



Cochrane
Library

Cochrane Database of Systematic Reviews

Interventions to prevent misconduct and promote integrity in research and publication (Review)

Marusic A, Wager E, Utrobicic A, Rothstein HR, Sambunjak D

Marusic A, Wager E, Utrobicic A, Rothstein HR, Sambunjak D.
Interventions to prevent misconduct and promote integrity in research and publication.
Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: MR000038.
DOI: [10.1002/14651858.MR000038.pub2](https://doi.org/10.1002/14651858.MR000038.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	3
METHODS	4
RESULTS	7
Figure 1.	8
Figure 2.	10
Figure 3.	11
Figure 4.	12
Figure 5.	13
DISCUSSION	18
AUTHORS' CONCLUSIONS	20
ACKNOWLEDGEMENTS	20
REFERENCES	21
CHARACTERISTICS OF STUDIES	24
ADDITIONAL TABLES	55
APPENDICES	69
CONTRIBUTIONS OF AUTHORS	80
DECLARATIONS OF INTEREST	80
SOURCES OF SUPPORT	80
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	80
INDEX TERMS	80

[Methodology Review]

Interventions to prevent misconduct and promote integrity in research and publication

Ana Marusic¹, Elizabeth Wager², Ana Utrobicic³, Hannah R Rothstein⁴, Dario Sambunjak⁵

¹Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia. ²Sideview, Princes Risborough, UK. ³Central Medical Library, University of Split, School of Medicine, Split, Croatia. ⁴Zicklin School of Business, Baruch College, New York, NY, USA. ⁵Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia

Contact address: Elizabeth Wager, Sideview, 19 Station Road, Princes Risborough, Buckinghamshire, HP27 9DE, UK.
liz@sideview.demon.co.uk.

Editorial group: Cochrane Methodology Review Group.

Publication status and date: New, published in Issue 4, 2016.

Citation: Marusic A, Wager E, Utrobicic A, Rothstein HR, Sambunjak D. Interventions to prevent misconduct and promote integrity in research and publication. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: MR000038. DOI: [10.1002/14651858.MR000038.pub2](https://doi.org/10.1002/14651858.MR000038.pub2).

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Improper practices and unprofessional conduct in clinical research have been shown to waste a significant portion of healthcare funds and harm public health.

Objectives

Our objective was to evaluate the effectiveness of educational or policy interventions in research integrity or responsible conduct of research on the behaviour and attitudes of researchers in health and other research areas.

Search methods

We searched the CENTRAL, MEDLINE, LILACS and CINAHL health research bibliographical databases, as well as the Academic Search Complete, AGRICOLA, GeoRef, PsycINFO, ERIC, SCOPUS and Web of Science databases. We performed the last search on 15 April 2015 and the search was limited to articles published between 1990 and 2014, inclusive. We also searched conference proceedings and abstracts from research integrity conferences and specialized websites. We handsearched 14 journals that regularly publish research integrity research.

Selection criteria

We included studies that measured the effects of one or more interventions, i.e. any direct or indirect procedure that may have an impact on research integrity and responsible conduct of research in its broadest sense, where participants were any stakeholders in research and publication processes, from students to policy makers. We included randomized and non-randomized controlled trials, such as controlled before-and-after studies, with comparisons of outcomes in the intervention versus non-intervention group or before versus after the intervention. Studies without a control group were not included in the review.

Data collection and analysis

We used the standard methodological procedures expected by Cochrane. To assess the risk of bias in non-randomized studies, we used a modified Cochrane tool, in which we used four out of six original domains (blinding, incomplete outcome data, selective outcome reporting, other sources of bias) and two additional domains (comparability of groups and confounding factors). We categorized our primary outcome into the following levels: 1) organizational change attributable to intervention, 2) behavioural change, 3) acquisition of knowledge/skills and 4) modification of attitudes/perceptions. The secondary outcome was participants' reaction to the intervention.

Main results

Thirty-one studies involving 9571 participants, described in 33 articles, met the inclusion criteria. All were published in English. Fifteen studies were randomized controlled trials, nine were controlled before-and-after studies, four were non-equivalent controlled studies with a historical control, one was a non-equivalent controlled study with a post-test only and two were non-equivalent controlled studies with pre- and post-test findings for the intervention group and post-test for the control group. Twenty-one studies assessed the effects of interventions related to plagiarism and 10 studies assessed interventions in research integrity/ethics. Participants included undergraduates, postgraduates and academics from a range of research disciplines and countries, and the studies assessed different types of outcomes.

We judged most of the included randomized controlled trials to have a high risk of bias in at least one of the assessed domains, and in the case of non-randomized trials there were no attempts to alleviate the potential biases inherent in the non-randomized designs.

We identified a range of interventions aimed at reducing research misconduct. Most interventions involved some kind of training, but methods and content varied greatly and included face-to-face and online lectures, interactive online modules, discussion groups, homework and practical exercises. Most studies did not use standardized or validated outcome measures and it was impossible to synthesize findings from studies with such diverse interventions, outcomes and participants. Overall, there is very low quality evidence that various methods of training in research integrity had some effects on participants' attitudes to ethical issues but minimal (or short-lived) effects on their knowledge. Training about plagiarism and paraphrasing had varying effects on participants' attitudes towards plagiarism and their confidence in avoiding it, but training that included practical exercises appeared to be more effective. Training on plagiarism had inconsistent effects on participants' knowledge about and ability to recognize plagiarism. Active training, particularly if it involved practical exercises or use of text-matching software, generally decreased the occurrence of plagiarism although results were not consistent. The design of a journal's author contribution form affected the truthfulness of information supplied about individuals' contributions and the proportion of listed contributors who met authorship criteria. We identified no studies testing interventions for outcomes at the organizational level. The numbers of events and the magnitude of intervention effects were generally small, so the evidence is likely to be imprecise. No adverse effects were reported.

Authors' conclusions

The evidence base relating to interventions to improve research integrity is incomplete and the studies that have been done are heterogeneous, inappropriate for meta-analyses and their applicability to other settings and population is uncertain. Many studies had a high risk of bias because of the choice of study design and interventions were often inadequately reported. Even when randomized designs were used, findings were difficult to generalize. Due to the very low quality of evidence, the effects of training in responsible conduct of research on reducing research misconduct are uncertain. Low quality evidence indicates that training about plagiarism, especially if it involves practical exercises and use of text-matching software, may reduce the occurrence of plagiarism.

PLAIN LANGUAGE SUMMARY

Preventing misconduct and promoting integrity in research and publication

Doctors and patients need to be able to trust reports of medical research because these are used to help them make decisions about treatments. It is therefore important to prevent false or misleading research. Problems with research include various types of misconduct such as altering results (falsification), making up results (fabrication) or copying other people's work (plagiarism). Good systems that produce reliable research are said to show 'research integrity'. We studied activities, such as training, designed to reduce research misconduct and encourage integrity. The effects of some of these activities on researchers' attitudes, knowledge and behaviour have been studied and we brought together the evidence from these studies.

Some studies showed positive effects on researchers' attitudes to plagiarism. Practical training, such as using computer programs that can detect plagiarism, or writing exercises, sometimes decreased plagiarism by students but not all studies showed positive effects. We did not find any studies on fabrication or falsification. Two studies showed that the way in which journals ask authors for details about who did each part of a study can affect their responses.

Many of the studies included in this review had problems such as small sample sizes or had used methods that might produce biased results. The training methods tested in the studies (which included online courses, lectures and discussion groups) were often not clearly described. Most studies tested effects over short time periods. Many studies involved university students rather than active researchers.

In summary, the available evidence is of very low quality, so the effect of any intervention for preventing misconduct and promoting integrity in research and publication is uncertain. However, practical training about how to avoid plagiarism may be effective in reducing plagiarism by students, although we do not know whether it has long-term effects.

BACKGROUND

Description of the problem or issue

The two World Conferences on Research Integrity and the Singapore Statement on Research Integrity, [Resnik 2011](#), called for the development of more comprehensive standards, codes and policies to promote research integrity both locally and on a global basis. However, there is little systematic evidence to guide the development and implementation of these standards. This systematic review assesses the existing evidence and identifies research questions that need to be addressed.

Description of the methods being investigated

Research integrity and responsible conduct of research emerged as a research topic in the 1990s after public reports of scientific fraud and the response by national policy makers to deal with the problem ([Steneck 2006](#)). Together with the growing understanding of research behaviour, there is also a growth of empirical information about interventions by institutions, funders, regulators, journal editors and other stakeholders to improve responsible conduct of research and foster research integrity. A systematic review and meta-analysis of the prevalence of fabrication, falsification or manipulation of research findings showed that an average of 1.97% (95% confidence interval (CI) 0.86 to 4.45) of scientists admitted such practices and that 9.54% (95% CI 5.15 to 13.94) admitted other questionable research practices ([Fanelli 2009](#)). Another meta-analysis on authorship in research publications also demonstrated a high prevalence of reported problems with authorship (pooled weighted average of 29% (95% CI 24 to 35) of authors reporting their own or others' experience of abuse of authorship) ([Marusic 2011](#)).

There is a growing body of evidence on attitudes and practices of researchers relating to the responsible conduct of research, serious forms of misconduct (i.e. fabrication, falsification and plagiarism or 'FFP') and so-called 'questionable research practices' ([Steneck 2006](#)). It is generally recognised that these behaviours form a continuum from outright misconduct at one end to ideal practices at the other, but that the boundaries between acceptable, careless, questionable and fraudulent behaviour are not universally defined.

How these methods might work

Interventions to foster responsible conduct of research or deter misconduct have been tested in a number of research settings and across research disciplines, although predominantly in medicine. Since some forms of misconduct and questionable research practices are assumed to result from ignorance rather than a deliberate intention to deceive, it is widely assumed that training will reduce such behaviours ([Anderson 2007](#); [Funk 2007](#); [Hren 2007](#); [Plemmons 2006](#)). Other interventions are based on the concept of screening acting as a deterrent, for example text-matching software may be used to screen material for plagiarism and it is believed that this will deter authors from these practices ([Bilić-Zulle 2008](#)). Since misconduct involves breaches in ethics, other interventions focus on establishing or reinforcing an ethical culture, for example by the use of honour codes ([Boyd 2004](#)).

Why it is important to do this review

Improper practices and unprofessional conduct in clinical research have been shown to waste a significant portion of healthcare

funds and adversely impact public health ([Angel 2004](#)). A synthesis of evidence for effective ways to promote responsible conduct of research and deter poor research practices should therefore have an important impact on the quality of research output. Understanding which techniques are effective will help institutions focus on these activities and also avoid interventions that may be counterproductive.

If published, fraudulent medical research can harm patients directly or indirectly. For example, a clinical trial that used a new method of predicting response to chemotherapy had to be stopped when the research on which the method was based was shown to be fraudulent. The paper was subsequently retracted but probably led to some cancer patients receiving suboptimal therapies ([Reich 2011](#)).

Although it is difficult to estimate the impact of misconduct outside of health research ([Steneck 2006](#)), a recent case study of the cost of research misconduct estimated the direct cost of an actual investigation to be USD 525,000 and indirect costs to be USD 1.3 million ([Michalek 2010](#)).

OBJECTIVES

Our objective was to evaluate the effectiveness of educational or policy interventions in research integrity or responsible conduct of research on the behaviour and attitudes of researchers in health and other research areas.

Since serious research and publication misconduct are relatively rare, and interventions may be designed to have long-term effects, the effects of interventions on the frequency of misconduct are hard to measure. Therefore we also considered surrogate endpoints such as researcher attitudes.

In this review, we particularly focused on interventions aimed at fostering research integrity, which views research behaviour from the perspective of professional standards ([Steneck 2006](#)). We also explored interventions to prevent unacceptable or fraudulent research and publication practices.

As there is significant variability in defining how researchers should behave when performing and communicating research ([Steneck 2006](#)), for the purpose of this review we used the following definitions:

- Responsible conduct of research was defined as "conducting research in ways that fulfil the professional responsibilities of researchers, as defined by their professional organizations, the institutions for which they work and, when relevant, the government and public" ([Steneck 2006](#)).
- Research integrity was defined as "the quality of possessing and steadfastly adhering to high moral principles and professional standards, as outlined by professional organizations, research institutions and, when relevant, the government and public" ([Steneck 2006](#)). Research integrity was differentiated from academic integrity, which is broader than research integrity and is defined as commitment to "six fundamental values: honesty, trust, fairness, respect, responsibility, and courage" ([ICAI 2013](#)); lack of academic integrity also includes engaging in plagiarism, unauthorized collaboration, cheating or facilitating academic dishonesty ([ICAI 2013](#)). We included a study dealing with academic integrity in this review when

it addressed plagiarism, because writing is an important component of authorship in research (ICMJJE 2014).

- Research ethics was defined as "moral problems associated with or that arise in the course of pursuing research" (Steneck 2006). Studies dealing with research ethics, regardless of how they were defined in the study, were included in the systematic review. We did not include studies dealing with professional ethics, which was defined as ethics related to professional work. While professional work generally includes research activities (Cogan 1953), we included studies which dealt with professional ethics only if research was explicitly specified as a study topic.
- Research misconduct included all misbehaviours in research, from fabrication, falsification and plagiarism (FFP) as defined by the US Office of Research integrity to a more inclusive list of misbehaviours as defined by research integrity bodies in Scandinavian countries (Steneck 2006).

METHODS

Criteria for considering studies for this review

Types of studies

Studies of educational or policy interventions are usually conducted in natural settings where a true randomized controlled design may not always be feasible. This is probably even truer for the topic of research integrity, which has only relatively recently become the focus of scientific investigations. For these reasons, we planned to include in our review not only randomized controlled trials, but also non-randomized controlled trials, such as controlled before-and-after studies, interrupted time series and regression discontinuity designs. We excluded observational (survey) data when there was no clear intervention or manipulation. We also excluded study designs without a comparison group. For controlled before-and-after studies we included studies that had before-and-after measurements for the intervention group and only one measurement for the control group. We examined quasi-experimental designs closely for threats to validity.

We included studies irrespective of publication status and language.

Types of data

We included studies that measured the effects of one or more interventions, such as teaching, training, mentoring, use of checklists, screening and policy, on research integrity or responsible conduct of research in its broadest sense, including 'questionable research practices' and publication misconduct. We considered interventions in any type of researcher or student, at any period of their research career, in all fields of research, including sciences, social sciences and humanities. We did not evaluate the effectiveness of reporting guidelines for improving the presentation of research data, as this is not directly related to research integrity, but rather to the quality of reporting, and has been covered by other systematic reviews (e.g. Plint 2006; Turner 2012).

Types of participants

The participants included any stakeholders in the research and publication process, such as: 1) students, who may or may not have an interest in becoming researchers, if they received an intervention related to research integrity; 2) health workers

involved in research; 3) researchers working at institutions or commercial research establishments; 4) authors, peer reviewers and/or editors of scholarly journals; 5) professional and/or research organizations; and 6) policy makers.

Types of interventions

Eligible interventions included any direct or indirect procedure that may have an impact on research integrity, from direct educational interventions, such as a formal course or training required by institutions or authorities (such as training required by Institutional Review Boards/Ethics Committees), to indirect interventions, such as policy change (e.g. introduction of statements on conflict of interest or authorship contribution declarations in journals).

Types of methods

We included the studies with comparisons of outcomes in intervention versus non-intervention groups or before versus after the intervention. We assessed the groups for baseline comparability, such as age, gender, educational/professional level and other relevant variables. We included studies that we judged to have reasonable baseline comparability and to be similar in important demographic characteristics that might reasonably be thought to influence response to the intervention or otherwise affect outcomes.

Types of outcome measures

The basis for our classification of outcomes was the four-level typology first described by Kirkpatrick (Kirkpatrick 1967) and modified by Barr et al (Barr 2000):

- Level 1 outcomes refer to learners' reaction to the intervention, including participants' views of their learning experience and satisfaction with the programme.
- Level 2a outcomes refer to modification of attitudes and/or perceptions regarding responsible conduct of research.
- Level 2b outcomes refer to acquisition of knowledge and/or skills related to responsible conduct of research.
- Level 3 outcomes refer to behavioural change transferred from the learning environment to the workplace prompted by modifications in attitudes or perceptions, or the application of newly acquired knowledge/skills in practice. We further divided this level into:
 - 3a – behavioural intentions; and
 - 3b – actual change in research or publication practices, or both.
- Level 4 outcomes refer to organizational changes attributable to the intervention.

We included outcomes at the individual level (e.g. individual behaviour change) and as aggregated units of analysis (e.g. frequency of retracted articles).

There was no outcome measure or set of outcome measures that we could consider 'standard' for the purposes of this review as we expected a wide range of different outcomes to be found in included studies. Our intention was to classify them in a theoretically grounded way, so that we could meaningfully present the study results if meta-analysis was not appropriate due to the high heterogeneity of the studies. Kirkpatrick's four-level model (Kirkpatrick 1967) is a standard approach in educational research (Barr 2000). As we assessed interventions aimed at reducing/

preventing misconduct, we considered actual change in behaviour, either on the individual level (3rd level in Kirkpatrick's model) or at the organizational level (4th level) as a hierarchically higher (or 'better', more desirable) outcome than participants' satisfaction with an intervention (1st level). As 1st level outcomes are the easiest to assess, they are commonly used in educational research. However, they are the least informative and relevant, so we categorized them as secondary outcomes.

The inclusion of studies addressing perceptions and attitudes in the systematic review was based on Ajzen's theory of planned behaviour, in which the main predictors of behavioural intentions are attitudes towards the behaviour, subjective norms and perceived behavioural control (Ajzen 2005; Armitage 2001). Harms outcomes or potentially adverse effects were not expected and we did not assess them in this review.

Primary outcomes

We assessed the following primary outcomes:

- Primary outcome 1: Organizational change (level 4 outcome according to the Kirkpatrick/Barr typology).
- Primary outcome 2: Behavioural change (level 3 outcome according to the Kirkpatrick/Barr typology).
- Primary outcome 3: Acquisition of knowledge and/or skills (level 2b outcome according to the Kirkpatrick/Barr typology).
- Primary outcome 4: Modification of attitudes and/or perceptions (level 2a outcome according to the Kirkpatrick/Barr typology).

Secondary outcomes

Level of satisfaction or participants' experience with the intervention (level 1 outcome according to the Kirkpatrick/Barr typology).

Search methods for identification of studies

As the concepts of research integrity and responsible conduct of research emerged in the scientific community only after the establishment and active work of the Office for Research Integrity (ORI) in the USA in 1989 (Steneck 2006) and in Denmark in 1992 (Nylenna 1999), we limited our search to 1990 to December 2014.

Electronic searches

We searched the following bibliographic databases:

- Cochrane Central Register of Controlled Trials (CENTRAL, December 2014) via OvidSP;
- MEDLINE via OvidSP (1946 to December 2014);
- LILACS via BIREME (to December 2014);
- CINAHL via EBSCOhost (1981 to December 2014).

We also searched the following specialized or general electronic databases:

- [Academic Search Complete](#) – multi-disciplinary full-text and bibliographical database from the EBSCO Publishing platform (1887 to December 2014).
- [AGRICOLA](#), multidisciplinary database from the US National Agricultural Library, available via the OvidSP platform (1970 to December 2014).

