. Adversarial collaboration offers a possible framework for this type of collaboration (Ellemers et al., 2020). The term “adversarial collaboration” (AC) was introduced by Kahneman in 2001 (Mellers et al., 2001) and refers to a collaboration between two or more research groups who have different explanations for a particular phenomenon. It goes without saying that this approach is not without forerunners, which, for instance, are found in the falsificationist (i.e., critical rationalism) philosophy of science of Popper (1945/2012, p. 489), who early on already highlighted the “friendly-hostile co-operation of many scientists” necessary for the progress of science and scientific objectivity and trustworthiness (in this context, see Trafimow (2025) for a broader comparative and evaluative view on this stance of Popper, alongside further philosophy of science perspectives due to Lakatos, Laudan, and Feyerabend on scientific collaboration and theory refutation attempts).

The aim of an AC is to test their different views in a jointly designed study and to publish the results regardless of the outcome of the experiments. Such collaboration can be challenging due to disagreements. Therefore, a neutral person (a mediator) can be involved, who can act as an intermediary between the research groups and is not committed to either of the controversial views (Mellers et al., 2001).

The AC format is relatively new and not widely used. Only 14 ACs have been conducted so far, and several consensus papers have been published about the AC approach. Clark and her colleagues have published several reviews that mainly focus on the advantages of ACs, whereas disadvantages and challenges are less frequently mentioned (Clark et al., 2022a, 2022b, 2022c; Clark & Tetlock, 2022). Rakow (2022) provided a first overview of the content of ACs that have been conducted thus far. His account summarized selected ACs and their results. However, this work did not include all retrievable ACs published to date. The studies were not directly compared to one another, and no clear overview was provided of how ACs differ or are similar, such as how often a mediator was involved, whether and on what type of formal agreements (e.g., preregistration) the researchers agreed, etc. So-called quasi adversarial collaborations (quasi ACs) are only

briefly mentioned. Such collaborations resemble the AC concept but do not meet all its requirements. For example, [Ellemers et al. \(2020\)](#) attempted to integrate different models of social evaluation into a common framework in a comprehensive project with representatives of five different theories. However, because these researchers did not conduct a hypothesis-testing experiment, this form of collaboration is referred to as a quasi AC. These types of collaborations provide valuable information for the implementation of classic (i.e., full-fledged) ACs, information that is missing from extant conceptual reviews and opinion papers that have been published thus far (e.g., [Rakow, 2022](#); [Clark & Tetlock, 2022](#)). Some authors who have conducted ACs have summarized their experiences and recommendations in additional opinion pieces ([Lindner & Tribble, 2020](#)). Such related accounts provide important and exciting insights, but systematic comparisons with other ACs are still lacking.

The present study is thus the first systematic mapping review ([Grant & Booth, 2009](#)) on the topic of AC (for the PRISMA 2020 checklist, see Table S1). The aim of this work is to dissect, summarize, and integrate (i.e., “to map out”) the specifics of, and findings from, studies conducted as ACs, as well as to collect and elaborate on some concerns and considerations about this procedure from additional papers that did not conduct ACs themselves. In contrast to prior related work on the topic of ACs, our work provides a comprehensive overview of bibliographic data. In the vein of a mapping review, it describes which scientific fields, journals, and authors have dealt with ACs thus far. This allows estimating how widely the concept of AC is actually spread in the scientific community. For the first time, not only is the study quality of ACs evaluated, but further (quasi) ACs are also captured and compared to one another. The data from all analyzed papers are distilled into best-practice recommendations, which can serve as a possible guide for future ACs. In addition, the advantages of ACs are described in detail, as well as their challenges. This systematic mapping review thus contains a unique and essential overview of current research evidence on ACs, providing not only a deeper understanding of the concept of AC, including laypeople, but also serving as a helpful guide for research groups planning to use this method in the future.

2. Materials and Methods

For this systematic mapping review, the databases Web of Science, Scopus, PubMed, PsychInfo, and Google Scholar were used to find articles that were published since 1988 and used the term “adversarial collaboration” in the title, abstract, or keywords. The first AC was conducted in 1988 ([Latham et al., 1988](#)). All types of publications were considered, including editorial material, book chapters, and unpublished works. The literature search was finalized in December 2022, and the most recently published article eligible for this review was [Clark et al. \(2022b\)](#).

After the databases were searched and all relevant articles collected in a separate document, duplicates of articles were removed and all studies were excluded that used “adversarial collaboration” to describe other phenomena (e.g., business dealings in the film industry; [Watson, 2004](#)) than those examined within the scope of this review. [Heyman et al. \(2020\)](#) describe the benefits of ACs for comments; this paper was excluded because instructions or recommendations for writing comments are not the subject of this work. The included studies were all in English, although this was not an explicit search criterion. Backward citation searching and cited-reference searches were utilized to identify additional relevant works. One reason why this was necessary was that the term “adversarial collaboration” was only introduced by Kahneman in 2001, whereas the first collaboration of this type had already been conducted in 1988 ([Gilovich et al., 1998](#); [Mellers et al., 2001](#)). Additionally, publications were found on websites of organizations that promote ACs (e.g., the

“Adversarial Collaboration Project” at the University of Pennsylvania; [Clark et al., 2022a](#)). The search criteria were applied by one researcher (author A.S.). A total of 258 papers were identified, of which 89 were included in this systematic mapping review after applying the exclusion criteria (see Figure 1, PRISMA 2020 flow diagram). From the search criteria onwards, all further decisions in the course of this review work were validated by two further researchers (authors U.S.T. and M.V.), following consensus procedures (i.e., entailing independent evaluation and joint discussion of any substantive discrepancies until reaching agreement).

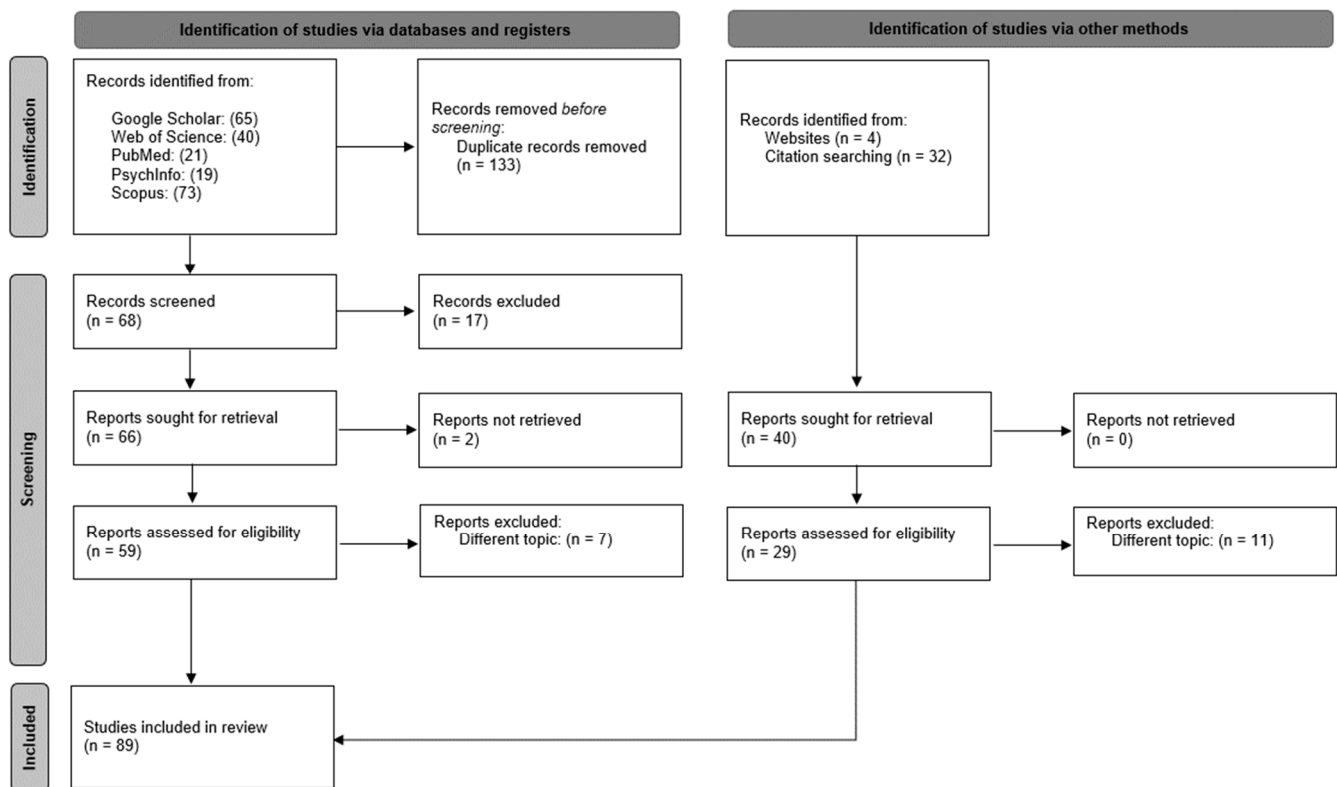


Figure 1. PRISMA 2020 flow diagram ([Page et al., 2021](#)) of literature search results.

In this review, high-quality publications are those that provide particularly informative content about the implementation, benefits, and drawbacks of an AC. Most of the included studies are research articles without new data that base their conclusions primarily on the papers that actually conducted an AC. There currently are no quality assessment instruments for studies conducted as ACs. To still assess their study quality, a list of AC quality categories was created based on our own considerations, alongside the prerequisites for ACs mentioned by [Mellers et al. \(2001\)](#).

To assess the quality of the ACs performed, the following aspects were examined (all coded with yes/no):

Roles: Was it clearly stated which party held which views? This captures potential bias, conflicts of interest, or similar issues and ensures an open and ethical approach.

Theory: Were the views supported by scientific sources? This ensures the scientific nature of the study and prevents the investigation of mere opinions.

Communication: Was the communication process and/or the form in which information was exchanged documented?

Protocol: Was a protocol like an AC agreement or a preregistration included? This ensures transparency and replication possibilities.

Decision-making: Explanation of why the parties chose the methods used to maintain the validity and reliability of the experimental design.

Discussion: Clear separation of where the parties agree and where they differed.

General information—Definition: Is there a clear definition of AC?

Procedure: Was the implementation of the AC described?

Article: Did the authors publish additional articles focusing on the implementation of the AC?

Methodological rigor (e.g., randomization, reliability, and validity of methods or analytical methods) was not assessed, as the parties involved in an AC usually check and question each other's assumptions to obtain the best possible outcome (Mellers et al., 2001). It should be noted that the above list of AC quality features only constitutes an initial attempt of ours to categorize the quality of AC studies and that in this case study, quality refers not to the reliability and validity of study methods and results but rather to how accurately the studies describe the prerequisites of the AC implemented.

The following characteristics of the included publications were collected: publication year, type of publication, journal and research field, authorship, study quality, implementation of an AC, role of the mediator, advantages and challenges of ACs and quasi ACs. As elaborated on above, our term quasi adversarial collaborations (quasi ACs) refers to collaborations that are similar to ACs but do not meet all of its requirements. According to the criteria for ACs, as originally advanced by Mellers et al. (2001), the five most important requirements include (1) a minimum of two research groups with different explanations for the same phenomenon; (2) development and documentation of a joint study design; (3) formulation of mutually exclusive hypotheses; (4) testing of hypotheses through an experiment; and (5) joint publication of results, regardless of the outcome. In what follows, we use all these features to provide a comprehensive overview of the important aspects of individual ACs and the respective underlying the AC literature at large.

3. Results

3.1. Emergence and Time Trends

The publication period from 1988 to 2022 was reviewed. Figure 2 provides an overview of the publication timeline of known ACs, with ACs and quasi ACs presented separately from the other papers. The oldest study, describing the collaboration between parties with opposing views, was published in 1988, whereas the now common term “adversarial collaboration” was only introduced and established by Kahneman more than a decade later (Kahneman & Anderson, 2003; Latham et al., 1988; Mellers et al., 2001). It is interesting to note that this term was already mentioned back in 2000; however, at that time, it referred to collaboration in legal firms and did not mention clear rules for scientific collaboration (Cohen et al., 2000). For these reasons, studies published prior to 2001 could only be identified through backward citation searching. It is thus conceivable that during this period, additional ACs may have been conducted that then remained unidentified and are not covered in the present review.

From 2001 onwards, the number of research papers, as well as related accounts, increased slightly. A somewhat stronger increase was observed in 2009 and 2015. This latter observation may well be related to the emerging replication debate and confidence crisis in psychology and other empirical disciplines (from 2011/12 onwards) and to the publication of a landmark paper in response to these events, the Reproducibility Project Psychology in particular (Open Science Collaboration, 2012, 2015).

In 2021, the “Adversarial Collaboration Project” was launched at the University of Pennsylvania to promote ACs, which may well have contributed to the increase in written discussions on this topic in 2022 (Clark et al., 2022a). However, this interpretation should

be taken with caution, as it could be an outlier, particularly due to the numerous reactions to the publication of an overview on ACs (Clark et al., 2022b, 2022c; Ceci & Williams, 2022; Cowan, 2022; Marsh, 2022; Melloni, 2022; Tatlidil & Sloman, 2022; Tullett, 2022; Vlasceanu et al., 2022). In 2022, no new ACs emerged, but several ACs were planned for the coming years (Greeno et al., 2022; Melloni, 2022; Research, n.d.).

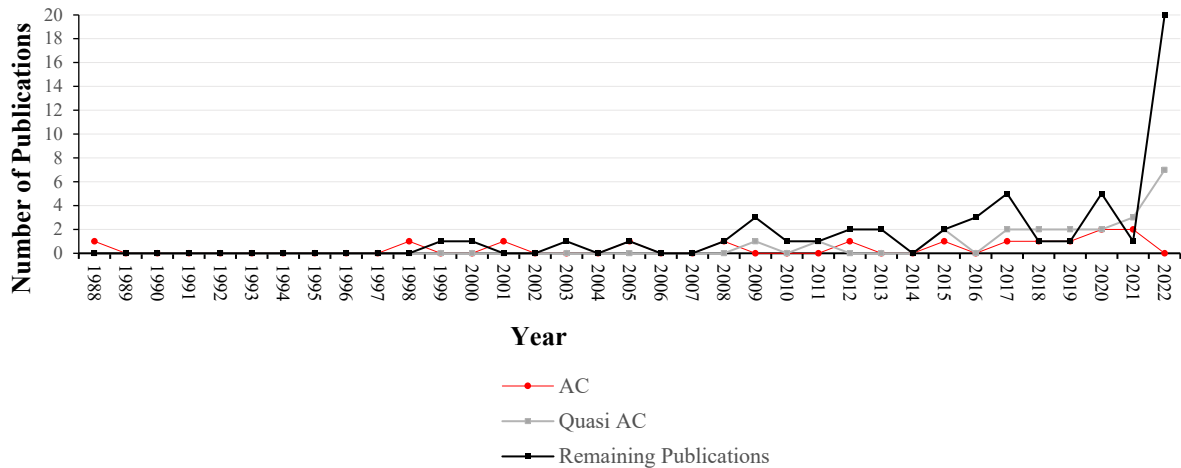


Figure 2. Number of AC publications per year.