- [GeoRef](#) – database from the American Geosciences Institute, available via the EBSCO Publishing platform (1933 to December 2014).
- [PsycINFO](#) – database from the American Psychological Association, available via the OvidSP platform (1806 to December 2014).
- [ERIC](#) – database of education literature, via OvidSP platform (1965 to December 2014).
- [SCOPUS](#) – citation database from Elsevier.
- [Web of Science \(WoS\)](#): SCI-EXPANDED, SSCI, A&HCI – citation database from Thomson Reuters.

We did not separately search the EMBASE bibliographic database because SCOPUS includes EMBASE data (Burnham 2006).

For the identification of studies to be included or considered for this review, we developed separate search strategies for each database searched. These were based on the search strategy developed for MEDLINE but revised appropriately for each database to take account of differences in the controlled vocabulary and syntax rules.

The subject search used a combination of controlled vocabulary, if appropriate, and free-text terms based on the search strategy for searching MEDLINE. The initial search strategy for MEDLINE, as presented in the protocol, was partly changed and further developed. The search strategies we used for each database are reported in [Appendix 1](#). We performed all searches in January 2013 and updated them in April 2015, searching for articles published between 1990 and 2014, inclusive. There were no language restrictions.

Searching other resources

We searched conference proceedings and abstracts in the following resources:

- Research presented at the Office of Research Integrity (ORI) Research Integrity Conferences (<https://ori.hhs.gov/conference>).
- Peer Review Congresses (<http://www.peerreviewcongress.org/previous.html>).
- Public Responsibility in Medicine and Research Conferences (<http://www.primr.org/conferences/>).
- Research Ethics site (<http://researchethics.ca/>).

We also searched a book on promoting research integrity by education (Institute of Medicine 2002) and publications from ORI-funded research, listed at: http://ori.dhhs.gov/research/extra/rri_publications.shtml.

Finally, we handsearched the electronic tables of contents of the following journals that regularly publish on research integrity topics:

- *Journal of Empirical Research on Human Research Ethics* (available online from volume 1 in 2006).
- *Science and Engineering Ethics* (available online from volume 1 in 1995).
- *Accountability in Research* (available online from volume 1 in 1989).
- *Ethics and Behavior* (available online from volume 1 in 1991).

- *Journal of Higher Education* (last available volume: 2002).
- *Journal of Medical Ethics* (available online from volume 1 in 1975).
- *Academic Medicine* (available online from volume 1 in 1926).
- *Medical Education* (available online from volume 1 in 1966).
- *Medical Teacher* (available online from volume 1 in 1979).
- *Teaching and Learning in Medicine* (available online from volume 1 in 1989).
- *Professional Ethics: A Multidisciplinary Journal* (available online from 1992; merged in 2004 with *Business and Professional Ethics*, which is available online since 1981).
- *American Psychologist* (available online from volume 1 in 1946).
- *Journal of Business Ethics* (available online from volume 16 in 1997).
- *Journal of Academic Ethics* (available online from volume 1 in 2003).

We also searched the references of all studies analysed in full text, as well as retrieved review articles, using both 'forward' (through citation databases such as Web of Science) and 'backward' (examining reference lists) citation searching ([Horsley 2011](#)).

Data collection and analysis

Following the execution of the search strategies, we collated the identified records (titles and available abstracts) in an EndNote database for de-duplication ([Thomson Reuters 2011](#)). We created the final unique record set and full text of potentially eligible studies as a separate EndNote file, from which we carried out screening of records and extraction of data from included articles.

Selection of studies

We expected a large volume of records from the initial database search because of the broad and general nature of many search terms. For this reason, we first conducted an initial screening of the titles only, including the screening of the abstract when the title was uninformative. In the next step, we screened the abstracts (and full text when needed, particularly in publications outside of biomedicine and health, which often had brief abstracts). After exclusions at these two steps, we reviewed all remaining potentially eligible articles in full text for eligibility, using a priori defined eligibility criteria. Two review authors (AM and EW) independently performed each step. Disagreements between review authors at the stage of full-text analysis were resolved by consensus or by a third member of the team (HRR).

Data extraction and management

Two members of the review team (AM and EW) independently extracted the data on study characteristics and outcome data. Disagreements were resolved by consensus or by consultation with the third team member (HRR). The review authors were not blinded to the authors, interventions or results obtained in the included studies.

We extracted and entered the following data in a customized collection form:

- Study design (e.g. randomized controlled trial, controlled before-and-after, etc.), date and length of follow-up.

- Participants: a) sample size and b) inclusion and exclusion criteria, demographic characteristics of participants: age, sex, country of origin, ethnicity, gender, field of research, academic level or research experience.
- Setting: type of institution or broader setting where the intervention(s) took place.
- Interventions: details of the type and duration of intervention and comparisons.
- Outcomes: detailed description of the outcomes of interest, including the method and timing of measurement.
- Source of funding.

We extracted results for pre-specified types of outcomes of interest. We extracted the raw data available in the study, such as means and standard deviations for continuous outcomes and number of events and participants for dichotomous outcomes. As meta-analysis of retrieved studies was not possible, we did not attempt to obtain original data if they were not adequately presented in the retrieved publication.

For studies that presented results in graphs ([Brown 2001](#); [Landau 2002](#)), we generated numerical values using digitizing software ([Huwaldt 2014](#)). For one study ([Brown 2001](#)), which presented data for individual scales but performed statistical analysis for composite scores, we averaged mean scores and standard deviations for composite scores using appropriate formula ([Headrick 2010](#)).

We designed the data extraction form for this review based on forms available from other Cochrane Review Groups, and piloted it before use. When several articles reported different outcomes of the same study, we considered them a single entry in the data extraction form.

Assessment of risk of bias in included studies

For randomized controlled trials, we assessed the risk of bias using Cochrane's 'Risk of bias' tool, which addresses the following domains: random sequence generation, allocation concealment, blinding (separately for researcher-assessed and self reported outcomes), incomplete outcome data, selective outcome reporting and other sources of bias.

For the assessment of non-randomized studies, we used a modified Cochrane 'Risk of bias' tool, in which we used four out of six original domains (blinding, incomplete outcome data, selective outcome reporting, other sources of bias) and two additional domains (comparability of groups and confounding factors) to assess the risk of bias in the included non-randomized studies. For the additional two domains we used the following questions for assessment: "Were the study groups comparable at baseline?" and "Were potential confounding factors adequately addressed?". As recommended in Chapter 13 of the *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2011](#)), we collected factual information on the confounders considered in the included studies and reported them in a table. We did not assess the risk of bias on the domains of sequence generation and concealment of allocation sequence, as a high risk on these domains is inherent in the design of non-randomized studies and therefore expectable by default. We participated in the pilot study for a Cochrane Risk Of Bias Assessment Tool for Non-Randomized Studies of Interventions - ACROBAT-NRSI tool ([Sterne 2014](#)), but found that it was not fully

suitable because the published articles on studies included in the review did not report many of the items needed for the NRSI tool.

For all study designs, we assessed compliance with intervention and possible contamination (spillover effect) between the groups under the domain of other sources of bias. We recorded each piece of information extracted for the 'Risk of bias' tool together with the precise source of this information. We first tested data collection forms and assessments of the risk of bias on a pilot sample of articles. Two assessors (AM and DS) independently carried out the assessment of risk of bias. The assessors were not blinded to the names of the authors, institutions, journal or results of a study. We tabulated risk of bias for each included study, along with a judgement of low, high or unclear risk of bias, as described in Chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Measures of the effect of the methods

In the meta-analyses planned in the protocol for this systematic review, we had intended to use standardized mean differences (SMD) with 95% confidence intervals. However, we decided that meta-analysis would not be appropriate because of the heterogeneity of the included studies.

Unit of analysis issues

The unit of analysis was generally the individual study participant. Where the unit of assignment to an intervention was not the individual, but some larger entity (e.g. a department or an institution), the unit of analysis was the larger entity.

Dealing with missing data

In the protocol, we planned to address the problem of missing data required for data synthesis or 'Risk of bias' assessment by contacting the authors to request the data (Young 2011). However, because most of the studies with missing information were published a long time ago, often did not show an e-mail address for the corresponding author and were published in fields outside biomedicine and health, we did not systematically contact study authors. In a single case where we contacted two authors of a study, we did not get a response.

Data synthesis

There were significant differences in the study designs and outcomes among included studies, therefore it was not possible to carry out a meta-analysis. The results are presented descriptively for individual studies, grouped according to four-level typology for educational outcomes (Barr 2000). This did not lend itself to building a 'Summary of findings' table. In 'Additional tables', we noted whether there was a significant effect or not for each outcome. We are aware that "vote counting" is not recommended in data analysis, but the tables make the findings more understandable within the context of our qualitative synthesis.

Quality of the evidence

We followed the GRADE approach in the assessment of the quality of evidence (Grade Working Group 2004). We considered limitations of included studies (risk of bias), indirectness of evidence, inconsistency (heterogeneity), imprecision of effect estimates and potential publication bias. We considered randomized controlled trials to start from 'high quality' and observational studies from 'low quality'. As we did not pool the data and the outcomes in the included studies were very diverse, we could only make a general statement on the quality of the evidence as a whole, taking into account the different GRADE domains. Also, we commented on possible factors affecting the quality of the evidence.

RESULTS

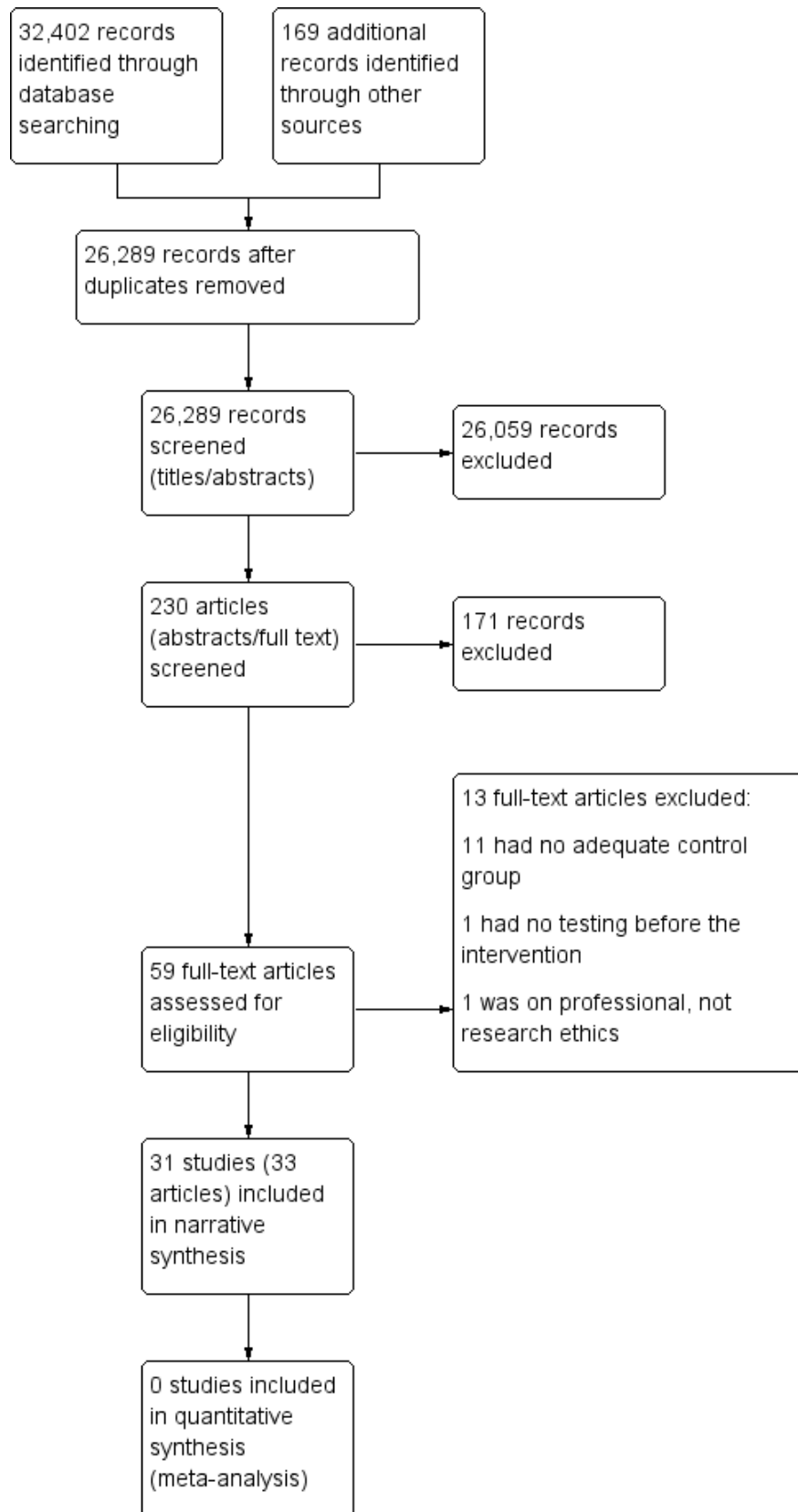
Description of studies

See: [Characteristics of included studies](#) and [Characteristics of excluded studies](#). We ordered the included studies so that randomized studies are presented first, followed by other types of study designs.

Results of the search

We identified 20,888 unique published records, of which we excluded 20,658 after title screening and a further 188 after abstract/full-text screening. This left 42 potentially eligible publications for detailed independent eligibility assessment (Figure 1).

Figure 1. Study flow diagram.



Included studies

Thirty-one studies involving 9571 participants, described in 33 articles, met the inclusion criteria. All were published in English. Fifteen studies reported the dates of the study, which ranged from 2001 to 2009 (Arnott 2008; Ballard 2013; Belter 2009; Bilić-Zulle 2008; Brown 2001; Dee 2012; Estow 2011; Ivaniš 2008; Kose 2011; Marshall 2011; Marušić 2006; Moniz 2008; Risquez 2011; Roberts 2007; Rolfe 2011).

Randomized studies

Fifteen studies were randomized controlled trials, involving 5275 participants (Aggarwal 2011; Ballard 2013; Brown 2001; Compton 2008; Dee 2012; Ivaniš 2008; Landau 2002; Marušić 2006; Moniz 2008; Newton 2014; Risquez 2011; Roberts 2007; Rose 1998; Schuetze 2004; Youmans 2011). One of these was presented in two published articles (Roberts 2007). Six had two intervention arms (Aggarwal 2011; Dee 2012; Ivaniš 2008; Newton 2014; Rose 1998; Schuetze 2004); four had three arms (Brown 2001; Marušić 2006; Moniz 2008; Roberts 2007), and four had four arms (Ballard 2013; Compton 2008; Landau 2002; Youmans 2011). All randomized controlled trials, except one (Aggarwal 2011), had only post-test evaluation without baseline measurement. The trials took place mostly in the USA (nine studies: Ballard 2013; Compton 2008; Dee 2012; Landau 2002; Moniz 2008; Roberts 2007; Rose 1998; Schuetze 2004; Youmans 2011), one each in Australia (Newton 2014), India (Aggarwal 2011), Ireland (Risquez 2011), and the UK (Brown 2001), and two were performed in an international general medical journal published in Croatia (Ivaniš 2008; Marušić 2006). Participants included undergraduate students from different disciplines, such as psychology, humanities, communication and entrepreneurship (10 studies: Ballard 2013; Brown 2001; Compton 2008; Dee 2012; Landau 2002; Moniz 2008; Newton 2014; Risquez 2011; Schuetze 2004; Youmans 2011), medical students (one study; Roberts 2007), graduate students in graduate students in physical, biological, engineering and social science fields (Rose 1998), and scientists/journal authors (three studies: Aggarwal 2011; Ivaniš 2008; Marušić 2006). The sample size ranged from 58 to 1462 participants (median 211, interquartile range (IQR) 91 to 398). Funding was reported for five out of the 15 randomized controlled trials; of the 10 studies without stated funding, one reported that the course was developed by a spin-off company. Interventions were related to plagiarism, research ethics courses and journal authorship.

In studies related to plagiarism, the comparisons included inoculation of educational statements about plagiarism in teaching material versus no intervention (Brown 2001), use of Turnitin in preventing plagiarism (Ballard 2013; Newton 2014), fear-based, guilt-based and reason-based messages versus no intervention (Compton 2008), online tutorial on understanding and avoiding plagiarism versus no intervention (Dee 2012), plagiarism detection exercise with feedback, examples or their combination versus no intervention (Landau 2002), direct lecture or student-centred, practical instruction on plagiarism versus standard teaching (Moniz 2008), in-class tutorial on plagiarism prevention versus no instruction (Risquez 2011), composite intervention of instruction and homework on plagiarism versus standard teaching on citations and referencing (Schuetze 2004), and brief paraphrasing instruction versus no intervention (Youmans 2011). The studies assessed a variety of outcomes (median of 2 per study, IQR 1 to 3).

Studies on plagiarism mostly investigated the effects of intervention on students' knowledge and attitudes toward plagiarism and its prevalence and prevention, as well as the extent of plagiarism in submitted student works (Ballard 2013; Brown 2001; Compton 2008; Dee 2012; Landau 2002; Moniz 2008; Newton 2014; Risquez 2011; Schuetze 2004; Youmans 2011).

Two studies investigated the effects of research ethics courses, one comparing an online versus on-site course and assessing the knowledge at the end of the course and three months after the course (Aggarwal 2011), and the other comparing criteria-oriented and participant-oriented teaching on evaluating the ethical soundness of research on humans versus no ethics instruction, assessing medical student's rating of the significance of ethical problems and their attitudes towards ethical conduct of clinical research (Roberts 2007). One study investigated the impact of authorship policies on students' judgements of authorship in graduate student-professor collaboration (Rose 1998). Finally, two studies in the same general medical journal investigated satisfaction of authorship criteria from the International Committee of Medical Journal Editors (ICMJE) in different situations of authorship declaration: one study compared a categorical or instructional declaration form versus an open-ended form for authorship declaration (Marušić 2006), and the other compared a binary (yes/no) to an ordinal rating scale of authorship contributions (Ivaniš 2008). In both studies the reporting outcome was the number of authors satisfying ICMJE authorship criteria.

Non-randomized studies

Sixteen studies had other study design types: nine were controlled before-and-after studies involving 2293 participants (Arnott 2008; Chertok 2014; Clarkeburn 2002; Estow 2011; Fisher 1997; Hull 1994; May 2013; Strohmets 1992; Walker 2008), four were non-equivalent controlled studies with a historical control involving 1897 participants (Belter 2009; Bilić-Zulle 2008; Marshall 2011; Rolfe 2011), one was a non-equivalent controlled study with a post-test only (Chao 2009) (116 participants) and two were non-equivalent controlled with pre- and post-test for the intervention group and post-test for the control group, involving 108 participants (Barry 2006; Kose 2011).

One study was presented in two published articles (Bilić-Zulle 2008). The non-randomized studies took place mostly in the USA (11 studies: Arnott 2008; Barry 2006; Belter 2009; Chao 2009; Chertok 2014; Estow 2011; Fisher 1997; Hull 1994; May 2013; Strohmets 1992; Walker 2008), two were performed in the UK (Clarkeburn 2002; Marshall 2011; Rolfe 2011), and one each in Croatia (Bilić-Zulle 2008) and Turkey (Kose 2011). Participants were mostly undergraduate students, predominantly from psychology (11 studies: Aggarwal 2011; Barry 2006; Belter 2009; Chao 2009; Clarkeburn 2002; Estow 2011; Fisher 1997; Kose 2011; Rolfe 2011; Strohmets 1992; Walker 2008); one study involved medical students (Bilić-Zulle 2008), one study involved health students (Chertok 2014), three involved postgraduate students, including masters degree students (Hull 1994; Marshall 2011; Rose 1998), and one had a mixture of undergraduate and postgraduate students (May 2013). The sample size ranged from 36 to 1085 participants (median 146, IQR 65 to 385.5). Funding was reported by six out of 16 studies. Interventions included prevention of plagiarism (10 studies) and research ethics training (six studies).

Plagiarism comparisons included different forms of paraphrasing exercises versus standard instruction on plagiarism (Barry 2006; Chao 2009; Chertok 2014; Kose 2011; Marshall 2011; Rolfe 2011; Walker 2008), as well as instructions on plagiarism and its prevention versus no intervention (Belter 2009; Bilić-Zulle 2008; Estow 2011). Research ethics training comparisons included dialogue-based computer tutoring systems versus standard teaching (Arnott 2008), ethics discussion embedded into science courses versus no intervention (Clarkeburn 2002), or embedded ethics training or stand-alone ethics training versus no ethics instruction (May 2013), ethics-enhanced discussions versus standard ethic course (Fisher 1997), research ethics course versus no intervention (Hull 1994), and role-playing in research ethics training versus no intervention (Strohmetz 1992).

The studies assessed different outcomes (median of 1 per study, 95% CI 1 to 2.5). In studies on plagiarism, the outcomes were mostly some form of estimation of plagiarism in submitted student work (Barry 2006; Belter 2009; Bilić-Zulle 2008; Chao 2009; Kose 2011; Marshall 2011; Rolfe 2011; Walker 2008), or students' knowledge about and perceptions of plagiarism (Barry 2006; Chertok 2014; Estow 2011). Interventions in research ethics training assessed knowledge (Aggarwal 2011; Fisher 1997; May 2013), moral

reasoning or socio-moral reflection (Clarkeburn 2002; Hull 1994; May 2013), ethical sensitivity (Clarkeburn 2002; Fisher 1997), or assessment of research study utility or cost in relation to ethics (Strohmetz 1992).

Five out of 31 studies also assessed satisfaction with or experience of the intervention (Aggarwal 2011; Dee 2012; Kose 2011; Moniz 2008; Rolfe 2011).

Excluded studies

We excluded studies described in 13 articles: eight did not have an adequate control group (Ali 2014; Bagdasarov 2013; Harkrider 2012; Kligyte 2008; McDonalds 2010; Mumford 2008; Powell 2007; Vallero 2007), two were on professional or academic ethics (excluding plagiarism) and not research ethics/integrity (Goldie 2001; Gurung 2012; May 2014), and two did not include testing before the intervention (Hren 2007; Pollock 1995).

Risk of bias in included studies

See Figure 2 and Figure 3 for illustration of the risk of bias rating for each study and across studies for each domain for randomized studies and Figure 4 and Figure 5 for non-randomized studies.

Figure 2. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included randomized studies.

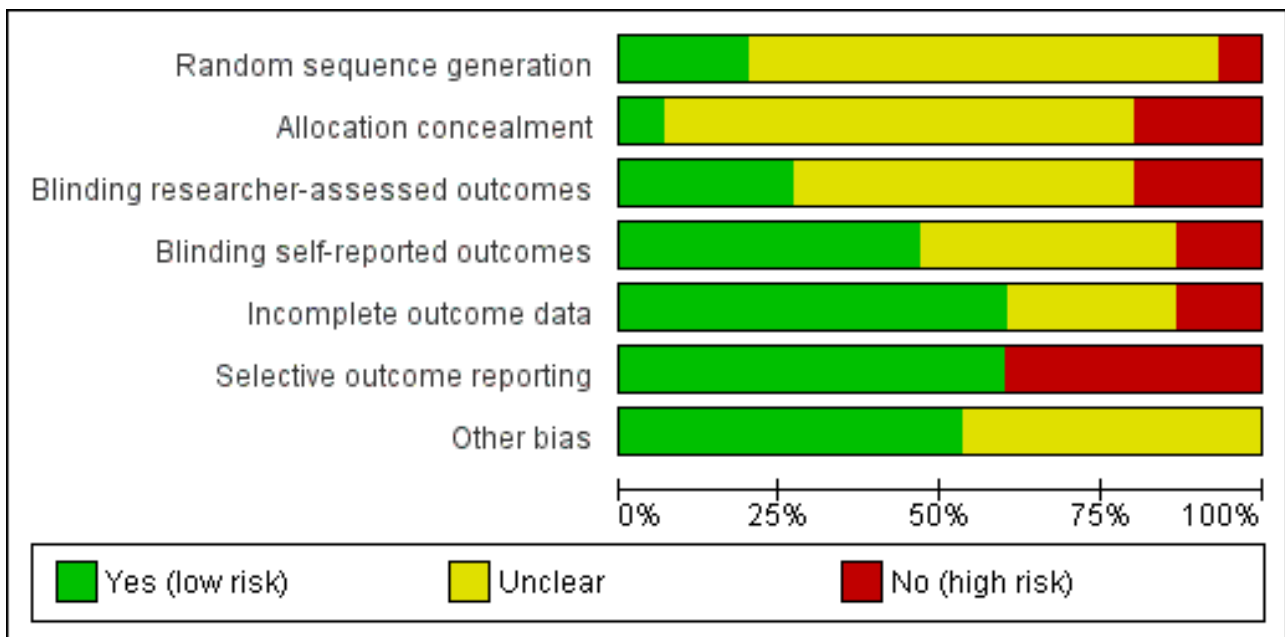


Figure 3. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included randomized study.

	Random sequence generation	Allocation concealment	Blinding researcher-assessed outcomes	Blinding self-reported outcomes	Incomplete outcome data	Selective outcome reporting	Other bias
Aggarwal 2011	+	?	?	-	+	+	?
Ballard 2013	?	?	?	+	?	+	?
Brown 2001	?	?	+	+	-	+	?
Compton 2008	?	?	?	?	?	+	+
Dee 2012	?	?	-	+	+	-	+
Ivaniš 2008	+	-	-	+	+	+	+
Landau 2002	?	?	+	?	?	-	+
Marušić 2006	+	-	-	+	+	+	+
Moniz 2008	?	?	?	?	+	-	?
Newton 2014	?	?	?	?	+	-	?
Risquez 2011	-	?	?	?	+	-	?
Roberts 2007	?	?	?	?	+	+	+
Rose 1998	?	?	?	+	-	+	+
Schuetze 2004	?	-	+	+	?	+	?
Youmans 2011	?	+	+	-	+	-	+

Figure 4. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included non-randomized studies.

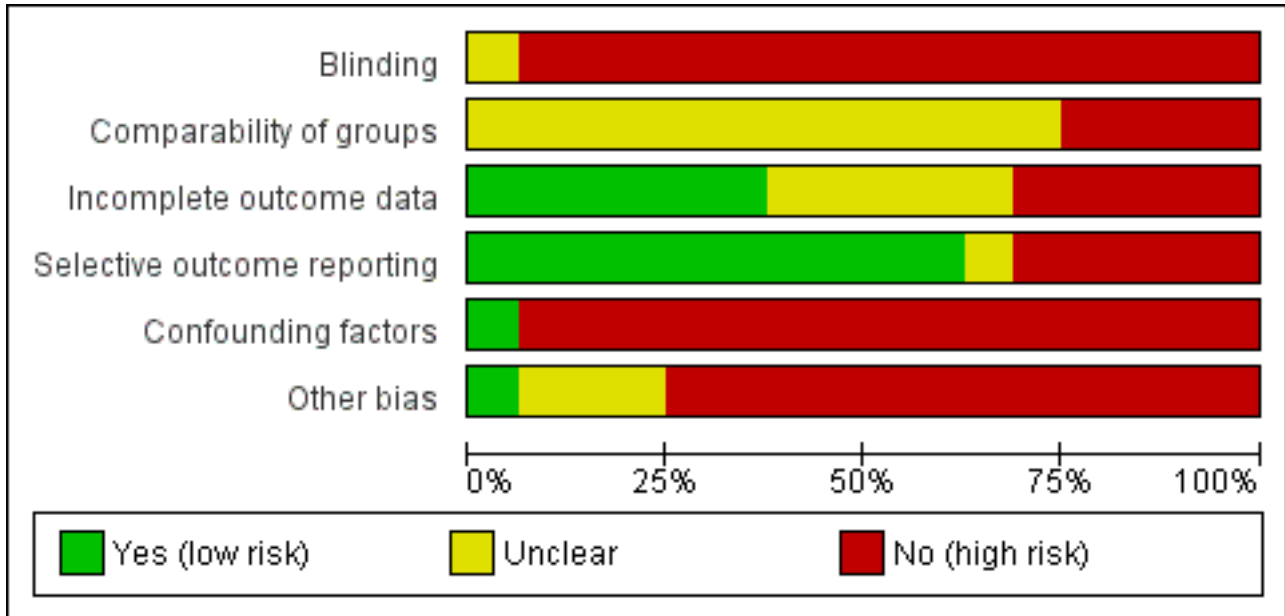


Figure 5. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included non-randomized study.

	Blinding	Comparability of groups	Incomplete outcome data	Selective outcome reporting	Confounding factors	Other bias
Arnott 2008	-	?	?	-	-	-
Barry 2006	-	?	-	-	-	-
Belter 2009	-	-	+	+	-	-
Bilić-Zulle 2008	-	-	+	+	-	-
Chao 2009	-	?	+	+	-	-
Chertok 2014	-	?	+	+	-	+
Clarkeburn 2002	-	-	-	?	-	-
Estow 2011	-	?	-	+	-	?
Fisher 1997	-	?	-	+	-	?
Hull 1994	-	?	?	-	-	?
Kose 2011	-	?	?	-	-	-
Marshall 2011	-	-	?	+	-	-
May 2013	?	?	-	+	+	-
Rolfe 2011	-	?	+	+	-	-
Strohmetz 1992	-	?	?	-	-	-
Walker 2008	-	?	+	+	-	-

Allocation

A. Randomized studies

Random sequence generation

Only three included randomized studies had a properly conducted and described random sequence generation procedure (Aggarwal 2011; Ivaniš 2008; Marušić 2006). Most other studies failed to describe the procedure and just mentioned that participants were randomly assigned (Ballard 2013; Compton 2008; Landau 2002; Moniz 2008; Newton 2014; Roberts 2007; Rose 1998; Schuetze 2004; Youmans 2011). One study described the procedure as block randomization with pairing on baseline traits, which we judged to be of insufficient clarity (Dee 2012). Another study distributed the intervention and control booklets from a "randomized pile", without explaining how the pile was randomized (Brown 2001). In one study, some intervention groups were assigned randomly and some non-randomly, resulting in a high risk of bias in this domain (Risquez 2011).

Allocation concealment

No study described an attempt to conceal the allocation sequence. In one case, the selection of participants inferred that the allocation sequence would not be known to the researchers (Youmans 2011). In two randomized controlled trials, authors of this review who had participated in the included study confirmed that there was no allocation concealment, so we rated the risk of bias for these studies as high (Ivaniš 2008; Marušić 2006). For another study, it was clear from the report that no allocation concealment was attempted (Schuetze 2004). Although the risk of bias in this domain for all other studies had to be marked as unclear due to lack of detail provided in the reports, it seems likely that the actual risk of bias in this domain was high for all studies.

Blinding

A. Randomized studies

Researchers or outcome assessors were described as being blinded in four randomized controlled trials (Brown 2001; Landau 2002; Schuetze 2004; Youmans 2011), and participants were blinded in seven studies (Ballard 2013; Brown 2001; Dee 2012; Ivaniš 2008; Marušić 2006; Rose 1998; Schuetze 2004). Some studies either explicitly stated that there was no blinding of researchers (Dee 2012; Marušić 2006) or participants (Youmans 2011), or it was clear from the report that there was no blinding of researchers (Aggarwal 2011; Ivaniš 2008) or study participants (Aggarwal 2011). The rest of the studies were reported in insufficient detail to discern whether the researchers (Rose 1998), participants (Landau 2002), or both researchers and participants (Compton 2008; Moniz 2008; Newton 2014; Risquez 2011; Roberts 2007), were blinded.

B. Non-randomized studies

In all but one included non-randomized study there was no active attempt by the researchers to blind the participants or outcome assessors, so risk of bias in this domain had to be judged as high. Participants were reportedly blind to the research hypothesis in only one study, but there was no description of how blinding was achieved, so we judged the risk of bias for this study to be unclear (May 2013). However, in many cases the study design was such that participants were probably not aware that there was another intervention or control group or even that they were participating in a study, so the effect of non-blinding was probably marginal.

Incomplete outcome data

A. Randomized studies

In most of the randomized controlled trials the attrition was small and we judged the risk of bias due to incomplete outcome data to be low (Aggarwal 2011; Brown 2001; Dee 2012; Ivaniš 2008; Marušić 2006; Moniz 2008; Newton 2014; Risquez 2011; Roberts 2007; Youmans 2011). The low attrition rate could in many cases be ascribed to the study design, such as when interventions were associated with mandatory courses that students had to complete in order to earn their grades. Data on the number of participants were reported insufficiently or inconsistently in four studies (Ballard 2013; Compton 2008; Landau 2002; Schuetze 2004), so the risk of bias in this domain had to be judged as unclear. In one study the response rate was low, with a high risk of bias due to incomplete outcome data (Rose 1998).

B. Non-randomized studies - incomplete outcome data

In six non-randomized studies the attrition was non-existent or very small (Belter 2009; Bilić-Zulle 2008; Chao 2009; Chertok 2014; Rolfe 2011; Walker 2008). In another five studies the risk of bias due to incomplete outcome data was unclear, as the attrition was not clearly reported (Hull 1994; Kose 2011; Marshall 2011; Strohmets 1992), or there were some unexplained imbalances in attrition rates between the study arms (Arnott 2008). A high risk of bias due to a considerable attrition rate was present in five studies (Barry 2006; Clarkeburn 2002; Estow 2011; Fisher 1997; May 2013), although in one study there was an attempt to analyse and explain such high attrition rates (May 2013).

Selective reporting

A. Randomized studies

All outcomes were adequately reported in most of the included randomized controlled trials (Aggarwal 2011; Ballard 2013; Brown 2001; Compton 2008; Ivaniš 2008; Marušić 2006; Roberts 2007; Rose 1998; Schuetze 2004). In one study only results of statistical analyses were reported, without absolute numbers (Dee 2012). We judged five studies to have a high risk of bias due to selective outcome reporting, as one or more outcomes of interest were reported incompletely and could not be entered in a meta-analysis. In one of those studies, means were presented without standard deviations, non-significant results were not reported, and the results for one of the outcomes was presented only for all participants and not for each studied group (Landau 2002). One study failed to report the numerical results altogether (Moniz 2008), and in another means were reported without standard deviations (Risquez 2011). Two studies did not report the number of participants per group (Newton 2014; Youmans 2011), and one of them did not report non-significant outcomes (Newton 2014).

B. Non-randomized studies

In 10 included non-randomized studies all outcomes were adequately reported (Belter 2009; Bilić-Zulle 2008; Chao 2009; Chertok 2014; Estow 2011; Fisher 1997; Marshall 2011; May 2013; Rolfe 2011; Walker 2008). In six studies, we judged the risk of bias due to selective outcome reporting to be high. Three of these studies only reported results of statistical analyses without providing scores by study group (Arnott 2008; Barry 2006; Strohmets 1992), one reported only means without standard deviations (Hull 1994), one did not fully report all of the outcomes

(Clarkeburn 2002), and one reported results only as percentages and provided ranges without medians (Kose 2011).

Other potential sources of bias

A. Randomized studies

We identified no further source of bias in seven randomized controlled trials (Compton 2008; Dee 2012; Ivaniš 2008; Landau 2002; Marušić 2006; Roberts 2007; Rose 1998; Youmans 2011). The outcome measurement instrument had not been validated (or was only partially validated) in six studies (Aggarwal 2011; Brown 2001; Moniz 2008; Newton 2014; Risquez 2011; Schuetze 2004). Course attendance rate was not reported in two studies (Aggarwal 2011; Ballard 2013), and there was a risk of contamination between the intervention and control groups in two studies (Ballard 2013; Schuetze 2004). In two studies there was an unclear or questionable baseline comparability between the study arms due to the units of randomization (Moniz 2008), or inadequate randomization procedure (Risquez 2011). We judged the risk of bias to be unclear for all studies in regard to other possible sources of bias.

B. Non-randomized studies

Comparability of groups

Comparability of groups was unclear in 13 out of 16 non-randomized studies (Arnott 2008; Barry 2006; Chao 2009; Chertok 2014; Estow 2011; Fisher 1997; Hull 1994; Kose 2011; May 2013; Rolfe 2011; Strohmets 1992; Walker 2008). There was either no comparison of demographic characteristics between the groups, or the number of compared characteristics was very limited. Often the participants were recruited from different study programmes, years or even universities. In some cases these differences were such that we judged the risk of bias due to questionable comparability of groups to be high (Belter 2009; Bilić-Zulle 2008; Clarkeburn 2002; Marshall 2011). Risk of bias under this domain was considerable for all the included non-randomized studies, as baseline differences between the study groups were likely even among the studies whose risk of bias had to be judged as unclear due to insufficient information provided in the report.

Confounding factors

All but one of the non-randomized studies failed to adequately control for confounding factors and we judged them to have a high risk of bias under this domain. A single study reported basic demographic characteristics of participants and used them in a regression analysis (May 2013). Although some other studies collected and reported basic demographic characteristics, none of them actually used the data to control for confounding factors.

Other sources of bias

While for a single study we could not identify other sources of bias (Chertok 2014), most of the non-randomized studies had a high risk of bias related to other reasons, such as unvalidated outcome measurement instrument (Arnott 2008; Barry 2006; Belter 2009; Bilić-Zulle 2008; Chao 2009; Kose 2011; Rolfe 2011; Strohmets 1992), poor or unknown attendance rate (Arnott 2008; Barry 2006; Clarkeburn 2002; Kose 2011; Marshall 2011; May 2013), interventions provided by different instructors or in different settings (Arnott 2008; Belter 2009; Chao 2009; Marshall 2011; May 2013; Strohmets 1992), potentially biased selection of participants in study arms (Arnott 2008; May 2013), failure to conduct both pre-test and post-test of all the participants (Barry 2006; Belter

2009; Bilić-Zulle 2008; Kose 2011; Rolfe 2011; Walker 2008), possible contamination between the groups (Kose 2011), and the use of a historical control (Bilić-Zulle 2008; Marshall 2011).