3.2. Publication Types

The 89 analyzed publications include 52 articles, of which 14 describe ACs and 23 quasi ACs. Evidently, the number of quasi ACs outweighs the number of ACs, because the former type has to meet fewer requirements than the latter one. The remaining articles were mostly general overviews of procedures to make research more transparent (e.g., Duygu, 2022; Edlund et al., 2022; Gordon et al., 2015; Koole & Lakens, 2012; Makel et al., 2019, 2020; VanLange, 2013), or ACs were mentioned as a way to clarify open questions (e.g., Cleeremans, 2022; Engelhardt, 2020; Leykin & DeRubeis, 2009; D. O. Martschenko & Trejo, 2022; Taylor & Munafò, 2016). In addition, some articles have been written by authors who conducted ACs themselves and therein describe their experiences (e.g., Cowan et al., 2020; Ellemers et al., 2020; Lindner & Tribble, 2020). Further publications mainly included commentaries, book chapters, and various editorial material. For a more detailed overview of the types of publications collected, such as the number of comments or book chapters, see Figure 3.

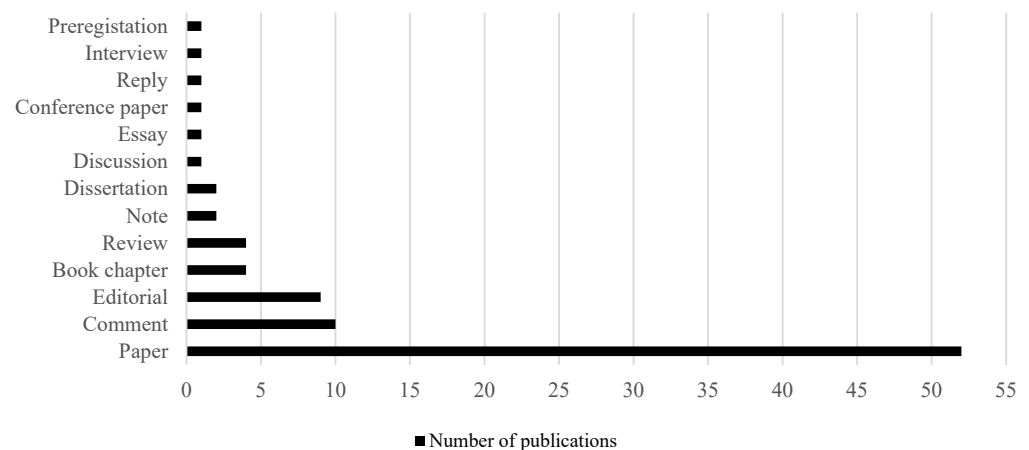


Figure 3. Number of AC publications according to type of publication.

3.3. Research Field and Journals

For the overview of the AC literature by research field, it was first recorded which articles had been published in which journal; then, the journals were categorized into the corresponding research fields, using the categories of the Journal Citation Report (Reuters, 2022). Publications that did not appear in journals were categorized based on the present authors' own assessment.

The results are presented in Table 1. Of note, most of the literature is from the field of psychology or from an intersection of psychology and other fields, such as economics (Corrigan et al., 2012; Duarte et al., 2015). A prerequisite of AC is the collaboration of different parties; presumably for this reason, most of the publications can be found in the field of multidisciplinary psychology. Papers from other fields, unrelated to psychology, are virtually nonexistent, with a few exceptions (Berzon, 2012; Tetlock & Mellers, 2011). Specifically, it is obvious that, for reasons that remain to be elucidated, ACs have not yet been implemented in disciplines like biology and physics (although the approach of “strong inference”, which can be taken as a precursor to the idea of ACs, originated from, and specifically was directed to, the field of biology and the life sciences more generally; see (Platt, 1964; Fudge, 2014)).

Table 1. Overview of research fields and selected journals from which AC publications originate.

Journals	Research Field	Total Number of Publications		Number of	
		<i>n</i>	%	AC	Quasi AC
<i>Advances in Methods and Practices in Psychological Science; American Psychologist; BMC Psychology; Emotion Review; Journal of Comparative Psychology; Judgement & Decision Making; Perspectives on Psychological Science; Psychological Inquiry; Psychology of Popular Media; Psychological Review; Psychological Science; The Psychologist</i>	Multidisciplinary Psychology	23 ^a	25.8	3	7
<i>Attention Perception & Psychophysics; Journal of Applied Research in Memory and Cognition; Journal of Experimental Psychology: General; Journal of Experimental Psychology: Learning Memory and Cognition; Nature Human Behaviour; Thinking and Reasoning</i>	Experimental Psychology	19 ^b	21.3	3	2
<i>American Journal of Psychiatry; Clinical Psychology: Science and Practice; European Journal of Psychotherapy & Counselling; Journal of Personality Assessment; Psychological Medicine</i>	Clinical Psychology/Psychiatry	11	12.4		5
<i>Comprehensive Results in Social Psychology; Journal of Experimental Social Psychology; Sex Roles; Personality and Social Psychology Review; Social Psychological and Personality Science</i>	Social Psychology	9	10.1	4	
<i>Journal of Applied Psychology; Journal of Behavioral Decision Making; Research in Organizational Behavior</i>	Applied Psychology	4	4.5	1	

Table 1. Cont.

Journals	Research Field	Total Number of Publications		Number of	
		<i>n</i>	%	AC	Quasi AC
<i>International Journal of Transpersonal Studies; Journal of Scientific Exploration</i>	Transpersonal Studies	4	4.5		3
<i>AREA Open; Education & Educational Research; Journal for the Education of the Gifted; Teachers College Record</i>	Education	4	4.5		2
<i>American Journal of Agricultural Economics; Journal of Public Economics; Public Choice</i>	Economics	3	3.4	3	
<i>Journal for Philosophy of Science; Synthese</i>	Philosophy	3 ^c	3.4		2
<i>Behavioral and Brain Sciences; Brain Sciences</i>	Neurosciences	2	2.2		
<i>Family Relations</i>	Family Studies	1	1.1		
<i>Human Genetics and Genomics Advances</i>	Genetics	1	1.1		
<i>California Law Review</i>	Law	1	1.1		
<i>Management Science</i>	Management	1	1.1		1
<i>Proceedings of the National Academy of Sciences of the USA</i>	Multidisciplinary Sciences	1	1.1		
<i>British Politics</i>	Political Science	1	1.1		1
<i>Journal of Consciousness Studies</i>	Social Sciences	1	1.1		

Note. *N* = 89 publications in total. Not all of these were published as journal articles. Book chapters, academic theses, and study preregistrations categorized under ^a Multidisciplinary Psychology (Thompson, 2021; Clark et al., 2022c); ^b Experimental Psychology (Gandalf, 2020; Rakow, 2022; Greeno et al., 2022); ^c Philosophy (Leal & Marraud, 2022).

For a broad overview of the topics covered by the ACs, see Table 2. Some details on research fields and topics covered include the following observations: In the field of clinical psychology, the problem of an allegiance effect in psychotherapy has been frequently addressed. Researchers belonging to a particular school of psychotherapy tend to rate it as most effective in both individual studies and meta-analyses on this topic (Berman & Reich, 2010; Clark et al., 2022c; Hollon, 1999; Leykin & DeRubeis, 2009). Also, a large-scale project examined the appropriateness of using certain clinical tests in forensic settings in the form of quasi ACs (Roberts et al., 2022; Kahneman & Klein, 2009; Sellbom et al., 2022). In the field of law, one review was published, describing how an AC might work in court cases, and another article described how an AC could work in a law firm (Berzon, 2012; Cohen et al., 2000). Even philosophy has already dealt with ACs: In two independent quasi ACs, one paper analyzed a debate of earlier philosophers, and the other paper explored the use of language (Fischer et al., 2021; Leal & Marraud, 2022). For more information about quasi ACs, see Section 3.6, Quasi Adversarial Collaborations.

Interestingly, there seem to be no journals that publish papers on ACs at an increasing rate. Most journals have published only one or two articles on this topic. The *Journal of Applied Research in Memory and Cognition* has the most writings in this regard, followed by *Perspectives on Psychological Science* and *Clinical Psychology: Science and Practice*.

Table 2. Numbers of parties involved in ACs, their research topics, whether a mediator was involved, and degree of agreement achieved by AC parties reached regarding the outcome.

AC	Number of Parties	Topic	Mediator	Degree of Unity ^a
Latham et al. (1988)	2	By what is the engagement to pursue a goal influenced?	E.A. Locke	3
Gilovich et al. (1998)	2	Actions as well as inactions can be regretted. How do these forms of regret develop over time?	-	3
Mellers et al. (2001)	2	To better understand the Linda problem, the effect of frequency representations (e.g., “How many of the 200 women are bank tellers and . . .”) on the violation of the conjunctions rule is examined.	B. Mellers	2
Bateman et al. (2005)	2	Is spending money on public goods experienced as a loss or not?	-	3
Cadsby et al. (2008)	2	What influences the decision to provide grants for a public project?	-	3
Corrigan et al. (2012)	2	How rationally do subjects behave with and without price feedback in an auction?	-	3
Matzke et al. (2015)	2	Can horizontal eye movements facilitate memory retrieval?	M.W. van der Molen	3
Dwyer et al. (2017)	2	How similar are happiness and meaning?	-	3
Van Dessel et al. (2017)	2	Receiving instructions to approach a specific stimulus and avoid another stimulus causes an implicit and explicit preference for the former stimulus. How does this work?	-	3
Kerr et al. (2018)	2	Not every study could replicate the ingroup favoritism effect. Potential moderators of this effect are examined.	-	4
Doherty et al. (2019)	3	What limits the capacities of working memory?	postdoctoral researchers	2
Koch et al. (2020)	2	Discusses the disagreement between two theories of social evaluation.	-	3
Lindner et al. (2020)	2	Does exposure to sexualized female characters in a Tomb Raider game affect women’s body image and their behavior towards other women?	-	4
Stern and Crawford (2021)	2	Testing a model of how ideological conflict and perceived dissimilarities trigger negative emotions and how this leads to prejudice as well as affiliative intentions. Differing opinions on how these factors play out between liberals and conservatives.	-	2

^a 1 = Debate was not solved; 2 = Debate was partially resolved; 3 = Debate was largely solved; 4 = Debate was fully resolved.

3.4. Researchers

It is not clearly determinable how the authors of publications about ACs are connected to each other. There are authors who publish somewhat frequently on this topic, whereas others only briefly touch the subject of ACs in their papers (e.g., [Marewski et al., 2018](#)). The following section focuses on some scientists who have published multiple times on this topic.

Firstly, Nobel Prize laureate Daniel Kahneman (1934–2024) must be singled out; he was involved in three ACs and one quasi AC and, as previously mentioned, coined the term “adversarial collaboration” (Bateman et al., 2005; Gilovich et al., 1998; Kahneman & Anderson, 2003; Kahneman & Klein, 2009; Mellers et al., 2001). In his article, Brockman (2022) referenced a lecture in which Kahneman talks about his experiences and conducted ACs.

Barbara Mellers was involved in Kahneman’s second AC, wherein she held the role of the mediator (Mellers et al., 2001). Similarly to Kahneman’s appeals, she continues to advocate for the implementation of ACs to this day. She is part of the Advisory Board of the “Adversarial Collaboration Projects” and is married to Philip Tetlock, who is the chairman of the research team there (Advisory Board, n.d.; Clark et al., 2022a; Research Team, n.d.).

Other members of this project, such as Cory C. Clark, have published several papers on ACs together with Tetlock, describing the general benefits of ACs and necessary steps to implement them (Clark et al., 2022a, 2022b; Clark & Tetlock, 2022; Costello et al., 2022). However, so far, neither Tetlock nor Clark was involved in a specific AC.

Chris Starmer conducted an AC with Kahneman and later participated in a quasi AC (Alempaki et al., 2019). Jarret T. Crawford, who also published an AC, was part of a later quasi AC and likely will publish another AC (Stern & Crawford, 2021; Reinero et al., 2020; Research, n.d.).

The researchers Dominic M. Dwyer and Jan De Houwer, who participated in two different ACs, wrote an article together in which ACs are mentioned as a way to resolve controversies (Beckers et al., 2016).

The research group of Doherty et al. (2019) published a more detailed description of the AC (Cowan et al., 2020), and similar work has been performed by Danielle Lindner (Lindner & Tribble, 2020; Lindner et al., 2020).

Susan T. Fiske (affiliated with Princeton University, like Kahneman was at some time) has authored a number of topical papers in this area, and in addition to her AC (Koch et al., 2020), she was involved in a larger quasi AC and recorded her experiences in several articles (Abele et al., 2021; Ellemers et al., 2020; Fiske, 2017; Koch et al., 2020).

3.5. Study Quality

The quality of the 14 ACs was assessed based on nine dichotomous features (applies [✓] vs. does not apply [×]). If a feature applied, the AC received a score of 1; if a feature did not apply, it received a score of 0. The scores for each study were then added up to obtain a total score. The studies fulfilled between four and eight features. In addition, the overall quality of the studies was calculated based on the percentage of features that were fulfilled across studies. The 14 AC studies fulfilled a total of 77 of the 126 (9 features times 14 studies) features; hence, the overall quality of the AC studies was 61% (77/126).

These results show that many of the studies only roughly describe the implementation of their specific AC; see Table 3 (e.g., Bateman et al., 2005; Cadsby et al., 2008; Corrigan et al., 2012; Gilovich et al., 1998). For example, it was not always reported whether the study design and the predictions of the researcher had been documented in a protocol (either via preregistration or with a specific AC agreement) before the study was conducted (Kerr et al., 2018; Koch et al., 2020). There are even cases where it is not clear who holds which view, or where the definition of an AC seems underspecified and thus inadequate (Dwyer et al., 2017; Koch et al., 2020; Lancaster & Friedman, 2017; Stern & Crawford, 2021). Despite these incomplete details, some studies still refer to themselves as ACs; here, such studies were summarized under the label quasi AC instead (see Section 3.6, Quasi Adversarial Collaborations). Future researchers should make sure to adequately label their ACs as such.