In three studies there were no other possible sources of bias except unvalidated outcome measurement instruments and/or unknown attendance rate (Estow 2011; Fisher 1997; Hull 1994). For those studies, we rated the risk of bias under this domain as unclear, rather than high.

Effect of methods

We assessed the effects of interventions for research integrity using outcome classification levels according to the Kirkpatrick/Barr typology (Barr 2000; Kirkpatrick 1967).

Primary outcomes

1. Organizational change attributable to the intervention

No studies tested interventions directed to organizational changes for preventing or fostering research integrity.

2. Behavioural change

One randomized controlled trial evaluated participants' intentions for a behavioural change (level 3a outcome) regarding research ethics (Rose 1998), showing that male graduate students who were aware of authorship policy when deciding on a case of authorship problems in student-faculty collaboration were more likely to report the professor who took first authorship on a student dissertation than students who were not informed about the policy. Female students aware of the policy were less likely to report such a case than those who were not aware of the policy. The awareness of policies had no effect on students' estimates of effectiveness or consequences of such actions.

Sixteen studies evaluated 21 different outcomes of actual behavioural change (level 3b outcome), related to research ethics/integrity interventions (in two studies) or plagiarism (in 14 studies). The results of the studies are summarized in Table 1.

a) Research ethics/integrity

Two randomized controlled trials in a single general medical journal evaluated whether changes in the authorship declaration form could influence authorship qualifications according to the standard definition of authorship in biomedicine from the International Committee of Medical Journal Editors (ICMJE). Fewer authors completing the instructional declaration form (which instructed the respondents about the number of contributions needed to satisfy ICMJE authorship criteria) failed to satisfy ICMJE criteria in comparison to authors who either answered an open-ended question about their contributions or had to pick from a list of predefined categories of contributions (Marušić 2006). When authors chose their contributions from a list of contribution categories, more of them satisfied the ICMJE authorship criteria when they could indicate the extent of their contribution than when they had to provide a binary (yes-no) response (Ivaniš 2008).

b) Plagiarism prevention

In a randomized controlled trial undergraduate students who received instruction about plagiarism definition and examples of a plagiarized text used fewer overlapping words and two-word or three-word strings in their assignments than students

receiving feedback about why the text overlap was considered plagiarism or both the example and the feedback (Landau 2002). In another randomized controlled trial, a one-hour in-class tutorial to undergraduate students on plagiarism prevention was not effective in comparison to no intervention in decreasing reported engagement in plagiarism (Risquez 2011). In another randomized controlled trial, undergraduate students who had to do a homework assignment of recognizing plagiarism after receiving a 30-minute presentation and handouts with plagiarism definition and citation guidelines had fewer citation problems in their term papers than students without the homework assignment (Schuetze 2004). In the Newton 2014 randomized controlled trial, a short training session on paraphrasing, patch-writing and plagiarism increased the performance of students in assignment writing, as measured by referencing source material, avoiding patch-writing and avoiding plagiarism in comparison to no intervention. In the Ballard 2013 randomized controlled trial, an interactive academic integrity module addressing plagiarism and paraphrasing, with or without examples from Turnitin text-matching software, was not successful in decreasing similarity index in course assignments by undergraduate students in comparison to no intervention. In the Youmans 2011 randomized controlled trial, warning about the use of Turnitin did not have an effect on plagiarism level in undergraduate students' writing assignments, regardless of whether they were instructed to cite sources or if citing was not mandatory.

In a controlled before-and-after study (Estow 2011), undergraduate students from a methods and statistics course with embedded plagiarism themes and assignments did not fare better than students from the control course in the quality of paraphrasing in their written assignments. In the Walker 2008 study, paraphrasing training to avoid plagiarism decreased the extent of plagiarism in the students' written course assignments compared to no training.

Two studies used a modification of the controlled before-and-after study design, with the control group of undergraduate students receiving post-intervention testing only. In the Chao 2009 study, compared to the no intervention group that was tasked with writing a report and warned about plagiarism, active training in preventing plagiarism involving extensive discussion and student practice work was more effective than active training involving instructions and examples in decreasing the proportion of students who had some extent of plagiarized course assignments. However, none of the interventions was effective in decreasing the mean proportion of plagiarized text in the assignments. In the Kose 2011 study, undergraduate students using Turnitin plagiarism detection software in writing their course works had a lower level of plagiarism in submitted essays than students without experience with Turnitin.

Four studies used a historical control in their before-and-after study design. Undergraduate students who completed an online academic integrity module during a course had fewer instances of plagiarism in their course papers than students attending a regular course (Belter 2009). Undergraduate students who were warned during the course that their essays would be examined for plagiarism and that plagiarism would be penalized had less plagiarized text in their essays than students who either received no warning or received an explanation about plagiarism and how to avoid it (Bilić-Zulle 2008). Similarly, warning about the use of Turnitin for checking students' assignments and additional

interactive seminars on plagiarism decreased the number of plagiarism occurrences and the amount of plagiarized text (Marshall 2011). Finally, use of Turnitin feedback to help students revise their course essays before final submission did not decrease the number the number of students who plagiarized compared to the student group without intervention. The number of submissions with plagiarism due to poor paraphrasing decreased, but not due to other causes (Rolfe 2011). Actually, plagiarism due to failure to acknowledge sources, measured as the lack of citations in the text, increased in the intervention group.

3. Acquisition of knowledge and/or skills related to responsible conduct of research

Ten studies evaluated 12 different outcomes at this level, related to research ethics training (in three studies) or plagiarism (in six studies). The results of individual studies are summarized in Table 2 and Table 3, respectively.

a) Research ethics/integrity

In a randomized controlled trial (Aggarwal 2011), onsite and online courses for researchers in research ethics increased knowledge immediately after the course, but with no difference in the median knowledge gain from the baseline to three months.

Two studies used a controlled before-and-after design to measure knowledge gain in research ethics/integrity. Undergraduate students receiving a dialogue-based computer tutoring system on research methodology, including research ethics, had greater knowledge gain than students receiving standard teaching (Arnott 2008). In the Fisher 1997 study, students attending courses with embedded research ethics modules had better knowledge of research ethics procedures and ability to weigh scientific responsibility and participant rights and welfare compared to standard ethics instruction.

b) Plagiarism prevention

Two randomized controlled trials failed to show positive effects of training undergraduate students in plagiarism prevention. There was no difference between direct instruction and student-centred teaching in improving student's functional understanding of plagiarism (Moniz 2008). Students who received a combination of lectures, handouts on plagiarism definition and citation guidelines, and a homework assignment to practise recognizing plagiarism, had similar scores on the knowledge test and understanding of the importance of proper citation in writing assignments to students without such training (Schuetze 2004).

In the Landau 2002 randomized controlled trial, when undergraduate students received instruction about plagiarism definition and examples of a plagiarized text, feedback about why the text overlap was considered plagiarism or both the example and the feedback, they had better knowledge about plagiarism, measured as Plagiarism Knowledge Survey (PKS) score, than students receiving no intervention.

In the Newton 2014 randomized controlled trial, short training session on paraphrasing, patch-writing and plagiarism increased referencing knowledge of undergraduate students in comparison to no intervention.

In a controlled before-and-after study (Estow 2011), undergraduate students from a methods and statistics course with embedded

plagiarism themes and assignments were better than students from the non-intervention course in identifying plagiarism in test examples. They also listed a similar number of strategies to avoid plagiarism while students from the control group listed fewer strategies after the course.

Practice of paraphrasing and source citation over six weeks of a course increased a composite score in a plagiarism definition test in undergraduate students compared to students from the control group, who were tested only after the course (Barry 2006).

The Chertok 2014 study used a controlled before-and-after design to demonstrate that a course on academic integrity using a combination of in-person and Internet-based lectures, discussion and participation did not increase knowledge about plagiarism.

4. Modification of attitudes and/or perceptions

Thirteen studies evaluated 34 different outcomes at this level, related to either research ethics training (in six studies) or plagiarism prevention (in seven studies).

a) Research ethics/integrity

Two randomized and four controlled before-and-after studies explored effects of interventions for fostering research ethics. The results are summarized in Table 4.

In a randomized controlled trial (Roberts 2007), medical students who received criteria-oriented (analytic-focused) instruction about ethically important considerations in human clinical research assigned greater significance to ethical problems presented in 10 vignettes than students receiving participant-oriented (empathy-focused) instruction or no instructions. In contrast, participants who received empathy-focused instruction but not those with analytic-focused instruction expressed stronger agreement with statements regarding research participant decision-making than the control group. They also had more positive attitudes toward towards the decisional capacity of seriously ill people as research participants: students in the participant-oriented (empathy) intervention considered that research participants with serious illness would be much less accepting of surrogate decision-making by research doctors than by family members, which was not the case for the criteria-oriented intervention group or the control group; they also considered that people with serious illness were much less accepting of surrogate decision-making by research doctors than by family members. The interventions had no effect on research participants' attitudes concerning research participants' vulnerability and acceptance of surrogate decision-makers, as well as ethically relevant attitudes related to clinical research, and the importance of the ethical duties of investigators.

A randomized controlled trial of the effect of authorship policies on students' judgement of authorship abuse in student-faculty collaboration, demonstrated different effects of policies on perceptions of male and female students (Rose 1998). Whereas the presence of an authorship policy did not affect male students' perception of the ethicality of assigning first authorship, female students had less positive attitudes to first authorship for a professor when they were aware of the professional policy, which favours student first authorship in collaborations leading to a dissertation. The policy had no effect on students' attitudes about the ethicality of professors to submitting a manuscript to a journal without discussing authorship with the student.

In a controlled before-and-after study (Clarkeburn 2002), undergraduate students receiving two-hour, face-to-face ethics and research integrity discussion within a science course increased in ethical sensitivity compared to the control group, which did not receive the intervention, but not in moral reasoning skills as measured by the Defining Issues Test (DIT) or meta-ethical understanding.

In the Hull 1994 study, a stand-alone course on research ethics, which combined lecture and discussion group formats, increased socio-moral reasoning of graduate students compared with students who did not attend the course. In the May 2013 study, undergraduate students receiving either stand-alone or course-embedded ethics modules showed enhanced perspective-taking compared to students receiving no intervention, as well as moral efficiency and moral courage, but there was no change in moral reasoning (as tested by the DIT) or moral meaningfulness. In the Strohmetz 1992 study, use of a role-play exercise in teaching students about the ethical complexity of research in comparison to a standard research ethics course increased students' perceptions of the ethical utility and decreased their perceptions on the ethical cost of problematic research studies.

b) Plagiarism prevention

Six studies used a randomized design and one used a controlled before-and-after design to test interventions to change perceptions of plagiarism. The results are summarized in Table 5.

In the Brown 2001 randomized controlled trial, students who received a detailed educational description of plagiarism and how it can be avoided scored higher on the ratings for severity of behaviour in examples of plagiarized text (verbatim copying or unacknowledged close paraphrasing) compared with students who received a simpler warning statement on plagiarism or no information when they were asked about a) their own views, b) perceived staff view and c) perceived frequency of behaviour among colleagues. Neither intervention influenced participants' perception of the necessity for citing sources when writing or their perception of the likelihood that colleagues would cite a source when writing and assignment.

In the Schuetze 2004 randomized controlled trial, undergraduate students who had to do a homework assignment of recognizing plagiarism after receiving a 30-minute presentation and handouts with plagiarism definition and citation guidelines had better perception of their own understanding of plagiarism and greater confidence in their ability to avoid plagiarism in future compared with the control group, which received the instruction but had no homework.

In a cluster-randomized controlled trial (Dee 2012), use of an online tutorial on understanding and avoiding plagiarism changed students' perception that they had good understanding of plagiarism in academic writing but not the perception of their actual behaviour of avoiding plagiarism or their confidence in knowledge of how to avoid it, or whether the instructor would notice or address plagiarism in writing assignments.

The Compton 2008 randomized controlled trial used a two-stage intervention to test the effect of messages about plagiarism (control, fear-based, guilt-based and cognitive-based experimental conditions) and plagiarism justification on the attitudes of undergraduate students. There was no significant effect of the

experimental condition on the attitudes towards plagiarism or its justification.

The [Newton 2014](#) randomized controlled trial demonstrated that a short training session on paraphrasing, patch-writing and plagiarism increased the confidence of undergraduate students in note-taking from sources and in their referencing but not in paraphrasing source material in course assignments in comparison to no intervention.

The [Risquez 2011](#) randomized controlled trial showed that an online, computer-based, one-hour tutorial on plagiarism prevention completed individually as a part of a course was not successful in changing the attitudes of undergraduate students towards plagiarism. The intervention increased their perception of the seriousness of breaching academic guidelines, either as their own view or as their assessment of lecturers' view but not their expectations from students in intervention examples.

The [Chertok 2014](#) study used a controlled before-and-after design to demonstrate that a course on academic integrity using a combination of in-person and Internet-based lectures, discussion and participation did not improve attitudes about academic integrity, including plagiarism.

Secondary outcome

1. Participants' reaction to the intervention

The results of seven studies that evaluated participants' reaction to interventions are summarized in [Table 6](#).

a) Research ethics/integrity

In a randomized controlled trial ([Aggarwal 2011](#)), participants were more satisfied with an on-site compared to an online course in research ethics, particularly because of the value of group discussions and the value of the course for work in biomedical research. They also reported that the speed of covering the course material in the online research ethics course was too fast.

In the [Fisher 1997](#) study, students and staff favourably responded to research ethics modules embedded within a standard course and considered that teaching and testing materials were appropriate. Students agreed that the ethics topics discussed in the modules increased their interest in research ethics, and staff agreed that ethics modules complemented and enhanced the curriculum.

b) Plagiarism prevention

In a randomized controlled trial of a web-based tutorial for reducing plagiarism ([Dee 2012](#)), the intervention had no effect on whether students liked the class, found it academically difficult or stressful, or procrastinated on writing assignments for the course. In another randomized study comparing different instructions about plagiarism ([Moniz 2008](#)), students were most satisfied with student-centred instruction involving group and role-playing, compared with lecture- and practice-based direct instruction or standard teaching, when assessing the interactivity of the course, seatwork exercises and group work.

In two studies of using Turnitin software to reduce plagiarism in student course work, participants' satisfaction was high. In the [Kose 2011](#) study, 81% of the students using the software reported that the software was easy to use and all thought that it was very useful for checking their course essays, and in the [Rolfe 2011](#), all

students reported positive experience in the formative feedback they received from Turnitin during drafting and final submission of their course essays. In the [Walker 2008](#) study, students who had paraphrasing practice to avoid plagiarism reported that the training was useful and helpful.

DISCUSSION

Summary of main results

Our systematic review identified a range of interventions aimed at reducing research misconduct. Most interventions involved some kind of training, but methods and content varied greatly. The training included face-to-face and online lectures, interactive online modules, discussion groups, homework and practical exercises. The success of the interventions was assessed on participants' attitudes, knowledge and behaviour. Most studies did not use standardized or validated outcome measures, although a few used the Plagiarism Knowledge Survey ([Roig 1997](#)). Participants included undergraduates, postgraduates and academics from a range of research disciplines and countries. It is therefore difficult to synthesize findings from studies with such diverse interventions, outcomes and participants.

Briefly, and at the risk of oversimplifying these complex findings:

- various methods of training in research integrity had some effects on participants' attitudes to ethical issues (two studies showed positive effects, four showed inconsistent effects) but minimal (or short-lived) effects on their knowledge (three studies);
- training about plagiarism and paraphrasing had varying effects on participants' attitudes towards plagiarism and their confidence in avoiding it. Training that included practical exercises appeared to be more effective (two positive studies, two with mixed effects, two with no effect);
- training on plagiarism had inconsistent effects on participants' knowledge about, and ability to recognize, plagiarism (four positive studies, three negative); active training, particularly if it involved practical exercises or use of text-matching software, generally decreased the occurrence of plagiarism although results were not consistent (nine positive studies, four negative studies, one with mixed effects);
- the design of a journal's author contribution form affected the truthfulness of information supplied about individuals' contributions and the proportion of listed contributors who met authorship criteria (two studies).

No studies assessed interventions designed to reduce other types of misconduct such as data fabrication or falsification. Our search did not include studies specifically dealing with training of research ethics committees, where the need for systematic training has recently been identified ([Mhaskar 2015](#)).

Overall completeness and applicability of evidence

Many of the interventions assessed in the studies were complex, such as lecture programmes, workshops or online training. In most cases these interventions were inadequately described, so it would be impossible to repeat the experiments or implement the intervention that had been tested. It seems reasonable to assume that many factors can influence the effectiveness of training, such as the skills of the trainer, the quality and relevance

of the material, and the training medium (online or face-to-face) and environment (group size, timing of sessions). The 'dose' of the training (length of sessions, number of sessions) may also influence effectiveness. Furthermore, external factors may influence attitudes and behaviour (e.g. a well-publicized or local misconduct case may raise awareness of the penalties for misconduct), making historical controls difficult to interpret. Not only do these factors make it difficult to combine or compare findings between studies, they also decrease the applicability of the evidence.

It was not possible to estimate the completeness of reporting or the incidence of publication bias since educational studies are rarely registered.

We cannot tell if the findings are applicable outside the populations included in the studies, which were mainly undergraduate students. Factors affecting the behaviour and attitudes of other populations, such as more experienced researchers, may be different.

Quality of the evidence

We found the quality of included studies to be highly variable. Methods and participants were often poorly described. Some studies used randomized designs with reasonable, contemporaneous controls. Many studies used weaker methods such as controlled before-and-after designs. Few studies on attitude or knowledge used standardized or validated outcome measures, making findings difficult to interpret or compare. Studies on plagiarism tended to use more robust designs and employed more objective outcomes such as the percentage of plagiarized text in an assignment.

Most included studies were short-term, measuring effects directly after training. One study showed an effect immediately after training but this had disappeared three months later. It is uncertain whether training of undergraduates or junior researchers will have lasting effects.

Overall, the quality of evidence for the effectiveness of interventions to prevent misconduct and promote integrity in research and publication was very low. Our downgrading of the quality of evidence, according to the GRADE approach, is due to high risk of bias, indirectness and imprecision of the evidence.

We judged most of the included randomized controlled trials to be at a high risk of bias in at least one of the assessed domains. In the case of the non-randomized trials, there were no attempts to alleviate the potential biases inherent in the non-randomized design.

The heterogeneity in the results of the studies ('inconsistency' in GRADE) was considerable, with large variation in the direction and magnitude of effects. However, much of this heterogeneity can be explained by differences in the interventions and outcomes.

The available evidence is indirect, as the majority of studies included only students (mostly undergraduate), while only a few studies had experienced researchers as participants. Also, most of the studies assessed surrogate outcomes (such as reaction to intervention, change in attitudes and knowledge) rather than outcomes relevant for practice (actual behavioural change), which were mostly limited to plagiarism in students' writing tasks.

The number of participants in the included studies was generally small, and because the heterogeneity was too large to allow pooling of results, we were unable to reduce this imprecision through meta-analysis.

We did not attempt a formal analysis of publication bias or other biases related to study size by assessing the asymmetry of a funnel plot. However, selective publication of positive findings related to commercial interests is not likely in this area of research.