Table 3. Study quality features of ACs.

AC	Roles	Theory	Communication	Protocol	Decision-Making	Discussion	Definiton	Procedure	Articles	Total Score
Latham et al. (1988)	✓	✓	✓	×	✓	✓	✓	×	×	7
Gilovich et al. (1998)	✓	✓	×	×	✓	✓	✓	×	×	5
Mellers et al. (2001)	✓	✓	✓	×	✓	✓	✓	×	×	7
Bateman et al. (2005)	✓	✓	×	×	✓	✓	✓	×	×	5
Cadsby et al. (2008)	✓	✓	×	×	✓	×	✓	×	×	4
Corrigan et al. (2012)	✓	✓	×	×	✓	×	✓	×	×	4
Matzke et al. (2015)	✓	✓	✓	✓	✓	✓	✓	×	×	8
Dwyer et al. (2017)	×	✓	×	✓	✓	×	✓	×	×	4
Van Dessel et al. (2017)	✓	✓	×	✓	✓	×	✓	×	×	5
Kerr et al. (2018)	✓	✓	×	×	✓	×	✓	×	×	5
Doherty et al. (2019)	✓	✓	×	✓	✓	✓	✓	×	✓	8
Koch et al. (2020)	✓	✓	✓	×	✓	×	×	×	✓	6
Lindner et al. (2020)	×	✓	×	✓	✓	×	✓	×	✓	5
Stern and Crawford (2021)	✓	✓	×	✓	✓	×	×	×	×	4

3.6. Quasi Adversarial Collaborations

Table 4 summarizes all 23 quasi ACs included in our review. It is evident that in most of this corpus of further (“AC-like”) studies, at least two AC conditions are met. Typically, parties with different positions work together and publish the study jointly. This is because the majority of papers of this type are reviews (e.g., [Finkel et al., 2015](#); [Kahneman & Klein, 2009](#); [D. Martschenko et al., 2019](#)). In four of these accounts, different authors with different views discussed the application of various psychological tests in forensic settings ([DeMatteo & Olver, 2022](#); [Marewski et al., 2018](#); [Roberts et al., 2022](#); [Sellbom et al., 2022](#)). These papers were part of a larger project in which other procedures have also been reviewed ([Neal et al., 2022](#)). One comparatively more detailed quasi AC describes how representatives of five different theories of social evaluation met to integrate their theories into a common framework ([Abele et al., 2021](#)). Another study describes how discussions were held with recognized experts, and the points of agreement or disagreement were recorded to better understand the priorities of particular theories and how they work ([Munter et al., 2015](#)).

In other quasi ACs, actual experiments were conducted. The parties came together during or after data collection because of similar hypotheses and study designs and published jointly ([Alempaki et al., 2019](#); [Reinero et al., 2020](#)). One study describes how, in one of several experiments, researchers instructed study participants to perform some form of AC. Although the study participants were against the establishment of universal basic income (UBI), they were asked to think of arguments that they thought might convince other opponents to support the UBI ([Johnson et al., 2023](#)). However, there are also papers in which crucial information (e.g., clear hypotheses) is missing to assess whether it is a proper AC or not ([Becker et al., 2016](#); [Coles et al., 2022](#)). Three quasi ACs have examined topics in transpersonal psychology. In these, the researchers provide an overview of their divergent opinions on various phenomena ([Lancaster & Friedman, 2017](#); [Laythe & Houran, 2022](#); [Robbins et al., 2019](#); [Thouin-Savard, 2017](#)). Many different types of collaborations

seem to fall under the term quasi AC, which, although they do not meet the requirements of a true AC, are referred to as such by the authors (e.g., [Lancaster & Friedman, 2017](#)). One possible reason for such instances of mislabeling might be found in incomplete reports of actual ACs: for example, sometimes only the adversarial character is highlighted as a requirement, whereas the other conditions of ACs (see above, Section 3.5, Study Quality) are not mentioned ([Stern & Crawford, 2021](#)).

Table 4. Summary of the topics covered in quasi ACs, as well as an overview of which and how many of the five AC requirements (feature similarities) were fulfilled.

Quasi AC	Research Topic	Feature Similarity to an AC ^a	Number of Feature Similarities (0–5)
Kahneman and Klein (2009)	A review of attempts to discuss and resolve differences between two approaches of intuition and expertise.	a; e	2
Marewski et al. (2011)	Overview of different viewpoints on recognition-based judgments.	a; e	2
Finkel et al. (2015)	Does familiarity promote interpersonal attraction?	a; e	2
Munter et al. (2015)	How should mathematics be taught?	a; e	2
Becker et al. (2016)	Is selective attention enough to produce automatic semantic priming effects, even in the absence of direct response competition?	a; d	2
Lancaster and Friedman (2017)	Discussion about the role of transpersonal psychology in the academic world (in relation to perennialism and essentialism approaches).	a	1
Steinert et al. (2017)	Comparison of the effectiveness of CBT, psychopharmacology, and psychodynamic therapy.	a; b; e	3
Joel et al. (2018)	Reports complications that can arise with open sharing of data and suggests ways to overcome them.	a; e	2
Alempaki et al. (2019)	Examination of contextual effects in decision-making.	a; d; e	3
D. Martschenko et al. (2019)	A review of how the increasing knowledge of molecular genetics could be relevant for education.	a; e	2
Robbins et al. (2019)	Overview of the perspectives of phenomenology and possible solutions to the problem of the subject/object dualism in transpersonal psychology.	a; e	2
Crusius et al. (2020)	Comparison of competing definitions of envy.	a; e	2
Reinero et al. (2020)	What impact does political ideology have on the quality of scientific articles?	a; e	2

Table 4. Cont.

Quasi AC	Research Topic	Feature Similarity to an AC ^a	Number of Feature Similarities (0–5)
Abele et al. (2021)	Integration of various theories on social evaluation into a common framework.	a; b; e	3
Fischer et al. (2021)	Examination of conditions under which competent language users make inappropriate stereotypical inferences that persistently influence further cognition.	a; c; d; e	4
Coles et al. (2022)	Testing the facial feedback hypothesis (the notion that individuals' subjective experience of emotion is influenced by their facial expressions), because it has not always been consistently replicated in the literature.	a; d; e	3
DeMatteo and Olver (2022)	Review of the application of the Psychopathy Checklist-Revised (PCL-R) in legal contexts.	a; e	2
Laythe and Houran (2022)	Capturing the best available evidence for life after death.	a; e	2
Leal and Marraud (2022)	Analysis of the Russell–Copleston debate of 1948 (arguments about the existence of God) from the point of view of two different argumentation theories.	a; e	2
Roberts et al. (2022)	Examination of the usefulness of the TSI-2 and TSI in forensic psychology.	a; e	2
Sellbom et al. (2022)	Review of the Millon Inventories for forensic applications.	a; e	2
Viglione et al. (2022)	Does the Rorschach test meet the legal standards for admissibility as evidence in court?	a; e	2
Johnson et al. (2023)	Investigates whether universal basic income (UBI) is unappealing to socially conservative voters.	-	0

^a a = minimum two research groups with different explanations for the same phenomenon; b = development and documentation of a joint study design; c = formulation of mutually exclusive hypotheses; d = testing of hypotheses through an experiment; e = joint publication of results, regardless of the outcome.

It goes without saying that this mixing of different forms of collaboration should be avoided in future research along these lines. It is therefore recommended to use more specific designations. For example, the designation *adversarial integration* might be used for collaborations that, although they do not conduct an experiment, at least aim to synthesize their different views into one shared concept ([Abele et al., 2021](#); [Laythe & Houran, 2022](#)). If there are different explanations for a research question on which different parties write an overview, the term *collaborative analysis* would be suitable ([Crusius et al., 2020](#); [DeMatteo & Olver, 2022](#); [Haaga, 2009](#); [Marewski et al., 2011](#)). Cases where researchers with different expertise collaborate to address a research question from all sides but do not test mutually exclusive hypotheses could alternatively be labeled *multiteam collaborations* ([Haaga, 2009](#); [Steinert et al., 2017](#)).

3.7. Best-Practice Recommendations

The following sections describe the steps involved in conducting an AC, with various descriptions of the necessary aspects, as—in the course of our systematic mapping review of the AC literature—compiled from selected articles and summarized by us into a coherent guideline, with examples provided. The information compiled and integrated here is based on actual AC-type collaborations, as well as on references and additions from other sources that were not ACs themselves. This was necessary because not all AC studies clearly report their implementation steps. Additionally, at the end of this section, we provide a brief overview of the ACs conducted so far (see Table 2). The best-practice recommendations were divided into 16 implementation steps for a planned AC:

- (1) Key concept.
- (2) Role of institutions and journals.
- (3) Initial contact between potential parties.
- (4) Prerequisite for participating scientists.
- (5) Appointment of a mediator.
- (6) Fostering teamwork and friendship.
- (7) Choosing a representative and distributing roles.
- (8) Comparing studies and defining discrepancies.
- (9) Formulating hypotheses and designing the study.
- (10) Drafting an Adversarial Collaboration Agreement.
- (11) Obtaining feedback and registration of the research plan.
- (12) Conducting the AC study.
- (13) Writing the article.
- (14) Formal aspects of the article.
- (15) Final submission and role of academic journals.
- (16) Failed AC.

These 16 steps are described in more detail below.

(1) *Key Concept*

If a researcher wants to experimentally refute a publication because it contradicts their own theory, a collaboration with the proponent of the opposing model should be considered (Alempaki et al., 2019; Bateman et al., 2005; Cadsby et al., 2008; Cowan et al., 2020; Doherty et al., 2019; Mellers et al., 2001; Nier & Campbell, 2013; Rakow et al., 2015; Melloni, 2022; Tetlock & Mitchell, 2009c). Disputes can arise not only because of competing theories, but also because of differing methods, measurements, perceived status of the phenomenon, or interpretations (Clark et al., 2022a; Rakow et al., 2015).

(2) *Role of Institutions and Journals*

ACs are very time-consuming, so universities and journals can play a central role in initiating them. Universities should, in hiring and promotion decisions, favor researchers who have participated in ACs and call on funding agencies to support such forms of collaboration (Tatlidil & Sloman, 2022; Vlasceanu et al., 2022). Journals could dedicate special issues or sections to studies in the form of an AC (Cowan, 2022; Vlasceanu et al., 2022). Editors should be aware of the strengths and weaknesses of ACs and evaluate them differently from traditional study designs, for example, by not expecting bold, far-reaching conclusions (Cowan, 2022).

(3) *Initial Contact Between Potential Parties*

An AC can consist of two or more parties (see Table 2). There can be different reasons to determine how and which research groups come to the decision to conduct an AC. A meeting at a scientific conference or workshop or through a guest lecture at a university can

present an occasion for an initial conversation on this topic (Ellemers et al., 2020; Gruber, 2022; Kerr et al., 2018; Koch et al., 2020; Makel et al., 2019; Tetlock & Mitchell, 2009a).

(4) *Prerequisite for Participating Scientists*

The adversarial character of an AC is based on the opposing viewpoints of a study; therefore only researchers who have different views of the same phenomenon should collaborate (Berzon, 2012). This way, the bias of the other party can be perceived and countered (Berzon, 2012; Beckers et al., 2016). The differences should not be too strong on a philosophical or meta-level, because that could, for example, complicate the operationalization of variables (Latham et al., 1988).

Curiosity, trust, willingness to compromise, and respectful interaction should form the basis of every collaboration (Latham et al., 1988; Makel et al., 2020; Matzke et al., 2015). Strong dogmatism is counterproductive, and methodological criticism must be separated from personal attacks (Clark et al., 2022a; Ellemers et al., 2020; Fiske, 2017; Rakow et al., 2015). The parties should be aware that one or even both may be wrong and that a single study in all likelihood will not be sufficient to resolve disputes (Cadsby et al., 2008; Matzke et al., 2015; Mellers et al., 2001; Nier & Campbell, 2013; Latham et al., 1988; Rakow et al., 2015). The goal of ACs should be to solve a problem and not to harm each other (Ellemers et al., 2020; Fiske, 2017; Makel et al., 2019).

(5) *Appointment of a Mediator*

If necessary, research groups can ask a neutral person to mediate in case of differences (Cowan et al., 2020; Makel et al., 2019; Mellers et al., 2001). See Section 3.8 (The Role of the Mediator) for more details.

(6) *Fostering Teamwork and Friendship*

To facilitate the collaboration, it is important to create positive experiences and to define a common identity. A fun activity, such as a shared dinner after each workday, gives participants an opportunity to relax together. This contributes to maintaining a positive atmosphere and makes the experience worthwhile (Ellemers et al., 2020).

(7) *Choosing a Representative and Distributing Roles*

To ensure equal collaboration away from the eyes of a wider audience, it is recommended to choose a representative for each theory. These representatives can then meet in a neutral location, such as a hotel, to plan their next steps. The other participants should be entrusted with specific tasks that correspond to their expertise, such as knowledge of accessing specialized literature, familiarity with specific statistical methods, access to resources for support, and so forth (Ellemers et al., 2020). Throughout the process, regular communication, as well as personal or digital meetings, are essential to promote trust and avoid misunderstandings (Cowan et al., 2020; Ellemers et al., 2020).

(8) *Comparing Studies and Defining Discrepancies*

The representatives of the different theories compare their previous research results to formulate precise hypotheses that confirm specific predictions of a theory and simultaneously reject those of the others (Costello et al., 2022; Ellemers et al., 2020; Koch et al., 2020; Makel et al., 2019; Matzke et al., 2015). Discrepancies between the models can pertain to methods, statistics, transparency, accessibility, or slightly deviating instructions for the study participants (Ellemers et al., 2020; Latham et al., 1988).

To provide a better overview, differences between previous study designs can be compared with each other in a tabular format (Kerr et al., 2018). Alternatively, each representative could compile a list of articles that support their own viewpoint and answer predefined questions about the origins, premises, methods, and evidence for their theories.

This way, a concise overview of the theories and research design can be created, and both models are brought into a congruent format (Ellemers et al., 2020). Disagreements should not be addressed at this stage; instead, the explanation should be given on a low and precise level of abstraction, explaining why certain approaches were chosen (Ellemers et al., 2020; Matzke et al., 2015).