It is important to keep in mind that research integrity and research misconduct are terms that are relatively new in the field of professional ethics (Komic 2015), and that there is little agreement on the goals and content of training for responsible conduct of research (RCR) (Kalichman 2014). Also, research misconduct is difficult to detect and measuring subtle changes in behaviour related to responsible conduct of research usually relies on surrogate or self-reported outcomes, or both (Fanelli 2009; Marusic 2011). This is the reason why it is difficult to design interventions to prevent research misconduct and foster research integrity.

Potential biases in the review process

Our initial literature search was designed to be broad and inclusive, since few specific keywords are available for studies of the type we were seeking. It therefore resulted in a large number of titles to check (more than 20,000). Although two review authors checked these titles, it is possible that some eligible studies were missed, although such omissions are unlikely to have led to any systematic bias in the review. Also, we did not plan or conduct a systematic search of grey literature (such as unpublished reports, dissertations or theses). Due to the range of study designs and deficient reporting in the publications, eligibility of the studies was discussed extensively usually among three or four of the review authors, including an expert methodologist, to ensure that only study designs specified in the protocol (Marusic 2013) were included in the review.

Some of the included studies were co-authored by the authors of this review, which may have introduced bias in the review process. To avoid this, the assessment of eligibility, data extraction and 'Risk of bias' assessment for those studies was done independently by at least one member of the review author team who was not a co-author of the included study.

Agreements and disagreements with other studies or reviews

Antes et al published a meta-analysis of the effectiveness of ethics instruction in sciences, in which they included 20 studies of different study designs (Antes 2009). Out of those, we included one study in this systematic review (Clarkeburn 2002), and excluded two studies after full-text assessment for eligibility (Kligyte 2008; Mumford 2008). We excluded other studies from this systematic review during abstract/full-text screening (Figure 1) because they did not fit the inclusion criteria, or were published before 1990. The scope of the Antes 2009 meta-analysis was to analyse training and instruction in ethics in sciences (also including professional ethics). They performed a meta-analysis based on expert reviewer opinion of the outcome of the studies, rather than by combining the actual outcomes. They concluded that the overall effectiveness of ethics instruction was modest, but that programmes were more likely to be effective if they were conducted as separate rather than

embedded activities in already existing courses, or if they were interactive and based on real world examples.

AUTHORS' CONCLUSIONS

Implication for methodological research

One of the major difficulties in attempting to synthesize the findings from this review is the lack of standardized and validated outcome measures, especially those assessing attitudes towards research integrity and misconduct. The Plagiarism Knowledge Survey is a rare example of a validated tool to assess an outcome relevant for research integrity (Roig 1997). However, it assesses only one aspect of research misconduct, namely plagiarism, and it is still not widely used among researchers. Development of other validated tools, especially to assess practice-relevant outcomes such as behavioural change, would greatly contribute to the advancement of knowledge in the field of research integrity.

Another problem in interpreting the findings of studies in this area is our lack of knowledge of the prevalence of various forms of research and academic misconduct. We excluded studies of student cheating (e.g. in exams) from this review but these show varying rates across disciplines and countries. While we included studies on undergraduate plagiarism, as we considered them relevant to other types of plagiarism (e.g. in scholarly journals), few studies interpreted their findings in the context of an overall or expected rate of plagiarism. One further difficulty is that definitions of

plagiarism vary and, while text-matching software is an invaluable screening tool, institutions generally recognize that output from such tools must be interpreted carefully using expert judgement.

Studies on the prevalence of various types of misconduct have generally relied on surveys of self reported misconduct (Fanelli 2009; Marusic 2011), which are likely to be prone to serious biases. In most countries, institutions and funders rarely or never report on the number of cases of suspected misconduct that have been investigated or upheld, so it is impossible to make an accurate estimate of the 'usual' number of cases per thousand researchers per year. Greater transparency would facilitate the interpretation of studies designed to measure the effectiveness of interventions to reduce misconduct. However, given that serious misconduct appears to be rare (so far as we can tell), it is likely that only large, robust studies over a reasonable time would be sensitive to changes in prevalence. For example, a systematic follow-up of researchers undergoing federally required training in responsible conduct of research in the USA may provide better evidence for the effectiveness of training on research integrity and establish possible causative relationships.

ACKNOWLEDGEMENTS

We thank Prof. Melissa S Anderson from the University of Minnesota, Minneapolis, USA, for her contribution to the development of the protocol for this systematic review.

REFERENCES

References to studies included in this review

Aggarwal 2011 {published data only}

Aggarwal R, Gupte N, Kass N, Taylor H, Ali J, Bhan A, et al. A comparison of online versus on-site training in health research methodology: a randomized study. *BMC Medical Education* 2011;**11**:37.

Arnott 2008 {published data only}

Arnott E, Hastings P, Allbritton D. Research Methods Tutor: evaluation of a dialogue-based tutoring system in the classroom. *Behavior Research Methods* 2008;**40**:694-8.

Ballard 2013 {published data only}

Ballard IB. The impact of an academic integrity module and Turnitin on similarity index scores of undergraduate student papers. *Research in the Schools* 2013;**20**:1-13.

Barry 2006 {published data only}

Barry ES. Can paraphrasing practice help students define plagiarism. *College Student Journal* 2006;**40**:377-84.

Belter 2009 {published data only}

Belter RW, du Pre A. A strategy to reduce plagiarism in an undergraduate course. *Teaching of Psychology* 2009;**36**:257-61.

Bilić-Zulle 2008 {published data only}

Bilic Zulle L, Azman J, Frkovic V, Petrovecki M. Is there an effective approach to deterring students from plagiarizing?. *Science and Engineering Ethics* 2008;**14**:139-47.

Bilić-Zulle L, Frković V, Turk T, Ažman J, Petrovečki M. Prevalence of plagiarism among medical students. *Croatian Medical Journal* 2005;**46**:126-31.

Brown 2001 {published data only}

Brown VJ, Howell ME. The efficacy of policy statements on plagiarism: do they change students' views?. *Research in Higher Education* 2001;**42**:103-18.

Chao 2009 {published data only}

Chao C-A, Wilhelm WJ, Neureuther BD. A study of electronic detection and pedagogical approaches for reducing plagiarism. *The Delta Pi Epsilon Journal* 2009;**51**:31-42.

Chertok 2014 {published data only}

Chertok IRA, Barnes ER, Gilleland D. Academic integrity in the online learning environment for health science students. *Nurse Education Today* 2014;**34**:1324-9.

Clarkeburn 2002 {published data only}

Clarkeburn H, Downie JR, Matthew B. Impact of an ethics programme in a life sciences curriculum. *Teaching in Higher Education* 2002;**7**:65-79.

Compton 2008 {published data only}

Compton J, Pfau M. Inoculating against pro-plagiarism justifications: rational and affective strategies. *Journal of Applied Communication Research* 2008;**36**:98-119.

Dee 2012 {published data only}

Dee TS, Jacob BA. Rational ignorance in education: a field experiment in student plagiarism. *Journal of Human Resources* 2012;**47**:397-434.

Estow 2011 {published data only}

Estow S, Lawrence EK, Adams KA. Practice makes perfect: improving students' skills in understanding and avoiding plagiarism with a themed methods course. *Teaching of Psychology* 2011;**38**:255-9.

Fisher 1997 {published data only}

Fisher CB, Kuther TL. Integrating research ethics into the introductory psychology course curriculum. *Teaching of Psychology* 1997;**24**:172-5.

Hull 1994 {published data only}

Hull R, Wurm-Schaar M, James-Valutis M, Trigg RD. The effect of a research ethics course on graduate student's moral reasoning. http://www.richard-t-hull.com/publications/effect_research_ethics.pdf 1994.

Ivaniš 2008 {published data only}

Ivanis A, Hren D, Sambunjak D, Marusic M, Marusic A. A quantification of authors' contributions and eligibility for authorship: randomized study in a general medical journal. *Journal of General Internal Medicine* 2008;**23**:1303-10.

Kose 2011 {published data only}

Köse Ö, Arikan A. Reducing plagiarism by using online software: an experimental study. *Contemporary Online Language Education Journal* 2011;**1**:122-9.

Landau 2002 {published data only}

Landau JS, Druen PB, Arcuri JA. Methods for helping students avoid plagiarism. *Teaching of Psychology* 2002;**29**:112-5.

Marshall 2011 {published data only}

Marshall T, Taylor B, Hothersall E, Perez-Martin L. Plagiarism: a case study of quality improvement in a taught postgraduate programme. *Medical Teacher* 2011;**33**:e375-81.

Marušić 2006 {published data only}

Marušić A, Bates T, Anić A, Marušić M. How the structure of contribution disclosure statements affects validity of authorship: a randomized study in a general medical journal. *Current Medical Research and Opinion* 2006;**22**:1035-44.

May 2013 {published data only}

May DR, Luth MT. The effectiveness of ethics education: a quasi-experimental field study. *Science and Engineering Ethics* 2013;**19**:545-68.

Moniz 2008 {published data only}

Moniz R, Fine J, Bliss L. The effectiveness of direct-instruction and student-centered teaching methods on students' functional understanding of plagiarism. *College & Undergraduate Libraries* 2008;**15**:255-79.

Newton 2014 {published data only}

Newton FJ, Wright JD, Newton JD. Skills training to avoid inadvertent plagiarism: results from a randomized study. *Higher Education and Research* 2014;**33**:1180-93.

Risquez 2011 {published data only}

Risquez A, O'Dwyer M, Ledwith A. Technology enhanced learning and plagiarism in entrepreneurship education. *Education and Training* 2011;**53**:750-61.

Roberts 2007 {published data only}

Roberts LW, Warner TD, Dunn LB, Brody JL, Hammond KAG, Roberts BB. Shaping medical students' attitudes toward ethically important aspects of clinical research: results of a randomized, controlled educational intervention. *Ethics and Behavior* 2007;**17**:19-50.

Roberts LW, Warner TD, Hammond KAG, Brody JL, Kaminsky A, Roberts BB. Teaching medical students to discern ethical problems in human clinical research studies. *Academic Medicine* 2005;**80**:925-30.

Rolfe 2011 {published data only}

Rolfe V. Can Turnitin be used to provide instant formative feedback. *British Journal of Educational Technology* 2011;**42**:701-10.

Rose 1998 {published data only}

Rose MR, Fischer K. Do authorship policies impact students' judgments of perceived wrongdoings?. *Ethics and Behavior* 1988;**8**:59-79.

Schuetze 2004 {published data only}

Schuetze P. Evaluation of a brief homework assignment designed to reduce citation problems. *Teaching of Psychology* 2004;**31**:257-9.

Strohmetz 1992 {published data only}

Strohmetz DB, Skleder AA. The use of role-play in teaching research ethics: a validation study. *Teaching of Psychology* 1992;**19**:106-8.

Walker 2008 {published data only}

Walker AL. Preventing unintentional plagiarism: a method for strengthening paraphrasing skills. *Journal of Instructional Psychology* 2008;**35**:387-95.

Youmans 2011 {published data only}

Youmans RJ. Does the adoption of plagiarism-detection software in higher education reduce plagiarism?. *Studies in Higher Education* 2011;**36**:749-61.

References to studies excluded from this review
Ali 2014 {published data only}

Ali J, Kass NE, Sewakambo NK, White T, Hyder AA. Evaluating international research ethics capacity development: an empirical approach. *Journal of Empirical Research on Human Research Ethics* 2014;**9**:41-51.

Bagdasarov 2013 {published data only}

Bagdasarov Z, Thiel CE, Johnson JF, Connelly S, Harkrider LN, Devenport LD, et al. Case-based ethics instruction: the influence of contextual and individual factors in case content on ethical decision-making. *Science and Engineering Ethics* 2013;**19**:1305-22.

Goldie 2001 {published data only}

Goldie J, Schwartz L, McConnachie A, Morrison J. Impact of a new course on students' potential behaviour on encountering ethical dilemmas. *Medical Education* 2001;**35**:295-302.

Goldie J, Schwartz L, McConnachie A, Morrison J. The impact of three years' ethics teaching, in an integrated medical curriculum, on students' proposed behaviour on meeting ethical dilemmas. *Medical Education* 2002;**36**:489-97.

Gurung 2012 {published data only}

Gurung RAR, Wilhelm TM, Filz T. Optimizing honor codes for online exam administration. *Ethics and Behavior* 2012;**22**:158-62.

Harkrider 2012 {published and unpublished data}

Harkrider LN, Thie CE, Bagdasarov Z, Mumford MD, Johnson JF, Connelly S, et al. Improving case-based ethics training with codes of conduct and forecasting content. *Ethics and Behavior* 2012;**22**:258-80.

Hren 2007 {published and unpublished data}

Hren D, Sambunjak D, Ivaniš A, Marušić M, Marušić A. Perceptions of authorship criteria: effects of student instruction and scientific experience. *Journal of Medical Ethics* 2007;**33**:428-32.

Kligyte 2008 {published data only}

Kligyte V, Marcy RT, Waples EP, Sevier ST, Godfrey ES, Mumford MD, et al. Application of a sensemaking approach to ethics training in the physical sciences and engineering. *Science and Engineering Ethics* 2008;**14**:251-78.

May 2014 {published data only}

May DR, Luth MT, Schwoerer CE. The influence of business ethics education on moral efficacy, moral meaningfulness, and moral courage: a quasi-experimental study. *Journal of Business Ethics* 2014;**124**:67-80.

McDonalds 2010 {published data only}

McDonalds RJ, Neff KL, Rethlefsen ML, Kallmes DF. Effects of author contribution disclosures and numeric limitations on authorship trends. *Mayo Clinic Proceedings* 2010;**85**:920-7.

Mumford 2008 {published data only}

Mumford MD, Connelly S, Brown RP, Murphy ST, Hill JH, Antes AL, et al. A sensemaking approach to ethics training for scientists: preliminary evidence of training effectiveness. *Ethics and Behavior* 2008;**18**:315-39.

Pollock 1995 {published data only}

Pollock RE, Curley SA, Lotzova E. A short course in research ethics for trainees. *Academic Medicine* 1994;**69**:213-4.

Pollock RE, Curley SA, Lotzova E. Ethics of research training for NIH T32 surgical investigators. *Journal of Surgical Research* 1995;**58**:247-51.

Powell 2007 {published data only}

Powell ST, Allison MA, Kalichman MW. Effectiveness of a responsible conduct of research course: a preliminary study. *Science and Engineering Ethics* 2007;**13**:249-64.

Vallero 2007 {published data only}

Vallero DA. Beyond responsible conduct in research: new pedagogies to address macroethics of nanobiotechnologies. *Journal of Long-Term Effects of Medical Implants* 2007;**17**:1-12.

Additional references

Ajzen 2005

Ajzen I, Fishbein M. The influence of attitudes on behavior. In: Albarracin D, Johnson BT, Zanna MP editor(s). *The Handbook of Attitudes*. Mahwah: Lawrence Erlbaum Associates, 2005.

Anderson 2007

Anderson MS, Horn AS, Risbey KR, Ronning EA, De Vries R, Martinson BC. What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a National Survey of NIH-funded scientists. *Academic Medicine* 2007;**82**:853-60.

Angel 2004

Angel M. *The Truth About Drug Companies: How They Deceive Us and What to Do About It*. New York: Random House, 2004.

Antes 2009

Antes AL, Murphy ST, Waples EP, Mumford MD, Brown RP, Connelly S, et al. A meta-analysis of ethics instruction effectiveness in the sciences. *Ethics and Behavior* 2009;**19**:379-402.

Armitage 2001

Armitage CJ, Conner M. Efficacy of the theory of planned behaviour: a meta-analytic review. *British Journal of Social Psychology* 2001;**40**:471-99.

Barr 2000

Barr H, Freeth D, Hammick M, Koppel I, Reeves S. *Evaluations of Interprofessional Education: A United Kingdom Review of Health and Social Care*. London: Centre for the Advancement of Interprofessional Education (CAIPE), British Educational Research Association, 2000.

Boyd 2004

Boyd EA, Lipton S, Bero LA. Implementation of financial disclosure policies to manage conflicts of interest. *Health Affairs (Milwood)* 2004;**23**:206-14.

Burnham 2006

Burnham JF. Scopus database: a review. *Biomedical Digital Libraries* 2006;**3**:1.

Cogan 1953

Cogan ML. Towards a definition of a profession. *Harvard Educational Review* 1953;**23**:33-50.

Fanelli 2009

Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS One* 2009;**4**:e5738.

Funk 2007

Funk CL, Barrett KA, Macrina FL. Authorship and publication practices: evaluation of the effect of responsible conduct of research instruction to postdoctoral trainees. *Accountability in Research* 2007;**14**:269-305.

Grade Working Group 2004

Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. GRADE Working Group. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. *BMC Health Services Research* 2004;**4**:38.

Headrick 2010

Headrick TC. *Statistical Simulation: Power Method Polynomials and Other Transformations*. Boca Raton, FL, USA: Chapman & Hall/CRC, 2010:137.

Higgins 2011

Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Horsley 2011

Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database of Systematic Reviews* 2011, Issue 8. [DOI: [10.1002/14651858.MR000026.pub2](https://doi.org/10.1002/14651858.MR000026.pub2)]

Huwaldt 2014 [Computer program]

Huwaldt JA. Plot Digitizer 2.6.6. Sourceforge.net, 2014.

ICAI 2013

International Center for Academic Integrity, Fishman T (editor). *The Fundamental Values of Academic Integrity*. International Center for Academic Integrity, 2013.

ICMJE 2014

International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. <http://icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html> December 2014.

Institute of Medicine 2002

Committee on Assessing Integrity in Research Environments. *Integrity in Scientific Research*. Washington DC: Institute of Medicine, National Research Council, 2002.

Kalichman 2014

Kalichman M, Sweet M, Plemmons D. Standards of scientific conduct: are there any?. *Science and Engineering Ethics* 2014;**20**:885-96.

Kirkpatrick 1967

Kirkpatrick DL. Evaluation of training. In: Craig R, Bittel L editor(s). *Training and Development Handbook*. New York: McGraw Hill, 1967.

Komic 2015

Komic D, Marusic SL, Marusic A. Research integrity and research ethics in professional codes of ethics: survey of terminology used by professional organizations across research disciplines. *PLoS One* 2015;**10**:e0133662.

Marusic 2011

Marusic A, Bosnjak L, Jeroncic A. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. *PLoS One* 2011;**6**:e25258.

Marusic 2013

Marusic A, Wager E, Utrobicic A, Anderson M, Sambunjak D, Rothstein HR. Interventions to prevent misconduct and promote integrity in research and publication. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: [10.1002/14651858.MR000038](https://doi.org/10.1002/14651858.MR000038)]

Mhaskar 2015

Mhaskar R, Pathak EB, Wieten S, Guterbock TM, Kumar A, Djulbegovic B. Those responsible for approving research studies have poor knowledge of research study design: a knowledge assessment of Institutional Review Board members. *Acta Informatica Medica* 2015;**23**:196-201.

Michalek 2010

Michalek AM, Hutson AD, Wicher CP, Trump DL. The cost and underappreciated consequences of research misconduct: a case study. *PLoS Medicine* 2010;**7**:e1000318.

Nylenna 1999

Nylenna M, Andersen D, Dahlquist G, Sarvas M, Aakvaag A. Handling of scientific dishonesty in the Nordic countries. National committees on scientific dishonesty in the Nordic countries. *Lancet* 1999;**354**:57-61.