(9) *Formulating Hypotheses and Designing the Study*

Once the differences between the competing theories are clearly outlined, researchers formulate hypotheses that would confirm the views of one party and simultaneously would refute the prediction of the other party (Bateman et al., 2005; Costello et al., 2022; Latham et al., 1988). The focus should be on refuting a single phenomenon and not on disproving an entire theory, as this is hardly achievable by a single study (Cowan et al., 2020; Mellers et al., 2001). The hypotheses should be specific enough to prevent generalizations. Questions such as “How does priming affect behavior?” are more suitable than “What is gender?” (Tullett, 2022). The parties subsequently agree on a study design. The methods chosen to test the hypotheses should be designed to convince a broad audience, for example, through multiple operationalizations of the most important variables, including broader latent constructs (Lindner & Tribble, 2020; Vlasceanu et al., 2022). Enough time should be allowed for this step so that good conclusions can be drawn from the different viewpoints and data, and no hasty solutions are cobbled together under pressure (Costello et al., 2022; Ellemers et al., 2020; Gruber, 2022).

(10) *Drafting an Adversarial Collaboration Agreement*

The agreement on hypotheses and the study design are recorded in a protocol (Berzon, 2012; Cadsby et al., 2008; Matzke et al., 2015; Mellers et al., 2001). Akin to study preregistration and clinical trial protocols, an *Adversarial Collaboration Agreement* (AC agreement) or *Outline Collaboration Proposal* is created before the study is conducted and includes all agreements between the parties regarding hypotheses, study design, planned analyses and methods, and all other foreseeable aspects of the study (Clark et al., 2022a; Makel et al., 2019; Matzke et al., 2015; Rakow et al., 2015). The agreement describes all possible interpretations and which results correspond to or contradict the explanation of the respective party. Additionally, it includes a promise that the parties will publish the study regardless of the outcome, as well as a plan for communication between the parties (Bateman et al., 2005; Ellemers et al., 2020; Matzke et al., 2015; Rakow et al., 2015; the last reference describes the structure of the AC agreement in further detail). Explicit AC agreements help to avoid bias and future disagreements over interpretation and increase transparency. They can be drafted by the researchers or a mediator (Bateman et al., 2005; Berzon, 2012; Cadsby et al., 2008; Matzke et al., 2015; Mellers et al., 2001).

Only one (Matzke et al., 2015) out of the 14 ACs reviewed here contained an AC agreement. Many others preregistered their study instead. Some did not even mention a protocol (see Table 3). Since an AC agreement is more detailed than a preregistration and strongly contributes to transparency, it is recommended for future studies to create one or at least indicate that (and justify why) none is mentioned in the study.

(11) *Obtaining Feedback and Registration of the Research Plan*

The AC agreement is reviewed by coworkers who are not directly involved in the process in order to identify and address contentious issues or inconsistencies. It is then sent to a selected journal to receive potential improvement suggestions (Ellemers et al., 2020). If the editors approve the research questions and methods, they should agree to publish the study independently of the results, similar to the format of a Registered Report (Vlasceanu et al., 2022). To save time, it is recommended to create a general version of

the AC agreement first and then refine it in more detail after it has undergone a first peer review (Rakow et al., 2015).

(12) *Conducting the AC Study*

After finalizing study planning and approval of the AC agreement, the study can begin (Doherty et al., 2019). Data collection can be carried out by the parties themselves, research assistants, or, if involved, the mediator (Latham et al., 1988; Lindner & Tribble, 2020; Rakow et al., 2015). If the data are collected in different laboratories and/or universities, the parties monitor data collection in their own institution. Care should be taken to ensure that the load of data collection is evenly distributed (Doherty et al., 2019; Rakow et al., 2015). If a mediator is involved, they should have control over the data to publish, should one party refuse cooperation during the course of the study (Mellers et al., 2001).

During the data collection process, the AC agreement can be changed if both parties agree. For example, the stopping criteria and the strategy for recruiting participants were slightly modified in one of the studies reviewed here (Matzke et al., 2015; Rakow et al., 2015).

Once all collected data have been compiled, data analysis can begin. Post hoc (i.e., exploratory, supplemental) analyses are allowed if they are marked as such (Rakow et al., 2015). The agreed-upon deadlines should be met (Mellers et al., 2001). The allotted time for data collection and manuscript submission should be limited to a maximum of twelve months (Rakow et al., 2015).

(13) *Writing the Article*

After analyzing the data, the paper is written (Mellers et al., 2001). Who writes which sections is regulated in the AC agreement; for example, it can be the mediator's task to care for the write-up of the paper (Ellemers et al., 2020; Latham et al., 1988; Marewski et al., 2011).

Since ACs are not yet common research practice, a clear definition of the AC should be included in the text. The process of the AC should be carefully documented, i.e., which methods worked, how the communication process went, whether one party refused to cooperate during the study, as well as any other unexpected occurrences (Cowan et al., 2020; Mellers et al., 2001). Future ACs definitely benefit and learn from such documentation and will run more smoothly (Cowan et al., 2020).

A single AC is unlikely to end a scientific debate, so the points of disagreement between the parties should be explicitly listed at the end, which then can form the basis for further research (Makel et al., 2019; Matzke et al., 2015). See Table 2 for an overview of how much agreement there was among the AC research groups reviewed here in interpreting their respective results. Each party should also be free to propose an additional experiment to make use of the new findings (Clark et al., 2022a; Mellers et al., 2001).

If disagreements persist despite best efforts, each party can present their interpretations in their own section of the resulting paper (Latham et al., 1988; Makel et al., 2019; Marewski et al., 2011; Matzke et al., 2015; Mellers et al., 2001; Rakow et al., 2015). Space and length are predetermined for summarizing individual views and points of persisting disagreement. For example, Rakow et al. (2015) recommend that these sections should not exceed 20% of the word-count limit of the paper's discussion section (Costello et al., 2022; Ellemers et al., 2020). Further material, exceeding those limits, can be provided in a footnote (or, nowadays, in a supplemental file, or in external files openly accessible via research repositories, such as the Open Science Framework) to give readers the opportunity to obtain additional in-depth material (Mellers et al., 2001).

(14) *Formal Aspects of the Article*

Word choice: The researchers should not be referred to as supporters or critics of a model, as this can shift the focus from the models to the individuals (Marewski et al.,

2011). Authorship: The mediator is usually the first or last author, and the parties are listed in random order. If a party consists of more than one person, the authors are listed in order of contribution. Alternatively, names can be listed alphabetically (Rakow et al., 2015); nowadays, utilizing taxonomies of (paper) contributorship roles is recommended. Publication of materials: It is recommended to make data and materials publicly available to enable further research (Tullett, 2022).

(15) *Final Submission and Role of Academic Journals*

After the article is written, it is submitted to the journal. If there are deviations from the study protocol, an additional document is included, explaining why there were any changes in the experimental design. The manuscript is then subjected to peer review and revised if necessary (Rakow et al., 2015). Editors of academic journals should be aware that extensive revision requests after submission of the article can cause problems for the research groups, owing to the specifics of ACs (Mellers et al., 2001).

(16) *Failed AC*

If an AC fails to reach agreement, it would be beneficial for the research community at large when a brief report could be published to give interested readers insight into the differences in opinion that led to the dissolution (Ceci & Williams, 2022; Neal et al., 2022).

3.8. *The Role of the Mediator*

Sometimes, a neutral person, called the mediator (alternatively, arbiter or referee), is involved in an AC (Haaga, 2009; Matzke et al., 2015; Nier & Campbell, 2013). Interestingly, in the 14 examined ACs, a mediator was explicitly mentioned only four times (Doherty et al., 2019; Latham et al., 1988; Matzke et al., 2015; Mellers et al., 2001; see Table 2 for details). It appears that over time, the role of the mediator may well have become less significant due to the increasing popularity of preregistration.

3.8.1. *Requirements of the Mediator*

The mediator is an impartial person who has not committed to any of the competing views (Cowan et al., 2020; Makel et al., 2019; Mellers et al., 2001; Nier & Campbell, 2013). All parties should agree on the appointment of the mediator (Clark et al., 2022a; Leal & Marraud, 2022). The mediator must be familiar with the theories and should demonstrate a fair and effective way of working so that constructive solutions for disagreements, such as those related to statistical procedures, can be found (Makel et al., 2019; Nier & Campbell, 2013). Both outsiders and in-house personnel, like research assistants, can take on this role (Cowan et al., 2020).

3.8.2. *Responsibilities of the Mediator*

The mediator can be asked by research groups to participate in an AC, or they can be the one who initiates an AC by asking competing parties to participate in a joint study (Nier & Campbell, 2013). The mediator works with the parties on the design and methods for data collection, and they may be responsible for analysis and the results section of the study. Additionally, they provide impartial advice, monitor compliance with the AC agreement, and approve or decide on changes to the original action plan (Clark et al., 2022a; Rakow et al., 2015). At the outset, it should be agreed that the mediator will seek to publish the study even if one or more participants refuse to cooperate and withdraw (Clark et al., 2022a). They ensure that all parties can express their opinions freely, which is especially important for newly promoted researchers (Marsh, 2022). In other studies, mediators are not involved in data collection but receive weekly updates (Matzke et al., 2015). In some ACs, the mediator is not mentioned at all (e.g., Bateman et al., 2005; Corrigan et al.,

2012; Van Dessel et al., 2017). The tasks of the mediator differ depending on the degree of involvement.

The involvement of a mediator is not an essential requirement for an AC. There are no reports of collaborations failing due to the lack of a neutral person, so it is difficult to estimate how important the involvement of a mediator actually is (Rakow, 2022).

3.9. Advantages of Conducting an AC

A list of 18 advantages was extracted in the course of our review of the existing AC literature, which are listed and further elaborated on in the following:

- (1) Collaboration of competing parties.
- (2) Fair research setting.
- (3) Emotionally liberating.
- (4) Fostering collegiality.
- (5) Efficient competition of ideas
- (6) Increased accumulation of knowledge in a shorter time.
- (7) Maximization of expert knowledge.
- (8) Deep understanding of other theories.
- (9) Identification of limitations of one's own theory.
- (10) Improvement of one's own theory.
- (11) Identification of contextual factors.
- (12) Ongoing optimization of the study design.
- (13) Basis for future studies.
- (14) Reduction of bias.
- (15) Increase in transparency.
- (16) Better overview of theories for readership.
- (17) Standardization of methods.
- (18) Learning from failed collaborations.

The following section provides more detailed explanations of these advantages. For a short overview of all identified benefits and drawbacks of an AC and possible solutions, refer to Table 5.

Table 5. Overview of the benefits and drawbacks of an AC and possible solutions.

Advantages of ACs	Disadvantages and Challenges of ACs	Suggested Solutions
Working with opposing parties; helpful in areas with years of dispute	Time-consuming and labor-intensive, thus resulting in a lower publication rate	Increased promotion by institutions (journals, universities, etc.) (Tatlidil & Sloman, 2022)
Fair setting to resolve scientific disputes	Aversion against interpersonal conflict; pride is assaulted	No cooperation with adversaries who are temperamental or dogmatic; politeness on both sides (Clark et al., 2022a)
Emotionally liberating	Power imbalance due to different career levels; young researchers are more restricted in their research freedom	Mediator ensures freedom of speech and equality (Marsh, 2022)
Collegiality is cultivated, trust is built; provides enjoyment	Career risk as theory can be partially disproved; leads to loss of prestige	More open-minded view: updating beliefs is progress, not failure (Melloni, 2022)
Increased competition of ideas with transparent rules	Hard to prove independence; affects young scientists in particular	Faculty could treat all authors of an AC as joint first authors to honor effort (Marsh, 2022)

Table 5. Cont.

Advantages of ACs	Disadvantages and Challenges of ACs	Suggested Solutions
Through close collaboration, faster gain of knowledge than in traditional research	Mostly reduced to two views (coming from the same subject area)	Initialize larger research teams through institutionalized efforts, such as the journal <i>Comprehensive Results in Social Psychology</i> (Tullett, 2022)
Maximization of expert knowledge resulting in more comprehensive understanding of the phenomenon	Not filed against socio-political asymmetry of science; Republicans, women, young researchers, and unknown universities are less represented	Larger research teams (see above; Tullett, 2022); blind applicants for an AC, until agreed to; developing lab guidelines for these requests (Marsh, 2022)
Focus on real differences in the debate, rather than ideological or political differences	Socially controversial topics may be addressed less frequently	Difficult—apparently, so far none specified
Recognition of limitations of different theories; better resource management	Not appropriate for all research questions, e.g., exploratory work, lack of competing hypothesis, moral concerns, ideological conflicts	Respect limitations of ACs; they should not be praised as a cure for all kinds of scientific problems
Theories improve faster, become more similar; ideally, a new theory is formed	A single experiment cannot reject a theory; theories can be adjusted to fit results	Science/society must accept that progress is mostly achieved by improving a theory
Errors are detected early; study can be improved directly	Differentiation of theories: differences arise from different levels of explanation	Difficult—apparently, so far none specified
New findings provide basis for new studies	Dependence on explicitly contradictory hypotheses: blowing up small differences, less rigorous tests and thus more flexible results pose danger	Peer-review of the AC agreement by the journal (Rakow et al., 2015); control through the mediator
Reduction of many biases (e.g., hindsight, selection, or confirmation bias)	Disagreement between parties on research design	Mediator can help find compromises (Matzke et al., 2015)
High transparency thanks to AC agreement	Communication problems due to language barriers, time zones, etc.	Video conferencing, English language skills as a prerequisite
More credibility towards the readership; people from outside the field get a better overview of theories	Exact replications of experiments rarely possible	Involvement of authors of original study (Kerr et al., 2018)
Best possible standardization of methods; experimental designs are operationalized	Success (high citation numbers, high-impact journals, etc.) dependent on results	Difficult—apparently, so far none specified
Failed ACs can be an indication of vague claims, different phenomena, or ambiguous methods; improvement of research tools	Journal audiences and readership are not used to the rarely clear outcomes of ACs	Difficult—apparently, so far none specified
New awareness of the strong impact of unnoticed small contextual factors in research	Not every initialized AC is performed, may increase conflict between parties	Refraining from collaborating with very intolerant researchers; learning from failures (problems arose from ambiguous methods, vague language, etc.) (Clark et al., 2022a)

(1) *Collaboration of Competing Parties*

Collaboration between parties with different theories about the same phenomenon is made possible (Corrigan et al., 2012; Gilovich et al., 1998). Even if opinions differ greatly, a mediator can take on a mediating role between positions (Corrigan et al., 2012; Kerr et al., 2018; Latham et al., 1988; Matzke et al., 2015). Such collaborations are particularly helpful in areas where arduous disputes drag on for years, such as in education (Makel et al., 2019). If the opponents are facing each other personally, they tend to represent a more moderate view than they would against an imaginary opponent (Clark et al., 2022a).

(2) *Fair Research Setting*

AC is likely the fairest way to resolve a dispute between scientists (Nier & Campbell, 2013), as it provides the opportunity to resolve important research questions in a transparent process in which all sides have equal say (Clark et al., 2022a; Nier & Campbell, 2013).

(3) *Emotionally Liberating*

It can be emotionally liberating for researchers to pursue the answers to a question out of pure intellectual curiosity, regardless of the results. This neutrality can provide a break from the stress that arises from the emotional involvement in the results of classic studies (Haaga, 2009). The execution is described as exciting and enlightening, as science in its purest form (Gilovich et al., 1998; Latham et al., 1988).

(4) *Fostering Collegiality*

It is variously mentioned that the researchers participating in ACs had fun and enjoyed their work (e.g., Mellers et al., 2001). This is ensured through close teamwork, which allows for developing an understanding of contrary views, building mutual trust, and cultivating friendships between scientists (Beckers et al., 2016; Ellemers et al., 2020; Finkel et al., 2015; Kerr et al., 2018; Mellers et al., 2001). This not only improves the research process but also is perceived as entertaining (Kerr et al., 2018; Matzke et al., 2015).

(5) *Efficient Competition of Ideas*

The AC provides a stricter competitive setting with more transparent rules than classic studies. Bad ideas are discarded more quickly, while good ideas are clearer and more refined (Clark et al., 2022a, 2022c). As a result, uncertainties and contradictions are reduced in published literature, and better hypotheses are established (Clark et al., 2022a).

(6) *Increased Accumulation of Knowledge in a Shorter Time*

Scientific ideas are often exchanged with a time delay (Tetlock & Mitchell, 2009b). Peer reviews can make an important contribution to the debate, but they slow down the pace of publication, thus, in turn, slowing down the accumulation of knowledge (Rakow et al., 2015). In an AC, novel evidence is often obtained which neither party expected, and so the gain in knowledge can be faster and more satisfying (Clark et al., 2022c; Doherty et al., 2019; Lindner & Trible, 2020; Matzke et al., 2015; Mellers et al., 2001; Vlasceanu et al., 2022).

(7) *Maximization of Expert Knowledge*

The parties benefit from each other's knowledge, perspectives, and concerns that influence the research process. This way, they not only become aware of possible biases and systematic flaws in their own theories but also maximize the likelihood of designing a study that can resolve their differences (Clark et al., 2022b, 2022c; Ellemers et al., 2020; Makel et al., 2019; Matzke et al., 2015). If the authors from an original study participate, the experiments can be better replicated (Kerr et al., 2018). Not only participants but also the field as a whole gains a more comprehensive and richer understanding of the phenomenon through the different explanatory approaches (Edlund et al., 2022).

(8) *Deep Understanding of Other Theories*

Instead of getting bogged down in political or ideological disputes that ultimately only lead to divisions, stifle progress, and suppress fruitful debates, the focus is on the actual differences in scientific discourse (Bateman et al., 2005; Fiske, 2017; Nier & Campbell, 2013). In an AC, it must be explicitly stated what each party believes and what they do not. The emphasis is on true differences, not on merely perceived differences in opinion (Bateman et al., 2005; Finkel et al., 2015; Nier & Campbell, 2013).

(9) *Identification of Limitations of One's Own Theory*

Both parties become aware of the limitations of their own theories (Latham et al., 1988; Mellers et al., 2001). Their own hypotheses, arguments, and conclusions are put to the test to see whether they can withstand the criticism of other researchers (Berzon, 2012; Clark et al., 2022a, 2022c). As just one example, biasing allegiance effects in psychotherapy research can be prevented through collaboration of scientists from different psychotherapy schools (Leichsenring et al., 2017).

Additionally, identifying possible limitations can clarify which conditions need to be met for interventions to be effective, leading to improved implementation and resource management (Makel et al., 2019). On average, researchers are more open-minded and expect more empirical consistency than laypeople, which is why they should be more willing to consider new compelling data, even if these contradict their own pet theories and pet hypotheses (Reinero et al., 2020; Beckers et al., 2016).

(10) *Improvement of One's Own Theory*

The rigorous experimental design and high quality of the collected data lead to more reliable data and thus empirically better supported theories (Bateman et al., 2005; Cowan et al., 2020). By recognizing the limitations of a theory, its applicability is also questioned in other situations and, if necessary, improved (Cowan et al., 2020). New evidence gradually changes theories and makes them more similar to each other. This happens automatically by taking the joint data into account, even if these partially question one's own theory (Clark et al., 2022a; Cowan et al., 2020; Lindner & Tribble, 2020). In addition, it is possible that certain aspects of the other party's theory are useful for one's own explanations (Cowan et al., 2020; Finkel et al., 2015; Nier & Campbell, 2013). Ideally, this adaptation and improvement would lead to a new theory (Cowan et al., 2020). ACs also provide a more transparent overview of the changes to and modifications of a theory (Melloni, 2022).

(11) *Identification of Contextual Factors*

The evident effects of contextual factors are made clear, which scientists are not always aware of, because they themselves are part of the context. Latham et al. (1988), for example, describe how the scientists participating in their AC were impressed by the number of differences in their study procedures, even though both parties were studying the same phenomenon. Even small differences can add up and have an impact on the phenomenon under scrutiny. While it is true that such nuances could also be reported in traditional studies, they are often restricted by the limited word count of journal manuscripts and thus lack the space to add these details. The above-mentioned advantages of ACs offer a solution to this problem (Latham et al., 1988).

(12) *Ongoing Optimization of the Study Design*

ACs are better than mere peer reviews of traditional research, as errors are not only pointed out by overworked and underpaid reviewers after the study is completed. Pre-registrations serve a similar purpose, but ACs go further by integrating criticism into the research process from the beginning to the end (Clark et al., 2022a).

(13) *Basis for Future Studies*

The new, mostly unexpected findings from an AC lead to further questions, assumptions, and hypotheses. These, in turn, require further experiments and drive the scientific debate forward (Cadsby et al., 2008; Cowan et al., 2020; Doherty et al., 2019; Finkel et al., 2015).

(14) *Reduction of Bias*

Hypotheses and expectations must be precisely formulated in advance (Bateman et al., 2005). The parties choose appropriate procedures that confirm either only their own theory or the theory of the other party (Bateman et al., 2005; Berzon, 2012; Reinerio et al., 2020). The research groups monitor each other and question all decisions (Berzon, 2012; Clark et al., 2022c). This prevents hindsight bias, selection bias, confirmation bias, and groupthink (Bateman et al., 2005; Makel et al., 2020), whilst increasing validity and explanatory power (Berzon, 2012). The meticulous planning also leads to fewer post hoc explanations than in typical study procedures. Regardless of the outcome, at the end of the study, the results are published, even if they contradict one's own theory (Bateman et al., 2005).

(15) *Increase in Transparency*

The AC agreement, disclosure of experimental procedures, and statistical analyses increase the transparency of the investigation and guarantee a purely confirmatory nature of hypothesis testing (Matzke et al., 2015). Unexpected or unwanted results are openly addressed, instead of being dismissed as lack of accuracy, or hidden altogether (Bateman et al., 2005; Makel et al., 2019).

(16) *Better Overview of Theories for Readership*

ACs set different expectations for the scientific community. They often do not provide clear-cut results, and the parties are often in disagreement regarding their interpretations, so the readers must draw their own conclusions (Bateman et al., 2005). Experts who are not committed to either viewpoint can use this to form a more informed opinion about the phenomenon and the effect of contextual variables, based on the obtained data and related claims (Cowan et al., 2020; Melloni, 2022). Additionally, science gains more trust and respect from non-experts, because ACs not only collect data and provide a clear explanation but also reconcile facts, synthesize them, and point out the different interpretations (Clark et al., 2022c; Ellemers et al., 2020).

(17) *Standardization of Methods*

Experimental methods are standardized as best as possible, because the proponents of the different theories must agree on experimental plans and operationalizations of variables (Melloni, 2022).

(18) *Learning from Failed Collaborations*

If an AC fails (e.g., because scientists cannot agree on methods, or because the results lead to a dispute), lessons can be learned from the failures. If participants cannot agree on reasons why they disagree, there may not be any at all. Perhaps the adversaries are using the same vague language to describe two different phenomena, or vice versa (cf. jingle-jangle fallacies; Hanfstingl et al., 2024). A lack of agreement on methods may well indicate a lack of conceptual clarity. More precise statements from each side could resolve the disagreement (Clark et al., 2022a).

3.10. *Potential Difficulties of ACs*

A list of 18 disadvantages was extracted in the course of our review of the existing AC literature, which are listed and further elaborated on in the following:

- (1) Time-consuming and labor-intensive.

- (2) Aversion to interpersonal conflicts.
- (3) Status differences.
- (4) Occupational risk.
- (5) Demonstration of autonomy is difficult.
- (6) Not all sides are illuminated.
- (7) Sociopolitical asymmetry in science.
- (8) Societally controversial topics are less addressed.
- (9) Not all questions find an answer.
- (10) One experiment alone cannot reject a theory.
- (11) Differentiation of theories.
- (12) Dependency on clear contradictory hypotheses.
- (13) Disagreement between parties regarding research design.
- (14) Communication problems.
- (15) Replication problems.
- (16) Depends on results.
- (17) Other expectations of the readers.
- (18) Not every initiated AC is conducted.

The following section offers more detail about these challenges for ACs. For a concise overview of all the benefits and drawbacks of an AC and possible solutions for the latter ones, please refer to Table 5.

(1) *Time-Consuming and Labor-Intensive*

AC is labor-intensive and often time-constrained and requires more planning and anticipation than standard research (Marsh, 2022; Melloni, 2022; Vlasceanu et al., 2022). It can take months to several years to complete (Ellemers et al., 2020; Lindner & Tribble, 2020). Not only must there be agreement on the experimental design, but all foreseeable aspects of data collection and analysis should also be documented (Ellemers et al., 2020; Matzke et al., 2015). The publication rate may decrease, potentially reducing one's chances of obtaining a permanent position in academia. Even if the review process of the finished study is faster, it often does not alleviate these, and related, concerns (Marsh, 2022).

(2) *Aversion to Interpersonal Conflicts*

Interpersonal conflicts among scientists are often avoided, because they can lead to bitterness and long-term discomfort, for example, at conferences or in the context of other professional activities (Clark et al., 2022a). It is draining to be subjected to criticism from other colleagues (Cowan, 2022; Ellemers et al., 2020; Haaga, 2009; Vlasceanu et al., 2022). Therefore, hostility between opponents, or outright lack of curiosity about alternative viewpoints, hinders such a collaboration (Clark et al., 2022c; Ellemers et al., 2020; Latham et al., 1988). Frustration can arise when one party feels that their opponent is focusing on minor issues that seem to delay the progress, takes a long time to respond, or questions the neutrality of the mediator. The impression could be created that one party is putting in more effort than the other, which, in a worst-case scenario for an intended AC, could lead to complete withdrawal from the project (Tatlidil & Sloman, 2022; Vlasceanu et al., 2022).

(3) *Status Differences*

It can be challenging to balance differences in academic experience, status, and power. Each career stage brings different concerns, and scientists, regardless of their status, rely on others to recognize the validity of their ideas and appreciate their contributions. Employees may feel offended if they feel not adequately involved in the work or that their results are not presented appropriately (Ellemers et al., 2020).

Junior researchers feel less academic freedom than more senior colleagues who might hold a professorship and thus are less likely to contradict them, which could potentially affect their research independence (Marsh, 2022). Not all ACs that begin are completed. For younger researchers, this can pose a problem if they declare themselves as an “opponent” of an established researcher, but no collaboration ensues (Marsh, 2022; Melloni, 2022).

(4) *Occupational Risk*

AC can be a professional risk, especially for mediators, as they invest time and effort in a study that is not directly related to their ongoing empirical work. Particularly for new or very ambitious researchers, involvement in an AC can lead to a diversion of resources that conceivably could be put to better use elsewhere (Haaga, 2009). Many scientists focus in their research on one specific theory or intervention, often with career-long experience and tremendous investment. This eventually leads to presentations, publishing books, grants, awards, and a multitude of publications involving numerous colleagues (Clark et al., 2022a; Cowan et al., 2020; Makel et al., 2019). If parts of a theory are called into question by joint research, the researchers are at risk of losing prestige (Latham et al., 1988; Makel et al., 2019).

(5) *Demonstration of Autonomy is Difficult*

To find out whether a researcher is the leader of their own research, one can check whether they are named as corresponding (and/or first or last) author of a publication. Especially for young researchers, publications with a mediator may be considered less desirable, as the mediators are sometimes named first (as well as sometimes being the corresponding author), while the “opposing” parties are listed as co-authors. That way, the heuristics for independence, as demonstrated by corresponding (and/or first) authorship, are lost (Marsh, 2022; Mellers et al., 2001).

Another proof of intellectual independence is the writing of articles without senior colleagues. Newcomers to the profession must show that they can do it alone as well. If a young researcher’s theory is compared to that of an established researcher in an AC, the resulting joint publication lacks evidence of autonomy (Marsh, 2022).

(6) *Not All Sides are Illuminated*

An AC is meant to illuminate a question from two sides. While already informative, it is not the same as showing all perspectives. Especially in the case of controversial social issues, such as research on differences between women and men, there is a risk that debates will not be resolved, but rather reduced to polarizing pairs. In addition, someone with a degree in psychology may have different approaches than a biologist or a historian (Ceci & Williams, 2022; Tullett, 2022). If ACs are touted as objective decision-makers in big debates, this is inappropriate, as there is still room for bias in ACs (Tullett, 2022).

(7) *Sociopolitical Asymmetry in Science*

Scientific collaborations are not equally accessible to all researchers (Marsh, 2022). There is a sociopolitical asymmetry in science, especially in the social sciences and humanities. This leads to an imbalance in research questions, which even ACs cannot counteract. For example, for the field of psychology in the USA, there are data suggesting that seventeen researchers are registered as Democrats for every Republican. In some fields, the imbalance of political orientation in research is even stronger (such as in communication sciences or anthropology). This asymmetry between liberals and conservatives cannot be balanced by ACs, and they cannot prevent certain questions from being investigated less frequently due to ideological preferences (Ceci & Williams, 2022). Scientists who identify as women participate less in international collaborations, whereas researchers who identify as men are more likely to work with other male scientists (Marsh, 2022). Even review committees, funding agencies, reviewers, and editors have ideological preferences that could negatively

affect ACs (Tullett, 2022). Established scientists may not want to work with researchers who have just earned their doctorate. They may consider it less beneficial to their careers to work with someone who has hardly published anything. This would mean that unknown researchers have fewer opportunities to participate in opposing collaborations than their already established colleagues (Marsh, 2022).