Plemmons 2006

Plemmons DK, Brody SA, Kalichman MW. Student perceptions of the effectiveness of education in the responsible conduct of research. *Science and Engineering Ethics* 2006;**12**:571-82.

Plint 2006

Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Medical Journal of Australia* 2006;**4**:263-7.

Reich 2011

Reich ES. Cancer trial errors revealed. *Nature* 2011;**469**:139-40.

Resnik 2011

Resnik DB, Shamoo AE. The Singapore statement on research integrity. *Accountability in Research* 2011;**18**:71-5.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Roig 1997

Roig M. Can undergraduate students determine whether text has been plagiarized?. *The Psychological Record* 1997;**47**:113-22.

Steneck 2006

Steneck NH. Fostering integrity in research: definitions, current knowledge, and future directions. *Science and Engineering Ethics* 2006;**12**:53-74.

Sterne 2014

Sterne JAC, Higgins JPT, Reeves BC on behalf of the development group for ACROBAT-NRSI. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI), Version 1.0.0, 24 September 2014. Available from <http://www.riskofbias.info> (accessed 29 December 2015).

Thomson Reuters 2011

Thompson Reuters. EndNote X5. Thompson Reuters 2011.

Turner 2012

Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: [10.1002/14651858.MR000030.pub2](https://doi.org/10.1002/14651858.MR000030.pub2)]

Young 2011

Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: [10.1002/14651858.MR000027.pub2](https://doi.org/10.1002/14651858.MR000027.pub2)]

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Aggarwal 2011

Methods	Randomized controlled trial, with 2 arms: 1) onsite 3.5-day classroom training in Biostatistics + (a week later, from their own home or office settings) 3.5-week online training course in Research Ethics; 2) 3.5-week online training in Biostatistics + onsite 3.5-day training in Research Ethics
---------	--

Aggarwal 2011 (Continued)

Data	Participants: 75 Indian scientists invited via an Indian biomedical journal. Inclusion criteria: (i) a degree in medicine or a masters' degree in science, (ii) receipt of a graduate or postgraduate degree within the last 10 years, (iii) at least 1 year of experience in human health-related clinical or social science research, (iv) basic computer skills and availability of broadband internet access, (v) willingness to be randomized and to participate in the study, and (vi) willingness to undertake pre- and post-course evaluations. 60 volunteered, and 58 were finally randomized to 2 study arms.
Comparisons	<p>Intervention 1 (n = 29): research ethics course delivered online: 15 lectures with 8.75 hours of instruction, plus 5 1-hour interactive case discussions, requiring application of ethical analysis skills (1 dropout)</p> <p>Intervention 2 (n = 29): research ethics course delivered on-site: 15 lectures with 8.75 hours of instruction, plus 5 1-hour interactive case discussions, requiring application of ethical analysis skills</p>
Outcomes	<p>1. Knowledge of research ethics before the course, immediately after the course and 3 months after the course</p> <p>2. Gain in knowledge of research ethics 3 months after the course in comparison to baseline knowledge</p> <p>Secondary outcome: satisfaction with the course</p>
Sources of funding	In part by grants from the Fogarty International Center/USNIH (Grant#2D43 TW000010-22-AITRP) and the Division of AIDS/NIAID/USNIH (AIU01069497), and in part from funds provided through the NIH Office of AIDS Research
Setting	Onsite training in Lucknow, India. Study dates not reported.
Notes	The study also compared online and onsite training in Biostatistics, so that the online Research Ethics group received onsite training in Biostatistics and the onsite Research Ethics group received online training in Biostatistics. Only the results for Research Ethics were analysed in this review.

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Yes	Computer-based randomization procedure
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not described, probably no blinding
Blinding self reported outcomes?	No	Not described, probably no blinding
Incomplete outcome data?	Yes	Attrition small and explained Minor inconsistencies in the reported number of participants
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Outcome measurement instrument not validated. Course attendance rate not reported.

Arnott 2008

Methods	Controlled before-and-after study (non-equivalent control group design)
Data	Participants: 136 undergraduate psychology students attending research methodology course covering ethics, variables, reliability, validity and experimental design
Comparisons	Intervention (n = 83): research methodology tutor (RMT) - dialogue-based intelligent tutoring online computer system (ITS) that engages students in one-on-one dialogues about various topics in undergraduate psychology research methods, delivered either with the use of RMT animated agent or as text-only presentation on the screen Control (n = 53): standard teaching
Outcomes	Knowledge gain on a 106-item paper-and-pencil test on research methodology (including research ethics)
Sources of funding	Not stated
Setting	DePaul University, Chicago, Illinois, USA. Study dates: 2006 winter semester to 2007 spring semester.
Notes	Results for knowledge in research ethics were not separately presented. Only the primary outcome was analysed in this review.

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Unclear	More attrition in the study group than in the control group, reasons for attrition not explained
Selective outcome reporting?	No	Only results of statistical analyses reported, without providing scores by study arm
Other bias?	No	Not all courses were taught by the same instructor; course attendance rate not reported; outcome measurement instrument not well described and not validated; the control group was selected partly in the way that it favoured students who could not install the computer program.
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups; groups belonged to different student classes: 5 classes in winter (3 intervention, 2 control), 1 class in spring (1 control, additionally collected data)
Confounding factors?	No	Only basic regression analysis, without controlling for confounding factors

Ballard 2013

Methods	Cluster-randomized trial with 4 arms testing academic integrity module and course writing assignment check by Turnitin
Data	Participants: 96 undergraduate students of the same course, majoring in education
Comparisons	Intervention 1 (n = 28): academic integrity module + submission of assignment through Turnitin Intervention 2 (n = 17): academic integrity module + no submission of assignment through Turnitin

Ballard 2013 (Continued)

Intervention 3 (n = 24): no academic integrity module + submission of assignment through Turnitin
 Control (n = 27): no academic integrity module + no submission of assignment through Turnitin

Outcomes	Plagiarism rate as determined by similarity score from Turnitin (0% to 100%)
Sources of funding	Not stated
Setting	Southeastern university in the USA. Study date: 2002 fall semester.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	Randomization method not described
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not reported
Blinding self reported outcomes?	Yes	Students were not aware that they were a part of the study
Incomplete outcome data?	Unclear	Attrition not specified per group; 96 out of 109 students completed the study
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Course attendance rate not reported; possibility of contamination among the groups

Barry 2006

Methods	Non-equivalent control group study design, with control group tested only after intervention
Data	Participants: 68 freshman-level (1st-year) psychology students enrolled into introductory lifespan development course
Comparisons	Intervention (n = 35): over 6 weeks, students completed a series of weekly assignments in which they paraphrased quotes from important psychology publications. Each assignment consisted of one paragraph-long quote, containing approximately 100 to 125 words. Students paraphrased each quote in 2 different ways (to practice paraphrasing techniques) and provided a citation after each (to reinforce the notion of correct citation style); students then provided the reference for the source in APA style. Control (n = 33): no intervention during a 6-week course
Outcomes	Score on plagiarism definition test (total 6 points): (a) Recognition of plagiarism as a form of academic dishonesty (including "cheating" or "stealing"), 1 point; (b) understanding that it is important to give credit to the author/source, 1 point; (c) knowing that plagiarism is when you claim someone else's work as your own, 1 point; with (d) "work" described as someone else's words, 1 point; (e) someone else's ideas, 1 point; and/or (f) someone else wrote the paper (internet, friend, etc.), 1 point

Barry 2006 (Continued)

Sources of funding	Not stated
Setting	The Pennsylvania State University, Fayette, Pennsylvania, USA. Study dates not reported.
Notes	Intervention group was tested before and after the course, and the control group (students from another section of the course) was tested only after the course.

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	No	Attrition in the study group 7/35, reasons not explained
Selective outcome reporting?	No	Only results of statistical analyses reported, without providing scores by study arm
Other bias?	No	Course attendance rate not reported; outcome measurement instrument not validated. Not all participants had both pre-test and post-test.
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors; gender, ethnicity, age not reported

Belter 2009

Methods	Non-equivalent control group study design with a historical control
Data	Participants: 266 undergraduate students in the first author's abnormal psychology courses in spring 2004 and 2005 semesters and fall 2005 and 2006 semesters. The Spring 2004 class served as the control group, and the remaining classes served as the experimental group.
Comparisons	Intervention (n = 66): research integrity module within the course - an online self instructional module made up of 4 sections: (a) Plagiarism Defined and Strategies to Avoid It, (b) Cheating Defined and Strategies to Avoid It, (c) Penalties for Academic Misconduct, and (d) Academic Integrity Evaluation. Examples of paraphrasing, 11-question quiz that had to be retaken until all correct answers Control (n = 200): no module during the same course
Outcomes	Number of plagiarized course paper assignments, as measured by a combined use of commercial plagiarism detection service (Turnitin) and a Google Internet search of passages of having been plagiarized: "We only considered a paper to have been plagiarized if it included one or more passages that was word-for-word the same as another source without appropriate citation and quotation marks."
Sources of funding	Not stated
Setting	University of West Florida, Pensacola, Florida, USA. Study dates: 2004 and 2005 spring semesters and 2005 and 2006 fall semesters.
Notes	This study had a historical control group and both groups were tested after the courses, via course assignments

Belter 2009 (Continued)

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	Attrition very small, explained
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	Not clear if the courses were taught by the same instructors Outcome measurement instrument not validated. No pre-test, only post-test.
Blinding?	No	No blinding
Comparability of groups?	No	Only basic demographic comparisons (gender, race) compared between the groups, without actual numbers reported. Comparability questionable: groups belonged to different student classes - the first course (Spring 2004) served as control, and 3 subsequent courses (Fall 2004, Spring and Fall 2005) were intervention groups.
Confounding factors?	No	No regression analysis; no attempt to control for confounding factors

Bilić-Zulle 2008

Methods	Non-equivalent control group study design with a historical control
Data	Participants: 290 medical students attending second-year medical informatics courses in 2001/2002, 2002/2003 and 2004/2005 academic years
Comparisons	Intervention 1 (n = 87, 2002/2003 course): specific oral warning that plagiarism is forbidden Intervention 2 (n = 92, 2004/2005 course): specific oral warning that student essays will be examined by plagiarism detection software and that plagiarism will be penalized Control (n = 111, 2001/2002 course): no intervention
Outcomes	Proportion of plagiarized text in an essay based on a published scientific article (1 of 4 articles written in Croatian, 2 available only as printed copies and 2 in electronic format, posted at the School's website)
Sources of funding	Research project "Prevalence, features and attitudes toward plagiarism in biomedical sciences" supported by the Ministry of Science, Education and Sports of the Republic of Croatia, grant No. 0062044
Setting	University of Rijeka School of Medicine, Croatia. Study dates: 2001/2002, 2002/2003 and 2004/2005 academic years.
Notes	This study had a historical control group and all groups were tested after the courses, via course assignments. The results of the study were published in 2 articles, in 2005 and 2008.

Risk of bias

Bilić-Zulle 2008 (Continued)

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	No attrition
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	Outcome measurement instrument not validated, detecting plagiarism from the source articles only; no pre-test, only post-test; historical control
Blinding?	No	No blinding
Comparability of groups?	No	No comparison of demographic characteristics between the intervention groups; the groups were composed of students from different academic years
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors; age, gender stated but confounding factors not taken into consideration in analysis

Brown 2001

Methods	Quasi-randomized trial with 3 arms. Participants were assigned as they entered the classroom to interventions by being given a booklet from the top of a randomized pile so that the conditions were not in order and any particular student was as likely to receive a booklet from any condition.
Data	Participants: 218 undergraduate psychology students (11 incomplete responses)
Comparisons	<p>Intervention 1 (n = 51): educational statement on plagiarism - the passage (270 words) was designed to educate the student about the problem of plagiarism and how it can be avoided. The passage gave an extensive description and an example of the correct way to cite material. The tone of the passage was serious, referring to "scholarship" and "etiquette".</p> <p>Intervention 2 (n = 54): warning statement on plagiarism - the passage (137 words) contained a more limited definition of plagiarism, and there was no description of the correct way to cite material. The warning went on to refer to the frequency of plagiarism as being low. The tone of the passage was less serious, referring to plagiarism as "misbehaviour" and a "stupid" risk.</p> <p>Control (n = 102): "No information" condition - the student was given no prior instruction about plagiarism and proceeded to complete the questionnaire.</p>
Outcomes	<p>Ratings (0 to 100 continuous line scale) of students' views about examples of verbatim (identical copy) or closely paraphrased text without citations in comparison to the original text for the following outcomes:</p> <ol style="list-style-type: none"> 1. Seriousness of this behaviour as viewed by the respondent 2. Seriousness of the behaviour as it would be viewed by staff 3. Perceived frequency of similar behaviour by their colleagues 4. Perception of necessity to include citations when writing a text 5. Perception of whether their colleagues would cite sources in their written works
Sources of funding	Not stated
Setting	School of Psychology, University of St. Andrews, Scotland, UK. Study date: 2002

Brown 2001 (Continued)

Notes All 5 outcomes were analysed in this systematic review

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	Participants were assigned as they entered the classroom to interventions by being given a booklet from the top of a randomized pile so that the conditions were in random order and any particular student was as likely to receive a booklet from any condition.
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Yes	Experimenter not aware of the condition allocated to each student
Blinding self reported outcomes?	Yes	Participants not aware that there were different conditions in the booklets
Incomplete outcome data?	No	11 of 218 questionnaires were incomplete; breakdown by study groups not provided
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Outcome measurement instrument not validated

Chao 2009

Methods	Non-equivalent control group study design, with testing only after intervention
Data	Participants: 116 undergraduate business students enrolled in online or on-site courses
Comparisons	<p>Intervention 1 (n = 42): the instructional treatment for students included: 1) receiving instruction during the class and using reference and citation resources available at the Purdue University Online Writing Lab Website, 2) instructed to read the academic integrity policy in the University Code of Student Conduct Handbook as a homework assignment, and given a 10-question true-false quiz in the reading assignment, 3) brief discussion led by an instructor on what constitutes plagiarism, how to avoid it and proper steps for paraphrasing which included examples of properly and improperly paraphrased paragraphs, 4) originality report of an anonymous previous student's written submission shown and discussed during class as a negative example. Students were not required to purchase the APA Publication Manual for their course work.</p> <p>Intervention 2 (n = 41): the instructional treatment for students included: 1) review of the academic integrity policy in the University Code of Student Conduct Handbook and a quiz based on the policy; 2) extended in-class discussion on what constitutes plagiarism, how to avoid it, and proper paraphrasing techniques with examples; and 3) practice exercise in which students paraphrased 2 paragraphs. The exercise was graded by the instructor and returned to the students along with detailed feedback. Further, to orient students to the power of Turnitin plagiarism detection software, an originality report of an anonymous previous student's written submission was also shown and discussed during class.</p> <p>Control (n = 33): minimal instruction about avoiding plagiarism. The extent of the instruction consisted of an assignment to read the Code of Student Conduct for the university as it applied to plagiarism and cheating and to send an e-mail to the instructor identifying several types of plagiaristic behaviour identified in the Code of Student Conduct. This assignment came at the beginning of the course. Sever-</p>

Chao 2009 (Continued)

al weeks before a writing assignment ("the capstone report") was due, students were reminded not to plagiarize. Students had to purchase the APA Publication Manual and read assigned sections.

Outcomes	1. Number of students whose report ("capstone writing assignment") contained plagiarized text 2. Percentage of the text plagiarized
Sources of funding	Not stated
Setting	On-site courses: Indiana State University, Terre Haute, Indiana, USA. Study dates not reported.
Notes	All groups were tested after the intervention. Students from the control group were from an online course and the intervention groups were from on-site courses.

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	No attrition
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	Possible different treatment bias - control group attended online courses and the intervention group attended on-campus courses; not all courses were taught by the same instructor. Measuring instrument not validated.
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups
Confounding factors?	No	Only basic regression analysis, without controlling for confounding factors. Gender, age, ethnicity not reported.

Chertok 2014

Methods	Controlled before-and-after study
Data	Participants: 355 undergraduate health sciences (nursing, pre-medical, exercise physiology) students registered for health science courses
Comparisons	Intervention (n = 194): face-to-face education intervention about academic integrity (in addition to the course syllabus) at the first course class, including instruction and examples on online academic integrity and plagiarism Control (n = 161): no intervention (review of the course syllabus)
Outcomes	1. Attitude score on 17 questions (from 0 - most appropriate ethical attitude to 51 - most inappropriate ethical attitude) 2. Knowledge score on 13 true/false questions
Sources of funding	Not stated
Setting	West Virginia University, Morgantown, West Virginia, USA. Study dates not reported.

Chertok 2014 (Continued)

Notes —

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	All outcomes reported
Selective outcome reporting?	Yes	Some inaccuracies in data presentation, not affecting the results or conclusion
Other bias?	Yes	None identified
Blinding?	No	No blinding
Comparability of groups?	Unclear	Demographic characteristics reported only for the whole sample of participants; no comparison between the groups
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors

Clarkeburn 2002

Methods	Controlled before-and-after study
Data	Participants: 517 students attending honours courses in sciences
Comparisons	Intervention (n = 276): 3 structured 2-hour, face-to-face ethics discussions with students, covering scientific misconduct and integrity, during science courses Control (n = 241): no intervention
Outcomes	1. TESS – Test of Ethical Sensitivity in Science (students are asked to raise no more than 5 issues/questions they believe should be considered before making a research decision on the given scenario (which relates to genetic modification for the production of pharmaceutical milk) and these are rated according to the ethical elements recognised) 2. DIT – Defining Issues Test (test based on ratings and rankings of stage prototypical statements representing elements in moral decision-making in reference to a specific moral dilemma; scored using P-scores and Type-scores) 3. Perry Questionnaire (designed to assess students' meta-ethical understanding; Perry's scheme of ethical development can be divided into 3 different Types: A: basic dualism; B: confusion in relativism; C: commitment)
Sources of funding	Research was supported by the European Commission Marie Curie funding
Setting	University of Glasgow Institute of Biomedical and Life Sciences, Scotland, UK. Study dates not reported.
Notes	—

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Clarkeburn 2002 (Continued)

Incomplete outcome data?	No	Attrition not clearly reported, probably high
Selective outcome reporting?	Unclear	Some outcomes only partially reported; no breakdown of results by course type
Other bias?	No	Possible different treatment bias - 2 approaches to teaching analysed together; attendance reportedly less than perfect
Blinding?	No	No blinding
Comparability of groups?	No	Only basic demographic comparisons (honours course type) compared between the groups; differences between the groups very probable
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. Only gender stated, age and race not reported or taken into account in analysis.