(8) *Societally Controversial Topics are Less Addressed*

Some topics could be so sensitive or controversial that only a few scientists would be willing to take them on. This would be regrettable, but nothing new, considering that scientists have avoided investigating highly controversial topics in the past and still do so today (Clark et al., 2022a). Table 2 provides an overview of the topics covered in the ACs that have been conducted so far.

(9) *Not All Questions Find an Answer*

Research questions without competing hypotheses are not suitable for ACs. These include exploratory work or cases where the phenomenon has been proven beyond doubt (Clark et al., 2022a). Other research questions cannot be resolved because of moral concerns, for example, manipulating the experience of microaggression in the form of a long-term study. In this case, an AC would be more suitable for finding out whether common definitions of microaggression are consistent (Tullett, 2022). Ultimately, a dispute can also be based on ideological issues, which limits the possibilities of systematic data collection (Latham et al., 1988).

(10) *One Experiment Alone Cannot Reject a Theory*

ACs cannot unambiguously falsify a view. In most cases, the different theoretical concepts are not like card towers that collapse when one hypothesis is refuted (Cowan et al., 2020; Ceci & Williams, 2022). In some ACs, the results partly agree with all theories (Cowan et al., 2020; Doherty et al., 2019). In others, clear discrepancies cannot be resolved (Stern & Crawford, 2021). Even if the results clearly contradict a standpoint, the advocates of this approach are not always convinced (Matzke et al., 2015; Tatlidil & Sloman, 2022).

Theories are often flexible enough to accommodate new insights. As a result, several theories can sometimes fit the new data. To recognize only one theory as valid, researchers could agree to accept the simplest explanation for the outcomes obtained (i.e., Ockham's Razor). However, the parties often have different views on which interpretation is the simplest (Cowan et al., 2020). A radical change of view only happens when considerable and diverse data have been collected that outright contradict a theory (Cowan et al., 2020; Gilovich et al., 1998). An AC cannot eliminate a scientific debate but rather provides a first step for systematic and large-scale research (Cowan et al., 2020; D. Martschenko et al., 2019).

(11) *Differentiation of Theories*

To test the validity of different theories, these must first be compared to serve as a basis for hypotheses. Comparing theories can be challenging for several reasons. The main difference may be solely in how effects are interpreted (Cowan et al., 2020; Doherty et al., 2019). Differences can also be attributed to different levels of explanation (e.g., theory of brain function vs. theory of mind) rather than to a genuine conflict of ideas (Clark et al., 2022a; Cowan et al., 2020; Doherty et al., 2019). For these reasons, it can be challenging to find hypotheses that clearly differentiate between rival theories and allow predictions that would only corroborate one of the distinct views (Doherty et al., 2019).

(12) *Dependency on Clear Contradictory Hypotheses*

ACs are heavily dependent on the opposing character of the theories at their core, as well as on precise hypotheses. To avoid real conflicts and preserve their professional

prestige, parties could exaggerate insignificant differences between their theories and label themselves as opponents, then agree on imprecise hypotheses and less strict tests of theories. This results in flexible results that are open to interpretation (Cowan, 2022; Berzon, 2012; Tatlidil & Sloman, 2022). Also, this shortens discussion time and allows for more general conclusions, and no side ends up as a loser. Over time, “more flexible” research groups could emerge that are preferred as “opponents” (Tatlidil & Sloman, 2022).

(13) *Disagreement Between Parties Regarding Research Design*

It is a challenge to find a way to transfer differences in basic convictions into a joint study (VanLange, 2013). Due to different views, expectations, design plans, and working methods, it will take time and willingness to compromise to find a consensus with regard to the study design (Kerr et al., 2018; Rakow et al., 2015).

(14) *Communication Problems*

Communication problems can arise from language barriers, different time zones, or scheduling conflicts (Tatlidil & Sloman, 2022).

(15) *Replication Problems*

Even an AC can have difficulties replicating a study precisely (i.e., a direct or close replication), since most of the time not all necessary information from previous studies has been recorded, and the authors of the original study are not always involved in the experiment (Kerr et al., 2018).

(16) *Depends on Results*

The success of an AC also depends on how often it is cited and whether it is included in a journal with high prestige and/or a high journal impact factor, which in turn depends on whether significant results were found in the respective AC. In particular, scientists in early career stages may perceive an AC as too expensive for these reasons (Tatlidil & Sloman, 2022).

(17) *Other Expectations of the Readers*

Most readers of scholarly journals prefer studies with clear theoretical statements, whereas an AC often does not offer a clear result; ideally, ACs are nuanced and point out uncertainties about the correct interpretation (Cowan, 2022).

(18) *Not Every Initiated AC is Conducted*

It should also be noted that not every AC will work. In some research areas, the debate can turn into complete bitterness and possibly even into legal proceedings (Lindner & Tribble, 2020).

4. Discussion

AC is a method that is increasingly recommended in contemporary scientific research to deal with contrasting theories, to reduce bias, and to advance scientific progress more quickly. Transparent and collaborative research is of particular importance in today's polarizing and rapidly changing world. The aim of this systematic mapping review was to provide a comprehensive overview of the formal aspects of AC and to identify important factors that must be considered when applying this approach. Our systematic mapping review of the landscape of the AC literature provides a guide for researchers who consider embarking on an AC, as well as an introduction to the topic for laypeople.

Ours is not the first account dealing with the AC method; rather, as is evident from the present review, the number of conceptual, methodology, or opinion papers devoted to ACs is ever-growing. However, many of these articles only touch on the topic. In contrast, one

particularly detailed work is Rakow (2022), who summarized several already completed ACs, named the key features of ACs and their advantages, briefly outlined the tasks of a mediator, and explained the term quasi AC (as adopted by us for the current review). This work provides a useful overview, although it is limited in its intended scope. The list of ACs mentioned in Rakow (2022) is not exhaustive with respect to the available corpus of AC studies, implementation steps of ACs are not systematically compared, their advantages are not exhaustively listed, and there is no clear indication of the disadvantages surrounding ACs. Further, there are no quantitative overviews of bibliometric data and indicators; also, comprehensive analyses of the role of the mediator in ACs and of quasi ACs are lacking. All these features are included and discussed in the present systematic mapping review. Our work not only introduces the topic for laypeople but also offers researchers who plan to conduct an AC a clear and unvarnished view of the current state of research on ACs, a detailed guide for their optimal implementation, and a detailed list of advantages and challenges (including proposed solutions) to expect.

Other review articles focus more strongly on the positive aspects of ACs whilst, in our view, neglecting some of their potential negative sides (Clark et al., 2022a, 2022b, 2022c). Accounts in which authors of published ACs describe their individual experiences and recommendations offer important and exciting insights, but self-evidently lack systematic comparisons with other ACs (Lindner & Tribble, 2020). However, all of these works demonstrate that AC is gaining increased interest. This is clearly reflected in the temporal distribution of publications on this topic from 1988 to 2022 (see Figure 2). At the same time, despite numerous publications recommending ACs, including reviews, comments, and book chapters, the number of actually conducted ACs—perhaps surprisingly so—still is small (see Table 4). Although the research community seems to be aware of the numerous benefits of ACs, the costs of implementing them seem to outweigh the benefits (Rakow, 2022). Hence, the question arises as to how ACs can be better promoted to encourage an increase, as it is evidently not enough to just highlight them in a variety of accounts. Specifically, one viable avenue of future research and evaluation would be to design and test various incentives (e.g., specific editorial policies of journals, dedicated funding, and academic recognition vis-à-vis ACs).

On average, the majority of the journals covered here so far have published merely one or two papers on ACs, with the *Journal of Applied Research in Memory and Cognition* currently being the journal with the most published contributions on this topic (see Table 1). Evidently, the research field so far mainly concerned with ACs remains psychology, and this is particularly true for its experimental, social, clinical, and transpersonal psychology subfields, as well as industrial and organizational psychology (Bateman et al., 2005; Matzke et al., 2015; Mellers et al., 2001). When other disciplines were involved in ACs, it was at the intersection with psychology (Bateman et al., 2005; Cadsby et al., 2008; Fischer et al., 2021; Munter et al., 2015). Only one paper has discussed potential applications of ACs in the justice system (Berzon, 2012). It still appears unclear which additional strengths and/or obstacles the application of ACs would bring to other research fields besides psychology.

Our author-based analysis revealed that the majority of researchers in the AC literature corpus have only published one AC article, whereas only a few have published repeatedly in this area. Even without knowing the exact connections between all authors involved in this field, it is obvious from mere bibliometric analysis that a small portion of researchers are, or have been, repeatedly engaged with ACs, such as Kahneman, Mellers, Tetlock, Clark, or Fiske (Clark et al., 2022a; Mellers et al., 2001; Koch et al., 2020). In this context, we add that the early AC study by Mellers et al. (2001) has meanwhile been replicated (Chandrashekar et al., 2021), albeit not again in the form of an AC.

One of the basic requirements of ACs is the involvement of several researchers with different views, which could create more opportunities to promote inclusion and diversity in studies than is possible with traditional study designs. If two or more research groups with different explanations for the same phenomenon are found, they work together to develop a study that tests their predictions for validity. An independent person can participate to mediate possible disputes (Cowan et al., 2020; Latham et al., 1988). In the end, the researchers publish their study together, regardless of the outcome (Mellers et al., 2001). This concept is similar to that of “strong inference” (Platt, 1964; Fudge, 2014), in which distinct (optimally, contradictory) hypotheses are tested in a crucial experiment to obtain clear results that clearly support or reject one hypothesis or the other. This is also the basic tenet of the AC; however, Table 2 shows that despite this rigorous approach, researchers do not always agree, as theories cannot be simply discarded by refuting a single hypothesis.

Nevertheless, ACs have numerous advantages, including maximizing expert knowledge, better understanding of other theories, faster research progress, and transparency of the research process (Bateman et al., 2005; Cowan et al., 2020; Makel et al., 2019; Matzke et al., 2015). It goes without saying that there are also difficulties encountered with ACs, including their time-consuming nature, professional risks (e.g., the effects of potential power dynamics, or asymmetries, within AC projects, should both junior and senior researchers partake in these), and aversion to interpersonal conflicts (Clark et al., 2022a; Haaga, 2009; Marsh, 2022). As already elaborated on above, to make the implementation of ACs more attractive to researchers, journals, universities, and research funding agencies should begin actively promoting them, for example, by providing training and funding opportunities (Vlasceanu et al., 2022).

The overall quality score of the fourteen ACs analyzed here in detail amounted to 61%. This is because many projects did not describe certain important aspects of an AC in the published paper (e.g., which researchers represented which views, whether there was an AC agreement, how communication between the parties proceeded, etc.; Dwyer et al., 2017). Whilst the pioneering AC by Mellers et al. (2001) already discussed the role and importance of the mediator, more recent accounts (e.g., Bateman et al., 2005) still do not automatically refer to a mediator. In the case of the Bateman et al. (2005) AC, this is somewhat surprising, because Kahneman was involved in both Mellers et al. (2001) and Bateman et al. (2005) (see also Kahneman & Anderson, 2003). It remains unclear why the latter AC apparently did not utilize the AC mediator role. To ensure greater transparency, future ACs are recommended to document such and any further relevant factors in their respective AC agreement forms and publications.

Also, we have emphasized the existence of so-called quasi ACs, which do not meet all the requirements of a full-fledged AC, although quasi ACs by now seem to outnumber genuine ACs. For reasons of transparency, it is recommended that one refer to such collaborations not as ACs, but rather, for example, as *collaborative analysis* (Haaga, 2009).

This first systematic mapping review of the landscape of ACs (including quasi ACs and conceptual, theory, and method papers related to these) has several limitations. First, it is likely that the literature-search strategies applied did not capture all ACs that have been conducted to date. To identify eligible articles on this topic, the key search term used was “adversarial collaboration”, which, however, was only introduced in 2001 by Kahneman (Mellers et al., 2001). It is possible that eligible studies prior to the 2000s were not found, due to a lack of having been labeled as ACs. To overcome this limitation, backward-citation searches, as well as cited-reference searches, were applied. Still, it cannot be ruled out that early ACs, not yet labeled as such (i.e., hidden or clandestine ACs), have been missed with these literature-search strategies.

Second, it is clear that it cannot be expected that all conceivable factors that facilitate conducting an AC are actually described in the sources reviewed here. This constitutes an inevitable blind spot and limit of knowledge of the current review. For example, authors of various ACs may be in close contact with each other and may motivate each other to conduct an AC and share their experiences informally, which could potentially provide important information; however, they do not find their way into the final published record of the respective AC. Relatedly, it is a conspicuous fact that mainly the field of psychology (and, to some extent, the broader domain of the behavioral and social sciences, including economics) deals with ACs, whereas the life and natural sciences, as well as the medical sciences and further science domains (e.g., physics), so far seem to remain untouched by this methodology. There may be deeper, or tacit, reasons for this phenomenon, not explicitly retrievable from the AC literature at large. This means that while the current review provides a comprehensive overview and synthesis of extractable data and knowledge contents from ACs, quasi ACs, and scholarly literature related to these, it is not all-encompassing.

And third, research syntheses, such as the one at hand, which have not been planned and implemented as living systematic reviews (Bucher et al., 2023) but rather had a deadline for literature searches, coding, synthesis, interpretation, and write-up and thus remain static documents, will be gradually outdated by the natural growth of evidence. In this specific case, we do not anticipate that core evidence from our review work (e.g., the organized lists of AC-related advantages and challenges, alongside solutions to the latter ones; or the best-practice recommendations for conducting ACs) will be overthrown in its entirety. However, owing to the natural growth of research accounts, the bibliometric data from the AC literature will eventually be superseded, as more ACs amass over time (but, considering that ACs still are rare, this will occur slowly).

In the context of the study limitations elaborated on above, here we take the opportunity to illustrate the last one of these with three specific updates that fell past the deadline and timeframe of the current review. First, Ceci et al. (2023) presented a comprehensive review of the large body of evidence informative to the research question of gender-based biases across key domains of academia and science. Following the terminology applied here, this account is a quasi AC (e.g., no new data were generated).