Compton 2008

Methods	<p>Randomized controlled trial with 4 arms testing different inoculation messages against plagiarism. During Phase 1, participants completed questionnaires that gathered demographic information and assessed initial attitudes towards plagiarism. During Phase 2, occurring approximately 2 weeks after Phase 1, participants were presented with the inoculation treatment message and then completed scales designed to assess attitudes towards plagiarism. Participants in the control condition only completed the scales. Finally, during Phase 3, occurring approximately 2 weeks after Phase 2, all participants received an attack message, which attempted to justify plagiarism using either affective or rational justification. All participants then completed scales assessing attitudes towards plagiarism (in phase 2) and attitude toward position in attack message (in phase 3).</p>
Data	<p>Participants: 225 students from communication courses at a Midwestern university in the USA. Eligibility criteria: 1) at least 18 years of age, 2) signed informed consent.</p> <p>Participants were randomly assigned to 1 of the 3 treatment conditions and the 2 attack message conditions.</p>
Comparisons	<p>Intervention 1 (n = 64): fear-based inoculation treatment message: focused on the likelihood that students who plagiarize would be caught</p> <p>Intervention 2 (n = 66): guilt-based inoculation treatment message: used an anticipated guilt strategy, describing students' feelings of guilt after cheating and the hurt plagiarism causes professors</p> <p>Intervention 3 (n = 57): rational inoculation treatment message: used evidence and rational arguments and avoided affect-laden language</p> <p>Control (n = 59): no intervention</p>
Outcomes	<p>1. Attitudes towards plagiarism, assessed by the following measures (7 point Likert scale, from 0 = negative to 6 = positive):</p> <ul style="list-style-type: none"> a) Threat generated by inoculation messages (realization that one may encounter a message designed to change existing attitude) b) Attitude towards plagiarism c) Importance or salience of the topic of plagiarism (involvement in the issue) d) How often plagiarism is discussed (accessibility of the issue)

Compton 2008 (Continued)

- e) Vested interest in plagiarism (stake in the issue, thinking of the issue, being affected by the issue in present or future, personal ability to handle issue)
2. Attitudes towards justification of plagiarism (attack messages in phase 3), assessed by the following measures (7-point Likert scale, from 0 = more resistance to 6 = less resistance towards attack message):
- a) Attitude towards the justification of plagiarism
- b) Perceived credibility of the source of messages justifying plagiarism in relation to its character, competence and sociability

Sources of funding	Not stated
Setting	Described as a "midwestern university in the USA." Study dates not reported.
Notes	Study date

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The study just states that participants were randomly assigned.
Allocation concealment?	Unclear	Not described.
Blinding researcher-assessed outcomes?	Unclear	Not described
Blinding self reported outcomes?	Unclear	Not reported
Incomplete outcome data?	Unclear	Attrition not reported
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Yes	None identified

Dee 2012

Methods	Cluster-randomized trial; 28 courses were randomly assigned to treatment and control conditions
Data	Participants: 697 students from undergraduate social-science and humanities courses and their 1329 course written assignments
Comparisons	Intervention (n = not stated): blackboard (online) based tutorial on understanding and avoiding plagiarism (18 tutorial screens) Control (n = not stated): no intervention, students submitted the course paper via an online blackboard system
Outcomes	1. Course papers identified as having plagiarism content with similarity scores of 15 and higher or 8 and higher

Dee 2012 (Continued)

2. Score on scales for 6 questions about students' attitudes and perceptions about plagiarism (on a scale from 1 = strong disagreement to 5 = strong agreement)

Secondary outcome: satisfaction and experiences with the course (3 questions with scales from 1 = strong disagreement to 5 = strong agreement)

Sources of funding	Not stated
Setting	Described as a "single, highly selective postsecondary institution in the United States." Study date: 2007 fall semester.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	Block randomization, with matching participating courses on baseline traits
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	No	Course instructors were aware of the interventions in the study
Blinding self reported outcomes?	Yes	Students were not aware of the intervention
Incomplete outcome data?	Yes	Attrition small, explained
Selective outcome reporting?	No	Only results of statistical analyses reported, without reporting absolute numbers of plagiarized papers
Other bias?	Yes	None identified

Estow 2011

Methods	Controlled before-and-after study
Data	Participants: 62 undergraduate psychology major students enrolled in a research methods course
Comparisons	<p>Intervention (n = 27): Plagiarism-themed course: 1) the syllabus included a statement that all work should be done according to the principles in the college-wide academic honour code and referred students to the student handbook for further information, 2) discussed the importance of plagiarism avoidance in both conditions, 3) various study designs and statistics during the semester were covered where a variety of assignments were related to the plagiarism theme. The course lasted for one semester.</p> <p>Control (n = 17): No intervention, regular course: 1) the syllabus included a statement that all work should be done according to the principles in the college-wide academic honour code and referred students to the student handbook for further information, 2) discussed the importance of plagiarism avoidance.</p>
Outcomes	1. Number of sentences correctly identified as plagiarism on 2 homework assignments

Estow 2011 (Continued)

2. Quality of paraphrasing in homework assignments: 4-point scale, from 1 (direct copying of a significant portion of the original without quotation marks) to 4 (good paraphrasing)
3. Number of strategies to avoid plagiarism

Sources of funding	Authors' statement: "The authors received no financial support for the research, authorship, and/or publication of this article."
Setting	Described as a "small Southeastern liberal arts college"; USA. Study dates: 2007 and 2008 spring semesters.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	No	Considerable attrition, unexplained
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Outcome measurement instrument not validated; attendance rate not reported
Blinding?	No	No blinding
Comparability of groups?	Unclear	Only basic demographic comparison (grade point average). Students were similar, 2 classes were included from 2 years, not clear how they contributed to different groups.
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. Gender, age, ethnicity not reported

Fisher 1997

Methods	Controlled before-and-after study
Data	Participants: 585 undergraduate students enrolled in 24 introductory psychology sections taught during fall and spring college semester
Comparisons	<p>Intervention (n = not stated; "half of the students"): Ethics-enhanced instruction - 6 case study teaching modules based on a broad sample of "classic" empirical studies cited in a majority of introductory psychology textbooks. Students received a workbook that included (a) a brief abstract of each study; (b) a more detailed description of each experiment including the purpose of the study, primary hypothesis, participants, procedure, results and conclusions (Fisher & Fyrberg, 1994); and (c) homework assignments composed of 4 sets of focus questions requiring students to critically evaluate ethical issues derived from the Belmont Report (NCPHSBBR, 1978) and the APA Ethics Code (1992).</p> <p>Control (n = not stated; "half of the students"): No intervention - standard ethics instruction typically included a brief overview of informed consent requirements and the ethical issues associated with Milgram's (1963) use of deception in his classic obedience study.</p>
Outcomes	<ol style="list-style-type: none"> 1. Knowledge of specific ethics procedures (score on a 0 to 3 scale) 2. Ability to weigh scientific responsibility and participant rights and welfare (score on a 0 to 2 scale)

Fisher 1997 (Continued)

Secondary outcome: curriculum evaluation

Sources of funding	This research was supported by Grant SBR-9310458 from the National Science Foundation to Celia B. Fisher and Fordham University
--------------------	---

Setting	Fordham University, New York and Loyola University, Chicago, USA. Study dates not reported.
---------	---

Notes	—
-------	---

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	No	Considerable attrition, unexplained
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Outcome measurement instrument not validated; attendance rate not reported
Blinding?	No	No blinding
Comparability of groups?	Unclear	Only basic demographic comparison (grade point average). Students were similar, 2 classes were included from 2 years, not clear how they contributed to different groups.
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. Gender, age, ethnicity not reported.

Hull 1994

Methods	Controlled before-and-after study
---------	-----------------------------------

Data	Participants: 38 graduate students
------	------------------------------------

Comparisons	Intervention (n = 20): research ethics course that used combined lecture and discussion group format Control (n = 18): no intervention
-------------	---

Outcomes	Mean score on Gibbs' Sociomoral Reflection Objective Measure-Short Form (SROM-SF): the questionnaire consists of 2 moral dilemmas and 48 moral reasoning justifications
----------	---

Sources of funding	Not stated
--------------------	------------

Setting	University of New York at Buffalo, Buffalo, New York, USA. Study dates not reported.
---------	--

Notes	Participants in the intervention group were enrolled in health and biomedical sciences majors, and participants in the control group were enrolled in business and management majors.
-------	---

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Hull 1994 (Continued)

Incomplete outcome data?	Unclear	Attrition not clearly reported
Selective outcome reporting?	No	Means reported without standard deviations
Other bias?	Unclear	Course attendance rate not reported
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups, who differed in their majors
Confounding factors?	No	Only basic regression analysis, without controlling for confounding factors. Difference in age between groups (mean 38 versus 26), gender stated. Ethnicity and grade point average not stated.

Ivaniš 2008

Methods	Randomized controlled trial with 2 arms: 1) binary description (yes/no) of 12 possible authorship contributions and 2) ordinal description (scale from 0 - none to 4 - full) of 12 possible authorship contributions
Data	Data sources: 965 authors from 181 manuscript submitted to a general medical journal from January to July 2005
Comparisons	Intervention 1 (n = 87 manuscripts, 409 authors): binary rating of authorship contribution Intervention 2 (n = 94 manuscripts, 456 authors): ordinal rating of authorship contribution
Outcomes	Per cent of authors satisfying authorship criteria of the International Committee of Medical Journal Editors (ICMJE)
Sources of funding	Research grant from the Ministry of Science and Technology of the Republic of Croatia, No. 108-1080314-0140
Setting	Authors submitting work to <i>Croatian Medical Journal</i> , Croatia. Study dates: January to July 2005.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Yes	Random allocation using the method of randomly permuted blocks
Allocation concealment?	No	No allocation concealment
Blinding researcher-assessed outcomes?	No	Not stated, assessors were probably not blinded
Blinding self reported outcomes?	Yes	Participants were not aware of the intervention

Ivaniš 2008 (Continued)

Incomplete outcome data?	Yes	Attrition small, explained
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Yes	None identified

Kose 2011

Methods	Non-equivalent control group study design, with control group tested only after intervention
Data	Participants: 40 undergraduate university engineering students enrolled in academic writing course
Comparisons	<p>Intervention (n = 17): use of Turnitin in practising avoiding plagiarism in academic writing - no prior Turnitin experience, informed about the program, the originality reports and about how it works in detecting plagiarism in several lectures with PowerPoint assisted presentations; a folder "Self-Study" was created in one of these presentations and the students were informed that they could check their work's originality by submitting it to this folder and revise it if necessary before the final submission for grading; students also practised using Turnitin by submitting their essays.</p> <p>Control (n = 29): no intervention - no prior Turnitin experience, no information about using the program; not informed that their essays would be checked for originality using Turnitin.</p>
Outcomes	<p>Per cent plagiarized text in submitted course assignments</p> <p>Secondary outcome: views about experience of using Turnitin (only students in the intervention group)</p>
Sources of funding	Not stated
Setting	Middle East Technical University, Ankara, Turkey. Study dates: 2008 fall and 2009 spring semesters.
Notes	The intervention group was tested before and after intervention, control group only for post-intervention test

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Unclear	Attrition not clearly reported
Selective outcome reporting?	No	Results reported only in percentages, not in absolute numbers; ranges reported without medians; not clear if all questions from the survey have actually been addressed and reported
Other bias?	No	Outcome measurement instrument not validated, course attendance rate not reported, possible "contamination" between the groups. Only the experimental group has pre and post test, the control group only post test plagiarism check.
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups

Kose 2011 (Continued)

Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. No characteristics of the students reported.
----------------------	----	---

Landau 2002

Methods	Randomized controlled trial with 4 arms to assess whether teaching students about plagiarism by providing them with examples or providing them with feedback on a plagiarism detection exercise would increase their ability to identify plagiarism and avoid it when paraphrasing.	
Data	Participants: 94 undergraduate students attending freshman-level research course	
Comparisons	<p>Intervention 1 (n = not stated): feedback only - comparison of 2 texts to determine if plagiarized; for the first example, participants were informed via a PowerPoint presentation whether the passage had been plagiarized, and were then given a rationale for each answer (e.g. the passage was plagiarized because the words were not altered sufficiently)</p> <p>Intervention 2 (n = not stated): examples only - comparison of 2 texts to determine if plagiarized; for the first example, participants were presented a brief definition of plagiarism and 3 plagiarized examples (PowerPoint presentation about excerpts from a case in which one romance novelist stole several short passages from another writer)</p> <p>Intervention 3 (n = not stated): feedback/examples - comparison of 2 texts to determine if plagiarized; participants received both the feedback and the examples</p> <p>Control (n = not stated): no intervention - comparison of 2 texts to determine if plagiarized; no feedback of examples</p>	
Outcomes	<p>1. Score on Plagiarism Knowledge Survey (PKS)</p> <p>2. Plagiarism in students' exercises, measured as: a) mean number of words overlapping with original in students' paraphrasing exercise, b) mean number of plagiarized 2-word strings, c) mean number of plagiarized 3-word strings</p>	
Sources of funding	Not stated	
Setting	York College of Pennsylvania, York, Pennsylvania, USA. Study dates not reported.	
Notes	Randomization method not described	

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The study just states that participants were randomly assigned
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Yes	Rater was unaware of the aims of the study
Blinding self reported outcomes?	Unclear	Not described
Incomplete outcome data?	Unclear	Not clear as data were not reported in full (no numbers per group)

Landau 2002 (Continued)

Selective outcome reporting?	No	Means presented without standard deviations; non-significant results not reported; results for one of the outcomes presented only for all participants and not for each studied group
Other bias?	Yes	None identified

Marshall 2011

Methods	Non-equivalent control group study design with a historical control	
Data	Data sources: 1085 course assignments from students enrolled into a Masters in Public Health (MPH) programme (number of students not stated)	
Comparisons	<p>Intervention 1 (n = 279): warning about plagiarism involving 1) introductory lecture included an additional explanation of the need to avoid plagiarism; 2) students were specifically informed that the software was being used; 3) lecture later in the same academic year explained in some detail how the percentage text match was used to identify possible plagiarism; 4) all students were informed (without identifying individuals) when students had been penalized for plagiarism</p> <p>Intervention 2 (n = 515): training in avoiding plagiarism involving 1) a new 40-minute interactive plagiarism seminar was delivered during teaching time allocated to one of the compulsory MPH modules; 2) in the seminar, following a brief talk about plagiarism and Turnitin software, students were provided with an anonymised copy of a detailed Turnitin report from a previous student who had been penalized for plagiarism; students discussed the case in small groups and decided whether plagiarism had taken place; what (if any) action should be taken; what action should be taken if a similar Turnitin report was subsequently received for the same student, and how a student could avoid plagiarism; after feedback and discussion students were informed about the outcome of the case; 3) discussion on how to avoid plagiarism</p> <p>Control (n = 191): no intervention</p>	
Outcomes	1. Per cent text matching in assignments 2. Number of plagiarism occurrences	
Sources of funding	Not stated	
Setting	University of Birmingham, Birmingham, UK. Study dates: 2006-2009 academic years.	
Notes	The control group was historical - the first course in 2006, the course in 2007 received Intervention 1, courses in 2008-2009 received Intervention 2	

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Unclear	Attrition not clearly reported
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	Course attendance rate not reported; possible different treatment bias (changes in the course content over the years). Historical controls
Blinding?	No	No blinding

Marshall 2011 (Continued)

Comparability of groups?	No	No comparison of demographic characteristics between the groups; differences very probable (majors, year of enrolment)
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. No characteristics of the students reported.

Marušić 2006

Methods	Randomized controlled trial with 3 arms testing different formats of authorship contribution declaration
Data	Data sources: 1462 authors from 337 manuscript submitted to a general medical journal from January 2003 to June 2004
Comparisons	<p>Intervention 1 (n = 112 manuscripts, 523 authors): categorical form - the form listed 11 contribution categories and asked the respondents to choose what they contributed to the submitted manuscript; all contributions relevant to individual ICMJE criteria were included in the list</p> <p>Intervention 2 (n = 114 manuscripts, 471 authors): instructional form - the form grouped the contributions into ICMJE criteria and instructed the respondents how many contributions were needed to satisfy the authorship requirements</p> <p>Intervention 3 (n = 106 manuscripts, 468 authors): control - open-ended contribution declaration form (free-text answer to a question to describe in their own words their contributions to the submitted manuscript)</p>
Outcomes	Per cent of authors who did not satisfy authorship criteria of the International Committee of Medical Journal Editors (ICMJE)
Sources of funding	Research grant from the Ministry of Science and Technology of the Republic of Croatia, No. 0108182
Setting	Authors submitting work to Croatian Medical Journal, Croatia. Study dates: 2003 and the first half of 2004.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Yes	Random allocation using randomly permuted blocks
Allocation concealment?	No	No allocation concealment
Blinding researcher-assessed outcomes?	No	Study described as single-blinded, assessors were probably not blinded
Blinding self reported outcomes?	Yes	Participants were not aware of the intervention
Incomplete outcome data?	Yes	Attrition small, explained
Selective outcome reporting?	Yes	All outcomes adequately reported

Marušić 2006 (Continued)

Other bias?	Yes	None identified
-------------	-----	-----------------

May 2013

Methods	Controlled before-and-after study
Data	Participants: 442 students enrolled in graduate and undergraduate engineering, technology and science courses
Comparisons	<p>Intervention 1 (n = not stated): embedded ethics training - participants enrolled in a course where ethics was not the primary topic, but was addressed during the course</p> <p>Intervention 2 (n = not stated): stand-alone ethics training - participants enrolled in a stand-alone ethics course, where ethics content was the primary material for the course</p> <p>Control (n = not stated): no intervention - participants not enrolled in any ethics-based course</p>
Outcomes	<ol style="list-style-type: none"> 1. Mean score on N2 Index from the Defining Issues Test (DIT) of moral judgement 2. Number of correct answers on a knowledge test of responsible conduct of research 3. Mean score on a Perspective Taking test (6-item measure using 7-point scale from 1 - strongly disagree to 7 - strongly agree) 4. Mean score on a Moral Efficiency test (14-item measure using 5-point scale from 1 - not confident at all to 5 - very confident) 5. Mean score on a Moral Courage test (6-item measure using 7-point scale from 1 - strongly disagree to 7 - strongly agree) 6. Mean score on a Moral Meaningfulness test (4-item measure using 7-point scale from 1 - strongly disagree to 7 - strongly agree)
Sources of funding	Funded by the National Science Foundation (NSF) over a 3-year period (NSF Grant #0629443)
Setting	Described as "three different Midwestern universities, USA." Study dates not reported.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	No	High attrition due to complex survey instrument; analysed
Selective outcome reporting?	Yes	Outcomes adequately reported
Other bias?	No	Participation in study groups was possibly self selected; course attendance rate not reported; different treatment bias very probable, as no standardized course was delivered to experimental groups and intervention was provided by different instructors
Blinding?	Unclear	Participants were blind to the research hypothesis. No blinding described for the raters.

May 2013 (Continued)

Comparability of groups?	Unclear	Demographic characteristics reported only for the whole sample, without comparisons between the study groups. The student groups came from different universities, where they could have differences not only in the ethics training but also in other possible co-interventions. Researchers conducted a test for selection effect.
Confounding factors?	Yes	Characteristics of students reported: age, gender, level of education, race. A limited number of confounding factors controlled in the regression analysis: gender was found to have an effect.