Second, Ceci et al. (2026) published a valuable resource on ACs in the form of a conceptual paper, in which they provide various valuable background of the AC idea, including survey data from psychologists with regard to the current popularity (or unpopularity of ACs), as well as an itemized compilation of no less than 60 scientific controversies which seem amenable, or ripe, for ACs. Further, they dwell upon the manifold and interesting epistemic synergies between ACs on the one hand and varieties of study preregistration (in particular, the Registered Report [RR] and Registered Replication Report [RRR] formats) on the other, and on the similarities vs. differences between ACs on the one hand and approaches like meta-analysis (i.e., quantitative evidence synthesis), metascience (i.e., science about science and scientists), and multiverse-style approaches to data analysis on the other hand. Inter alia, Ceci et al. (2026) also draw attention to the seemingly neglected fact that ACs inherently, by the very nature of this research approach, will improve research validity and the credibility of science, and further shared the (surely not yet widely known) fact that a recent call for ACs, issued at a leading journal in research methods (*Advances in Methods and Practices in Psychological Science*) was met with no responses at all. All in all, Ceci et al. (2026) can be perceived as a much-needed call for action with regard to the AC method, while, as an opinion piece and conceptual and methodological paper, it does not follow the systematic review approach, thus aptly complementing, rather than substituting, the current systematic mapping review of ACs.

And third, a natural extension of the scope of the present review work was contributed by Isch et al. (2025), in which these researchers reported the very first expert survey of scholars who had participated in an already completed AC. Key findings of this survey work on the experiences of AC project participants include that interpersonal conflicts in the course of partaking in ACs were (perhaps unexpectedly) rare. Further, about one-third of scholars participating in ACs actually reported belief updating on the controversial research question under scrutiny in the respective AC, whereas the remainder reported no belief updating or that it was “complicated”. One strength of this first survey study on AC participants was the high response rate achieved (namely, 24 from the total of 29 scholars who co-authored the 13 AC projects that were in the sampling frame of Isch et al., 2025, completed the survey on their AC experiences). Limitations include the sampling frame of AC projects itself (the survey dates back to 2022 and thus coverage was not entirely up-to-date, only completed ACs were considered, and quasi ACs were not considered), the qualitative nature and brevity of the 10-item survey, and the fact that no pre-post data on the expectations and experiences of the researchers participating in ACs could be implemented. Still, with these caveats in mind, the survey work of Isch et al. (2025) provides manifold starting points for further inquiries.

Additionally, what is noteworthy about the first two of the above accounts is their highly visible placement. Ceci et al. (2023) appeared on the pages of *Psychological Science in the Public Interest*, a flagship journal of the APS (Association for Psychological Science), which boasts an imposing journal impact factor. Ceci et al. (2026) was published in the *American Psychologist*, the official society journal of the APA (American Psychological Association). The appearance of these two pieces in such highly visible and accessible scholarly outlets attests to the timeliness and value of ACs, a constructive, promising endeavor of team science to elucidate and resolve debates in science, which time appears to have come now.

5. Conclusions

On balance, the AC approach is a valuable tool for advancing research and enabling the integration and assessment of competing theories and therefore may well be valuable up to the level of further integration and convergence of the sciences themselves, in the sense of consilience phenomena (i.e., unity and synthesis of knowledge; Wilson, 1998). However, ACs also require a lot of time, effort, and willingness to resolve interpersonal conflicts. The broad results presented here can help researchers assess whether an AC would be suitable for their own research and, if so, how to effectively design and implement it. Furthermore, the information assembled in this systematic mapping review can serve as a basis for developing guidelines to promote ACs in academic circles.

There are several unanswered questions about ACs that future inquiry along these lines could fruitfully address. It would be helpful to investigate how this type of research collaboration differs across science domains, whether (and, if so, why) it is essential to have a mediator, and how personal involvement and experiences affect the process of ACs. Additionally, long-term benefits and limitations of ACs should be assessed in more detail, e.g., how ACs can be adapted to the needs of different research contexts and cultures, and how the impact of ACs on both the scientific community and society can be measured.

Research is constantly evolving. It is therefore important to investigate concepts that meet the new and complex challenges of science. The AC approach holds promise to represent such an approach. Accordingly, researchers from all disciplines should consider the AC method as a potential way to improve transparency, replicability, validity, credibility, and the impact of their work, ultimately contributing to better science.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/publications14030047/s1>, Table S1: PRISMA 2020 checklist.

Author Contributions: Conceptualization, A.S. and M.V.; methodology, A.S., U.S.T. and M.V.; software, A.S.; validation, U.S.T. and M.V.; formal analysis, A.S.; investigation, A.S.; resources, M.V.; data curation, A.S.; writing—original draft preparation, A.S.; writing—review and editing, A.S., U.S.T. and M.V.; visualization, A.S.; supervision, M.V.; project administration, M.V.; funding acquisition, M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Vienna.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

References marked with an asterisk (*) indicate studies included in the systematic mapping review.

- *Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, 128(2), 290–314. [CrossRef] [PubMed]
- Advisory Board. (n.d.). *Penn arts & sciences adversarial collaboration project*. Available online: <https://web.sas.upenn.edu/adcollabproject/undergraduate/> (accessed on 12 July 2026).
- *Alempaki, D., Canic, E., Mullett, T. L., Skylark, W. J., Starmer, C., Stewart, N., & Tufano, F. (2019). Reexamining how utility and weighting functions get their shapes: A quasi-adversarial collaboration providing a new interpretation. *Management Science*, 65(10), 4841–4862. [CrossRef]
- *Bateman, I., Kahneman, D., Munro, A., Starmer, C., & Sugden, R. (2005). Testing competing models of loss aversion: An adversarial collaboration. *Journal of Public Economics*, 89(8), 1561–1580. [CrossRef]
- *Becker, M., Klauer, K. C., & Spruyt, A. (2016). Is attention enough? A re-examination of the impact of feature-specific attention allocation on semantic priming effects in the pronunciation task. *Attention, Perception & Psychophysics*, 78(2), 396–402. [CrossRef] [PubMed]
- *Beckers, T., De Houwer, J., & Dwyer, D. M. (2016). Reasoning versus association in animal cognition: Current controversies and possible ways forward. *Journal of Comparative Psychology*, 130(3), 187–191. [CrossRef] [PubMed]
- *Berman, J. S., & Reich, C. M. (2010). Investigator allegiance and the evaluation of psychotherapy outcome research. *European Journal of Psychotherapy & Counselling*, 12(1), 11–21. [CrossRef]
- *Berzon, J. C. (2012). Dissent, “dissentals”, and decision making. *California Law Review*, 100(6), 1479–1492.
- Brockman, J. (2022). *Adversarial collaboration: An EDGE lecture by Daniel Kahneman*. Edge. Available online: <https://www.edge.org/adversarial-collaboration-daniel-kahneman> (accessed on 12 July 2026).
- Bucher, L., Burgard, T., Tran, U. S., Prinz, G. M., Bosnjak, M., & Voracek, M. (2023). Keeping meta-analyses alive and well: A tutorial on implementing and using community-augmented meta-analyses in PsychOpen CAMA. *Advances in Methods and Practices in Psychological Science*, 6(4), 25152459231197611. [CrossRef]
- *Buranyi, S. (2017, June 27). *Is the staggeringly profitable business of scientific publishing bad for science?* The Guardian. Available online: https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science?CMP=Share_iOSApp_Other (accessed on 12 July 2026).
- *Cadsby, C. B., Croson, R., Marks, M., & Maynes, E. (2008). Step return versus net reward in the voluntary provision of a threshold public good: An adversarial collaboration. *Public Choice*, 135(4), 277–289. [CrossRef]
- Ceci, S. J., Clark, C. J., Jussim, L., & Williams, W. M. (2026). Adversarial collaboration: An undervalued approach in behavioral science. *American Psychologist*, 81(3), 374–390. [CrossRef] [PubMed]
- Ceci, S. J., Kahn, S., & Williams, W. M. (2023). Exploring gender bias in six key domains of academic science: An adversarial collaboration. *Psychological Science in the Public Interest*, 24(1), 15–73. [CrossRef] [PubMed]
- *Ceci, S. J., & Williams, W. M. (2022). The importance of viewpoint diversity among scientific team members: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 35–40. [CrossRef]
- *Chandrashekar, S. P., Cheng, B. L., Cheng, Y. H., Fong, C. L., Leung, C. L., Wong, Y. T., & Feldman, G. (2021). Frequency estimation and semantic ambiguity do not eliminate conjunction bias, when it occurs: Replication and extension of Mellers, Hertwig, and Kahneman (2001). *Meta-Psychology*, 5, MP.2020.2474. [CrossRef]
- *Clark, C. J., Costello, T., Mitchell, G., & Tetlock, P. E. (2022a). Keep your enemies close: Adversarial collaborations will improve behavioral science. *Journal of Applied Research in Memory and Cognition*, 11(1), 1–18. [CrossRef]

- *Clark, C. J., Costello, T., Mitchell, G., & Tetlock, P. E. (2022b). The road less traveled: Understanding adversaries is hard but smarter than ignoring them. *Journal of Applied Research in Memory and Cognition*, 11(1), 50–53. [CrossRef]
- *Clark, C. J., Honeycutt, N., & Jussim, L. (2022c). Replicability and the psychology of science. In W. T. O'Donohue, A. Masuda, & S. Lilienfeld (Eds.), *Avoiding questionable research practices in applied psychology* (pp. 45–71). Springer.
- *Clark, C. J., & Tetlock, P. E. (2022). Adversarial collaboration: The next science reform. In C. L. Frisby, R. E. Redding, W. T. O'Donohue, & S. O. Lilienfeld (Eds.), *Ideological and political bias in psychology: Nature, scope, and solutions* (pp. 905–927). Springer.
- *Cleeremans, A. (2022). Theory as adversarial collaboration. *Nature Human Behaviour*, 6(4), 485–486. [CrossRef] [PubMed]
- *Cohen, A., Cash, D., & Muller, M. (2000, December 2–6). *Designing to support adversarial collaboration* [Conference Writing]. 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA, USA. [CrossRef]
- *Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R. M. K., ... Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour*, 6(12), 1731–1742. [CrossRef] [PubMed]
- *Corrigan, J. R., Drichoutis, A. C., Lusk, J. L., Nayga, R. M. J., & Rousu, M. C. (2012). Repeated rounds with price feedback in experimental auction valuation: An adversarial collaboration. *American Journal of Agricultural Economics*, 94(1), 97–115. [CrossRef]
- *Costello, T. H., Clark, C. J., & Tetlock, P. E. (2022). Shoring up the shaky psychological foundations of a micro-economic model of ideology: Adversarial collaboration solutions. *Psychological Inquiry*, 33(2), 88–94. [CrossRef]
- *Cowan, N. (2022). The adversarial collaboration within each of us: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 19–22. [CrossRef] [PubMed]
- *Cowan, N., Belletier, C., Doherty, J. M., Jaroslawska, A. J., Rhodes, S., Forsberg, A., Naveh-Benjamin, M., Barrouillet, P., Camos, V., & Logie, R. H. (2020). How do scientific views change? Notes from an extended adversarial collaboration. *Perspectives on Psychological Science*, 15(4), 1011–1025. [CrossRef] [PubMed]
- *Crusius, J., Gonzalez, M. F., Lange, J., & Cohen-Charash, Y. (2020). Envy: An adversarial review and comparison of two competing views. *Emotion Review*, 12(1), 3–21. [CrossRef]
- *DeMatteo, D., & Olver, M. E. (2022). Use of the psychopathy checklist revised in legal contexts: Validity, reliability, admissibility, and evidentiary issues. *Journal of Personality Assessment*, 104(2), 234–251. [CrossRef] [PubMed]
- *Doherty, J. M., Belletier, C., Rhodes, S., Jaroslawska, A., Barrouillet, P., Camos, V., Cowan, N., Naveh-Benjamin, M., & Logie, R. H. (2019). Dual-task costs in working memory: An adversarial collaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(9), 1529–1551. [CrossRef] [PubMed]
- *Duarte, J. L., Crawford, J. T., Stern, C., Haidt, J., Lee, J., & Tetlock, P. E. (2015). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*, 38, e130. [CrossRef] [PubMed]
- *Duygu, U. T. (2022). We should redefine scientific expertise: An extended virtue account. *European Journal for Philosophy of Science*, 12, 71. [CrossRef]
- *Dwyer, R. J., Dunn, E. W., & Hershfield, H. E. (2017). Cousins or conjoined twins: How different are meaning and happiness in everyday life? *Comprehensive Results in Social Psychology*, 2(2–3), 199–215. [CrossRef]
- *Edlund, J. E., Cuccolo, K., Irgens, M. S., Wagge, J. R., & Zlokovich, M. S. (2022). Saving science through replication studies. *Perspectives on Psychological Science*, 17(1), 216–225. [CrossRef] [PubMed]
- *Ellemers, N., Fiske, S. T., Abele, A. E., Koch, A., & Yzerbyt, V. (2020). Adversarial alignment enables competing models to engage in cooperative theory building toward cumulative science. *Proceedings of the National Academy of Sciences of the USA*, 117(14), 7561–7567. [CrossRef] [PubMed]
- *Engelhardt, P. E. (2020). Developmental dyslexia: Where do we go from here? *Brain Sciences*, 10(3), 151. [CrossRef] [PubMed]
- *Finkel, E. J., Norton, M. I., Reis, H. T., Ariely, D., Caprariello, P. A., Eastwick, P. W., Frost, J. H., & Maniaci, M. R. (2015). When does familiarity promote versus undermine interpersonal attraction? A proposed integrative model from erstwhile adversaries. *Perspectives on Psychological Science*, 10(1), 3–19. [CrossRef] [PubMed]
- *Fischer, E., Engelhardt, P., & Sytsma, J. (2021). Inappropriate stereotypical inferences? An adversarial collaboration in experimental ordinary language philosophy. *Synthese*, 198(11), 10127–10168. [CrossRef]
- *Fiske, S. T. (2017). Going in many right directions, all at once. *Perspectives on Psychological Science*, 12(4), 652–655. [CrossRef]
- Fudge, D. S. (2014). Fifty years of J. R. Platt's strong inference. *Journal of Experimental Biology*, 217(8), 1202–1204. [CrossRef] [PubMed]
- *Gandalf, N. F. (2020). *Spontaneous stereotype content: Measurement aiming toward theoretical integration and discovery* [Unpublished doctoral dissertation, Princeton University]. Available online: <https://dataspace.princeton.edu/handle/88435/dsp01zp38wg55d> (accessed on 12 July 2026).
- *Gilovich, T., Medvec, V. H., & Kahneman, D. (1998). Varieties of regret. *Psychological Review*, 105(3), 602–605. [CrossRef]
- *Gordon, R. A., Rowe, H. L., & Garcia, K. (2015). Promoting family resilience through evidence-based policy making: Reconsidering the link between adult-infant bedsharing and infant mortality. *Family Relations*, 64(1), 134–152. [CrossRef]

- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108. [CrossRef] [PubMed]
- *Greeno, D., Morey, C., & Popov, V. (2022). *Preregistration report: An adversarial collaboration contrasting two explanations for the effect of working memory load on distractor processing during encoding and retention* [preregistration]. Available online: <https://osf.io/4ds3a/files/osfstorage> (accessed on 12 July 2026).
- *Gruber, D. R. (2022). On integrated information theory (IIT) and adversarial collaboration: A conversation with Christof Koch, PhD. *Journal of Consciousness Studies*, 29(11), 174–185. [CrossRef]
- *Haaga, D. A. (2009). A timely reminder on research design and interpretation. *Clinical Psychology: Science and Practice*, 16(1), 66–68. [CrossRef]
- Hanfstingl, B., Oberleiter, S., Pietschnig, J., Tran, U. S., & Voracek, M. (2024). Detecting jingle and jangle fallacies by identifying consistencies and variabilities in study specifications: A call for research. *Frontiers in Psychology*, 15, 1404060. [CrossRef] [PubMed]
- Heyman, T., Moors, P., & Rabagliati, H. (2020). The benefits of adversarial collaboration for commentaries. *Nature Human Behaviour*, 4, 1217. [CrossRef] [PubMed]
- *Hollon, S. D. (1999). Allegiance effects in treatment research: A commentary. *Clinical Psychology: Science and Practice*, 6(1), 107–112. [CrossRef]
- Isch, C., Tetlock, P. E., & Clark, C. J. (2025). Reflections on adversarial collaboration from the adversaries: Was it worth it? *Theory and Society*, 54, 929–946. [CrossRef] [PubMed]
- *Joel, S., Eastwick, P. W., & Finkel, E. J. (2018). Open sharing of data on close relationships and other sensitive social psychological topics: Challenges, tools, and future directions. *Advances in Methods and Practices in Psychological Science*, 1(1), 86–94. [CrossRef]
- *Johnson, M., Johnson, E., & Nettle, D. (2023). Are ‘red wall’ constituencies really opposed to progressive policy? Examining the impact of materialist narratives for universal basic income. *British Politics*, 18(1), 104–127. [CrossRef] [PubMed]
- *Kahneman, D., & Anderson, N. B. (2003). Experiences of collaborative research. *American Psychologist*, 58(9), 723–730. [CrossRef] [PubMed]
- *Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6), 515–526. [CrossRef] [PubMed]
- *Kerr, N., Ao, X., Hogg, M. A., & Zhang, J. (2018). Addressing replicability concerns via adversarial collaboration: Discovering hidden moderators of the minimal intergroup discrimination effect. *Journal of Experimental Social Psychology*, 78, 66–76. [CrossRef]
- *Koch, A., Imhoff, R., Unkelbach, C., Nicolas, G., Fiske, S., Terache, J., Carrier, A., & Yzerbyt, V. (2020). Groups’ warmth is a personal matter: Understanding consensus on stereotype dimensions reconciles adversarial models of social evaluation. *Journal of Experimental Social Psychology*, 89, 103995. [CrossRef]
- *Koole, S., & Lakens, D. (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science*, 7(6), 608–614. [CrossRef] [PubMed]
- *Lancaster, B., & Friedman, H. L. (2017). Introduction to the special topic section: The Taylor-Hartelius debate on psychology and spirituality. *International Journal of Transpersonal Studies*, 36(2), 72–74. [CrossRef]
- *Latham, G. P., Erez, M., & Locke, E. A. (1988). Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez-Latham dispute regarding participation in goal setting. *Journal of Applied Psychology*, 73(4), 753–772. [CrossRef]
- *Laythe, B., & Houran, J. (2022). Adversarial collaboration on a Drake-S Equation for the survival question. *Journal of Scientific Exploration*, 36(1), 130–160. [CrossRef]
- *Leal, F., & Marraud, H. (2022). *How philosophers argue: An adversarial collaboration on the Russell-Copleston debate*. Springer. [CrossRef]
- *Leichsenring, F., Abbass, A., Hilsenroth, M. J., Leweke, F., Luyten, P., Keefe, J. R., Midgley, N., Rabung, S., Salzer, S., & Steinert, C. (2017). Biases in research: Risk factors for non-replicability in psychotherapy and pharmacotherapy research. *Psychological Medicine*, 47(6), 1000–1011. [CrossRef] [PubMed]
- *Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology: Science and Practice*, 16(1), 54–65. [CrossRef]
- *Lindner, D., & Tribble, M. (2020). Keep your friends close, your adversaries closer. *The Psychologist*, 33, 34–35.
- *Lindner, D., Tribble, M., Pilato, I., & Ferguson, C. J. (2020). Examining the effects of exposure to a sexualized female video game protagonist on women’s body image. *Psychology of Popular Media*, 9(4), 553–560. [CrossRef]
- *Makel, M. C., Smith, K. N., McBee, M. T., Peters, S. J., & Miller, E. M. (2019). A path to greater credibility: Large-scale collaborative education research. *AERA Open*, 5(4), 2332858419891963. [CrossRef]
- *Makel, M. C., Smith, K. N., Miller, E. M., Peters, S. J., & McBee, M. T. (2020). Collaboration in giftedness and talent development research. *Journal for the Education of the Gifted*, 43(2), 91–107. [CrossRef]
- *Marewski, J. N., Bröder, A., & Glöckner, A. (2018). Some metatheoretical reflections on adaptive decision making and the strategy selection problem. *Journal of Behavioral Decision Making*, 31(2), 181–198. [CrossRef]

- *Marewski, J. N., Pohl, R. F., & Vitouch, O. (2011). Recognition-based judgments and decisions: What we have learned (so far). *Judgment and Decision Making*, 6(5), 359–380. [CrossRef]
- Marks-Anglin, A., & Chen, Y. (2020). A historical review of publication bias. *Research Synthesis Methods*, 11(6), 725–742. [CrossRef] [PubMed]
- *Marsh, J. K. (2022). Clearing the obstacles to adversarial collaborations for early career researchers: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 31–34. [CrossRef]
- *Martschenko, D., Trejo, S., & Domingue, B. W. (2019). Genetics and education: Recent developments in the context of an ugly history and an uncertain future. *AERA Open*, 5(1), 233285841881051. [CrossRef]
- *Martschenko, D. O., & Trejo, S. (2022). Ethical, anticipatory genomics research on human behavior means celebrating disagreement. *Human Genetics and Genomics Advances*, 3(1), 100080. [CrossRef] [PubMed]
- *Matzke, D., Nieuwenhuis, S., van Rijn, H., Slagter, H. A., van der Molen, M. W., & Wagenmakers, E. J. (2015). The effect of horizontal eye movements on free recall: A preregistered adversarial collaboration. *Journal of Experimental Psychology: General*, 144(1), e1–e15. [CrossRef] [PubMed]
- *Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, 12(4), 269–275. [CrossRef] [PubMed]
- *Melloni, L. (2022). On keeping our adversaries close, preventing collateral damage, and changing our minds: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 45–49. [CrossRef]
- *Munter, C., Stein, M. K., & Smith, M. S. (2015). Dialogic and direct instruction: Two distinct models of mathematics instruction and the debate(s) surrounding them. *Teachers College Record*, 117(11), 1–32. [CrossRef]
- *Neal, T. M. S., Sellbom, M., & de Ruiter, C. (2022). Personality assessment in legal contexts: Introduction to the special issue. *Journal of Personality Assessment*, 104(2), 127–136. [CrossRef] [PubMed]
- *Nier, J., & Campbell, S. D. (2013). Two outsiders' view on feminism and evolutionary psychology: An opportune time for adversarial collaboration. *Sex Roles*, 69(9–10), 503–506. [CrossRef]
- Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, 7(6), 657–660. [CrossRef] [PubMed]
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. [CrossRef] [PubMed]
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, 372, n71. [CrossRef] [PubMed]
- Platt, J. R. (1964). Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, 146(3642), 347–353. [CrossRef] [PubMed]
- Popper, K. (2012). *The open society and its enemies*. Routledge. (Original work published 1945).
- *Rakow, T. (2022). Adversarial collaboration. In W. T. O'Donohue, A. Masuda, & S. Lilienfeld (Eds.), *Avoiding questionable research practices in applied psychology* (pp. 359–377). Springer.
- *Rakow, T., Thompson, V., Ball, L., & Markovits, H. (2015). Rationale and guidelines for empirical adversarial collaboration: A *Thinking & Reasoning* initiative. *Thinking & Reasoning*, 21(2), 167–175. [CrossRef]
- *Reinero, D. A., Wills, J. A., Brady, W. J., Mende-Siedlecki, P., Crawford, J. T., & Van Bavel, J. J. (2020). Is the political slant of psychology research related to scientific replicability? *Perspectives on Psychological Science*, 15(6), 1310–1328. [CrossRef] [PubMed]
- Research. (n.d.). *Penn arts & sciences adversarial collaboration project*. Available online: <https://web.sas.upenn.edu/adcollabproject/research/> (accessed on 12 July 2026).
- Research Team. (n.d.). *Penn arts & sciences adversarial collaboration project*. Available online: <https://web.sas.upenn.edu/adcollabproject/research-team/> (accessed on 12 July 2026).
- Reuters, T. (2022). *Journal citation reports*. Available online: <https://clarivate.com/products/journal-citation-reports/> (accessed on 12 July 2026).
- *Robbins, B. D., Friedman, H. L., Johnson, C. V., & Franco, Z. (2019). Subjectivity is no object: Can subject-object dualism be reconciled through phenomenology? *International Journal of Transpersonal Studies*, 37(2), 144–167. [CrossRef]
- *Roberts, J. M., Arbisi, P. A., John, M. A., & Seamans, N. R. (2022). The utility of the Trauma symptom Inventory as a primary and secondary assessment instrument for forensic practice in legal settings. *Journal of Personality Assessment*, 104(2), 221–233. [CrossRef] [PubMed]
- *Sellbom, M., Flens, J., Gould, J., Ramnath, R., Tringone, R., & Grossman, S. (2022). The Millon Clinical Multiaxial Inventory-IV (MCMI-IV) and Millon Adolescent Clinical Inventory-II (MACI-II) in legal settings. *Journal of Personality Assessment*, 104(2), 203–220. [CrossRef] [PubMed]

- *Steinert, C., Munder, T., Rabung, S., Hoyer, J., & Leichsenring, F. (2017). Psychodynamic therapy: As efficacious as other empirically supported treatments? A meta-analysis testing equivalence of outcomes. *American Journal of Psychiatry*, 174(10), 943–953. [CrossRef] [PubMed]
- *Stern, C., & Crawford, J. T. (2021). Ideological conflict and prejudice: An adversarial collaboration examining correlates and ideological (a)symmetries. *Social Psychological and Personality Science*, 12(1), 42–53. [CrossRef]
- *Tatlidil, S., & Sloman, S. (2022). Some collaborations just are not worth it: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 41–44. [CrossRef]
- *Taylor, A. E., & Munafò, M. R. (2016). Triangulating meta-analyses: The example of the serotonin transporter gene, stressful life events and major depression. *BMC Psychology*, 4(1), 23. [CrossRef] [PubMed]
- *Tetlock, P. E., & Mellers, B. A. (2011). Intelligent management of intelligence agencies: Beyond accountability ping-pong. *American Psychologist*, 66(6), 542–554. [CrossRef] [PubMed]
- *Tetlock, P. E., & Mitchell, G. (2009a). Adversarial collaboration aborted but our offer still stands. *Research in Organizational Behavior*, 29, 77–79. [CrossRef]
- *Tetlock, P. E., & Mitchell, G. (2009b). A renewed appeal for adversarial collaboration. *Research in Organizational Behavior*, 29, 71–72. [CrossRef]
- *Tetlock, P. E., & Mitchell, G. (2009c). Implicit bias and accountability systems: What must organizations do to prevent discrimination? *Research in Organizational Behavior*, 29, 3–38. [CrossRef]
- *Thompson, C. S. (2021). *Veterans treatment court stakeholder's experience of community collaboration serving gulf coast veterans* [Unpublished doctoral dissertation, Capella University]. Available online: <https://www.proquest.com/openview/bc54d4666146383776f5e74378b096db/1?pq-origsite=gscholar&cbl=18750&diss=y> (accessed on 12 July 2026).
- *Thouin-Savard, M. I. (2017). Adversarial collaboration: How free and open debate leads to better transpersonal ideas (Editor's introduction). *International Journal of Transpersonal Studies*, 36(2), iii–iv. [CrossRef]
- Trafimow, D. (2025). Theory-refuting findings in psychology: How much should they matter? *Journal for the Theory of Social Behaviour*, 55(1), e70000. [CrossRef]
- *Tullett, A. M. (2022). Adversarial collaborations will not solve society's moral debates: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 27–30. [CrossRef]
- *Van Dessel, P., Gawronski, B., Smith, C. T., & De Houwer, J. (2017). Mechanisms underlying approach-avoidance instruction effects on implicit evaluation: Results of a preregistered adversarial collaboration. *Journal of Experimental Social Psychology*, 69, 23–32. [CrossRef]
- *VanLange, P. A. M. (2013). What we should expect from theories in social psychology: Truth, abstraction, progress, and applicability as standards (TAPAS). *Personality and Social Psychology Review*, 17(1), 40–55. [CrossRef] [PubMed]
- *Viglione, D. J., de Ruiter, C., King, C. M., Meyer, G. J., Kivisto, A. J., Rubin, B. A., & Hunsley, J. (2022). Legal admissibility of the Rorschach and R-PAS: A review of research, practice, and case law. *Journal of Personality Assessment*, 104(2), 137–161. [CrossRef] [PubMed]
- *Vlasceanu, M., Reinero, D. A., & Van Bavel, J. J. (2022). Adversarial collaborations in behavioral science—Benefits and boundary conditions: Comment on Clark et al. *Journal of Applied Research in Memory and Cognition*, 11(1), 23–26. [CrossRef]
- Watson, G. (2004). Uncertainty and contractual hazard in the film industry: Managing adversarial collaboration with dominant suppliers. *Supply Chain Management*, 9(5), 402–409. [CrossRef]
- Wilson, E. O. (1998). *Consilience: The unity of knowledge*. Knopf.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.