Moniz 2008

Methods	Cluster-randomized controlled trial with 3 arms. Randomization included first the classes then the instructors to the classes.
Data	Participants: 289 fourth-year undergraduate university students enrolled in English Communications Skills course
Comparisons	<p>Intervention 1 (n = not stated): direct instruction on plagiarism - 1) lecture; 2) board-based instruction; 3) class work; 4) occasional student feedback</p> <p>Intervention 2 (n = not stated): student-centred approach - group and role-playing, with the instructor providing guidance, answering questions and helping students uncover information that they might have left out of their discussions and role-plays</p> <p>Control (n = not stated): standard teaching - PowerPoint lecture on plagiarism</p>
Outcomes	<ol style="list-style-type: none"> Score on the modified Plagiarism Knowledge Survey (PKS) Score on Theoretical Understanding of Plagiarism test Score on Knowledge of Content about plagiarism, paraphrasing and copyright <p>Secondary outcome: score on Fidelity of Treatment instrument to test participants' experiences with the activities during the course</p>
Sources of funding	Not stated
Setting	Johnson & Wales University, Charlotte Campus, Charlotte, North Carolina, USA. Study dates: 2005 to 2006.
Notes	The number of participants randomized to 3 arms was not stated, only the number of participants who completed pre- and post-intervention assessment.

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The article just states that classes and instructors were randomly assigned
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not described

Moniz 2008 (Continued)

Blinding self reported outcomes?	Unclear	Not described
Incomplete outcome data?	Yes	Attrition not addressed, probably none
Selective outcome reporting?	No	Numerical results not reported
Other bias?	Unclear	Only part of the outcome measurement instrument validated in previous research; unclear baseline comparability of units of randomization (courses)

Newton 2014

Methods	Randomized controlled trial with 2 arms to test training intervention for the prevention of plagiarism
Data	Participants: 260 newly enrolled students attending the first-year orientation programme for a business and economics undergraduate degree (52.7% response rate for final analysis of 137 participants)
Comparisons	Intervention (n = not stated): 30-minute procedural plagiarism training programme on paraphrasing, patch-writing and plagiarism, received before completing the survey Control (n = not stated): surveyed before intervention (30-minute procedural plagiarism training programme)
Outcomes	1. Score (on a scale from -3: strongly disagree to +3: strongly agree) of perception of: a) note-taking ability, b) confidence in paraphrasing source material; c) confidence in referencing sources in assignments 2. Knowledge score (average score from 0 to 5) about in-text referencing 3. Score (average score, marked by 2 assessors, score range not provided) on the quality of paraphrasing in a written assignment: a) referencing, b) patch-writing, c) plagiarism
Sources of funding	Monash University orientation special project grant
Setting	Monash University, Australia. Study dates not reported.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	Randomization method not described
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not reported
Blinding self reported outcomes?	Unclear	Not reported
Incomplete outcome data?	Yes	There was no attrition among students who accepted participation in the study

Newton 2014 (Continued)

Selective outcome reporting?	No	Number of participants in each study arm not reported; non-significant differences not reported
Other bias?	Unclear	Outcome measuring instrument not validated

Risquez 2011

Methods	Randomized controlled trial with 2 study arms to test a tutorial on plagiarism prevention	
Data	Participants: 434 second year undergraduate entrepreneurship students in a class setting	
Comparisons	Intervention (n = 77): in-class tutorial on plagiarism prevention as part of the course (computer-based 1-hour tutorial completed individually) Control (n = 211): no intervention, standard teaching	
Outcomes	1. Score on attitudes, ethical views and engagement in plagiarism (a series of scales from 1 = completely disagree to 5 = completely agree) 2. Score on perceived seriousness of breach of academic guidelines from their own and their lecturer's point of view (scales from 1 = no breach to 100 = extremely serious breach) 3. Reported engagement in plagiarism (a series of scales from 1 = never to 5 = always)	
Sources of funding	Not stated. The course was developed by a commercial spin-off company.	
Setting	University of Limerick, Limerick, Ireland. Study dates: spring semester of 2008/2009 academic year.	
Notes	The study also had a non-randomized part with voluntary non-randomized participation. Only the randomly assigned groups were included in this review. Baseline measures were not provided.	

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	No	Some intervention groups were randomly assigned, some non-randomly. The method of randomization is not described.
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not described
Blinding self reported outcomes?	Unclear	Not described
Incomplete outcome data?	Yes	No attrition
Selective outcome reporting?	No	Means reported without standard deviations
Other bias?	Unclear	Questionable comparability of study arms; outcome measurement instrument not validated

Roberts 2007

Methods	Randomized controlled trial with 3 arms to investigate different types of brief ethics training on medical students' attitudes toward ethically important considerations in participant decision making by seriously mentally or physically ill persons
Data	Participants: 83 from 300 medical students invited to the study
Comparisons	<p>Intervention 1 (n = 28): criteria-oriented teaching (analytic-focused) - 30-minute oral presentation of the RePEAT tool, a structured educational tool for investigators and institutional review boards to evaluate ethical soundness of human research studies</p> <p>Intervention 2 (n = 28): participant-oriented teaching (empathy-focused) - 30-minute oral presentation and video about the general subjective responses of seriously ill individuals who had participated in clinical trials</p> <p>Control (n = 27): no instruction; students read a paper on the biology of cell communication and implications for the future treatment of serious mental and physical disorders</p>
Outcomes	<p>1. Mean composite score on the global assessment of overall "significance of ethical problems" (scale from 1 - no problems/ethically acceptable to 9 - serious problems/ethically unacceptable)</p> <p>2. Mean score on attitude toward ethical conduct of clinical research (scale from 1 - strongly disagree to 9 - strongly agree)</p>
Sources of funding	National Institute of Mental Health (NIMH) and the National Institute for Drug Abuse (1R01DA13139). NIMH Career Development Award to one author (1K02MH0298). NIMH Mentored Career Development Award (K23MH66062) to another author
Setting	University of New Mexico School of Medicine, Albuquerque, New Mexico, USA. Study date: 2001.
Notes	The results of the study were published in 2 articles, in 2005 and 2007

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The articles just state that participants were randomly assigned
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not described
Blinding self reported outcomes?	Unclear	Not described
Incomplete outcome data?	Yes	No attrition
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Yes	None identified

Rolfe 2011

Methods	Non-equivalent control group study design with a historical control
Data	Participants: 156 first-year bioscience undergraduate university students
Comparisons	<p>Intervention (n = 76): students first received an introductory tutorial and received further guidelines on how to access the Turnitin originality report and interpret it for their draft course assignment. After making any changes they considered necessary, they then resubmitted their work in its final form to Turnitin.</p> <p>Control (n = 80): no intervention. Students received an introductory tutorial in the computer laboratory on academic writing and referencing, and submitted their final essays to Turnitin.</p>
Outcomes	<p>1. Number of students who plagiarized course work. Plagiarism was identified as: 1) 4 lines (or more) of word-for-word copying; (2) 4 lines (or more) of poor paraphrasing; 3) lack of a citation accompanying the evidence; 4) lack of a full reference relating to the evidence; 5) any combination of the above. Plagiarism was also defined by a cut-off point of a 30% similarity score.</p> <p>2. Number submissions with plagiarism by grammatical error</p> <p>3. Number submissions with grammatical error plus failure to acknowledge source</p> <p>4. Failure to acknowledge sources</p> <p>Secondary outcome: students' and staff's perceptions of using Turnitin (assessed only in the intervention group)</p>
Sources of funding	De Montfort University Faculty of Health and Life Sciences Teaching Quality Enhancement Fund
Setting	De Montfort University, Leicester, UK. Study dates: 2006 and 2007 academic years.
Notes	Historical control from 2006; intervention in the 2007 course. Outcomes measured only after intervention, no pre-intervention assessment.

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	Attrition small
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	Measuring tools not validated. Only post-test measure.
Blinding?	No	No blinding
Comparability of groups?	Unclear	Only basic demographic comparison (gender, age, average grade). Groups were from courses in different years.
Confounding factors?	No	No regression analysis; no attempt to control for confounding factors

Rose 1998

Methods	Randomized controlled trial with 2 study arms
---------	---

Rose 1998 (Continued)

Data	Participants: 211 out of 1289 graduate students in the physical, biological, engineering and social science fields
Comparisons	<p>Intervention (n = not stated): the survey included the text of the APA's (1992) "Ethical Principles of Psychologists and Code of Conduct," which was described as the policy on authorship at a "hypothetical university." The survey asked for opinions on authorship problems in vignettes about graduate student-faculty collaboration.</p> <p>Control (n = not stated): no policy enclosed with the survey</p>
Outcomes	<p>1. Mean score on attitude scale about ethicality of professor's first authorship place from work in collaboration with a student (scale from 1 - highly unethical to 7 - highly ethical)</p> <p>2. Mean score on the scale about possible response to an authorship problem (scale from 1 - highly unlikely to 7 - highly likely) for 3 actions: contacting a dean or other university authority, making a formal complaint and contacting the journal</p>
Sources of funding	Not stated
Setting	Described as a "large southwestern university in the USA." Study dates not reported.
Notes	Low response rates (211, 21% out of all invited participants)

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The article just states that participants randomly received 1 of 6 versions of the questionnaire
Allocation concealment?	Unclear	Not described
Blinding researcher-assessed outcomes?	Unclear	Not described
Blinding self reported outcomes?	Yes	Participants were unaware of their group assignment
Incomplete outcome data?	No	Small response rate
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Yes	None identified

Schuetze 2004

Methods	Cluster-randomized controlled trial with 2 arms to test instructional intervention on plagiarism prevention
Data	Participants: 76 undergraduate students in 2 lifespan developmental psychology classes
Comparisons	Intervention (n = 36): the students received 1) 30-minute presentation and handout on APA style in-text citations and referencing, which included definitions of various forms of plagiarism, including direct plagiarism (copying information verbatim without citing or including quotation marks), inappropriate-

Schuetze 2004 (Continued)

ly using work from other students without giving credit to the original student, vague or incorrect citations (not clearly indicating where material from another source begins and ends) and mosaic plagiarism (superficial changes to wording or sentence structure) as well as tips on how to avoid plagiarizing; 2) homework assignment which consisted of one page of text from an introduction section in a published manuscript by the author, with all citations removed. The paraphrased material could be easily identified in the chosen section. Other paraphrased sections were somewhat harder to identify but provided good examples for class discussion. Students indicated each sentence they believed should have a citation using consecutive numbers placed at the end of the appropriate sentence. The assignment and grading criteria were explained during approximately 5 minutes of class. Students completed 2 such graded homework assignments over the course of the semester. After the first assignment, students received feedback for missing and unnecessary citations.

Control (n = 40): no intervention. The students received the 30-minute presentation and handout on APA style in-text citations and referencing. There was no homework exercise.

Outcomes	<ol style="list-style-type: none"> 1. Mean score of citation problems in students' term papers (on a scale from 1 - consistent failure to cite properly throughout the paper to 5 - no apparent failures to cite) 2. Mean score of "Understanding of plagiarism" scale (on a scale from 1 - considerably worse than at the beginning of the course to 5 - considerably improved) 3. Mean score on "Confidence in ability to avoid plagiarism" scale (on a scale from 1 - considerably worse than at the beginning of the course to 5 - considerably improved) 4. Mean score on "Understanding of importance of proper citations" scale (on a scale from 1 - considerably worse than at the beginning of the course to 5 - considerably improved) 5. Mean number of correct answers in a knowledge test about plagiarism (out of maximum 5)
Sources of funding	Not stated
Setting	State University of New York College at Buffalo, Buffalo, New York, USA. Study dates not reported.
Notes	—

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	The article just states that classes were randomly assigned
Allocation concealment?	No	No allocation concealment
Blinding researcher-assessed outcomes?	Yes	Outcome assessor was blinded to group assignment
Blinding self reported outcomes?	Yes	Classes were unaware of their group assignment
Incomplete outcome data?	Unclear	Inconsistent reporting of number of participants and unclear attrition rate
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	Unclear	Outcome measurement instrument not validated; risk of contamination between the groups

Strohmetz 1992

Methods	Controlled before-and-after study
Data	Participants: 122 junior and senior college psychology majors attending an undergraduate research methods course
Comparisons	<p>Intervention (n = not stated): role-playing instruction in research ethics - 1) a lecture on research ethics based on the ethical principles of the American Psychological Association (1981), 2) discussion, which include a consideration of why proposed research must be evaluated in terms of the ethical costs and benefits of doing the study as well as the costs and benefits of not doing the study. Students were then given a homework assignment to find a recently published study that they consider to be unethical. They were told to read the study carefully and to be prepared to present it during the next class. During the next class, students present their studies in a discussion led by the instructor. Afterwards, they were asked to rate the cost and utility (benefit) of each study. Students were then asked to role-play as "devil's advocate" and defend the scientific value of their "unethical" study before the rest of the class. Following this role-play, each student re-rated the cost and utility of the studies. Finally, the re-evaluations were discussed to uncover how and why cost and utility ratings may have changed as a result of the students advocating studies that they originally viewed as unethical.</p> <p>Control (n = not stated): no intervention</p>
Outcomes	<p>1. Change in utility rating for presented cases on a scale ranging from no cost or utility (0) to highest cost or utility (100) after discussion of the case with the class as a peer-review board for students' decision, in comparison to the students' initial rating</p> <p>2. Change in cost rating for presented case on a scale ranging from no cost or utility (0) to highest cost or utility (100) after discussion of the case with the class as a peer-review board for students' decision, in comparison to the students' initial rating</p>
Sources of funding	The research was supported by Temple University fellowships awarded to both authors
Setting	Temple University, Philadelphia, Pennsylvania, USA. Study dates not reported.
Notes	The number of participants entering the study not clearly reported

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Unclear	Exact number of participants not provided, attrition unclear
Selective outcome reporting?	No	Only results of statistical analyses reported, without providing scores by study arm
Other bias?	No	Courses were taught by the same instructor; no. of studies discussed differed between the courses; outcome measurement not validated
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups
Confounding factors?	No	No comparison of demographic characteristics between the groups. No characteristics of the students reported.

Walker 2008

Methods	Controlled before-and-after study
Data	Participants: 36 undergraduate students enrolled in 2 sections of methods in psychology course
Comparisons	<p>Intervention (n = 19): paraphrasing training, consisting of: 1) a brief review of the University's Academic Integrity Policy and a discussion of the rules for proper citation, 2) introduction of definitions of plagiarism or patch-writing, 3) students' paraphrasing notes for their first course assignment and discussion about those, 3) practical work with paraphrasing notes for assignments during the courses, 4) guidelines for paraphrasing for final assignment during the semester.</p> <p>Control (n = 17): no intervention. The instructor for the control group presented a brief lecture on plagiarism and paraphrasing at the beginning of the semester. Similar to the training group, students in the control completed original research projects and wrote APA style empirical papers with similar requirements for the number of articles included in the literature review. Students in the control group, however, neither participated in an interactive session on plagiarism nor generated paraphrasing notes.</p>
Outcomes	<p>Plagiarism in students exercise, measured by: a) mean number of copied word strings in a student paper as a measure of plagiarism, b) mean number of word substitutions in a student paper as a measure of plagiarism, c) mean number of word deletions in a student paper as a measure of plagiarism</p> <p>Secondary outcome: usefulness of the course</p>
Sources of funding	Not stated
Setting	Described as "a New England university" (USA). Study dates not reported.
Notes	The assessments occurred in the middle of the semester and for the final course assignment. Final assessment also included 3 additional measures of plagiarism: 1) word additions, 2) word reversals and 3) average plagiarism score (average of all 5 measures added); plus the mean number of words of students' final course assignments.

Risk of bias

Item	Authors' judgement	Description
Incomplete outcome data?	Yes	Attrition small
Selective outcome reporting?	Yes	All outcomes adequately reported
Other bias?	No	No characteristics of the students reported. Only post-test measure.
Blinding?	No	No blinding
Comparability of groups?	Unclear	No comparison of demographic characteristics between the groups
Confounding factors?	No	No regression analysis, no attempt to control for confounding factors. No characteristics of the students reported.

Youmans 2011

Methods	Randomized controlled trial with 4 arms to test training in plagiarism prevention. Allocation to writing assignment was by class and to warning about the use of Turnitin by public randomization at the beginning of the course.
---------	---

Youmans 2011 (Continued)

Data	Participants: 90 undergraduate students majoring in psychology and enrolled in Experimental Psychology (n = 44) or Introduction to Human Factors course (n = 46)
Comparisons	<p>Intervention 1 (n = not stated): requirement to use citation in writing assignment + warning that Turnitin will be used</p> <p>Intervention 2 (n = not stated): requirement to use citation in writing assignment + no warning that Turnitin will be used</p> <p>Intervention 3 (n = not stated): no requirement to use citation in writing assignment + warning that Turnitin will be used</p> <p>Control (n = not stated): no requirement to use citation in writing assignment + no warning that Turnitin will be used</p>
Outcomes	Plagiarism, assessed as the per cent overall of the text in Turnitin, also judged by the researcher
Sources of funding	Not stated
Setting	California State University, Northridge, California, USA. Study dates not reported.
Notes	The course in Experimental Psychology was assigned the writing task with mandatory use of references and the course on Introduction to Human Factors was assigned the writing task where citing peer-reviewed work was encouraged but was optional.

Risk of bias

Item	Authors' judgement	Description
Random sequence generation?	Unclear	Randomization method not described
Allocation concealment?	Yes	Allocation sequence was not known to participants and investigators
Blinding researcher-assessed outcomes?	Yes	Grader was blind to the students' experimental condition
Blinding self reported outcomes?	No	No blinding for 1 of the 2 interventions
Incomplete outcome data?	Yes	No attrition in the study
Selective outcome reporting?	No	Number of participants in study arms not reported
Other bias?	Yes	Not identified

APA: American Psychological Association

ICMJE: International Committee of Medical Journal Editors

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Ali 2014	All participants in the study attended responsible conduct of research (RCR) training; no control group
Bagdasarov 2013	All participants in the study attended RCR training; no control group
Goldie 2001	The intervention included only a small element of research integrity and outcomes addressed only professional ethics, not research ethics
Gurung 2012	Not a before-and-after study design; addressing academic integrity in students, not research integrity
Harkrider 2012	All participants in the study attended RCR training; no control group
Hren 2007	No testing of the groups before the intervention
Kligyte 2008	No control group, only intervention group
May 2014	Study addresses professional ethics, not research integrity or research ethics
McDonalds 2010	Follow-up of journals; unclear interventions; no adequate control group
Mumford 2008	All participants in the study attended RCR training; no control group
Pollock 1995	No testing of the groups before the intervention
Powell 2007	All participants in the study attended RCR training; no control group
Vallero 2007	No control group

RCR: responsible conduct of research

ADDITIONAL TABLES

Table 1. Effects on behavioural change

Study (design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results (I versus C, or I1 versus I2 versus I _n versus C)	Significant difference
Intention for change:						
Rose 1998 (RCT)	Graduate students	I: Presence of authorship policy when judging authorship disputes in student-faculty collaboration	C: No authorship policy	Likelihood, effectiveness and consequences (mean score on a scale from 1 – low to 7 – very likely) of reporting a problem with authorship arrangement (in 3 vignettes with professor's first authorship: 1 - professor idea, student work; 2 - student idea, professor work;	Men I: 2.80 (SE 0.21) versus C: 1.89 (SE 0.22), $F(1,156) = 9.08$, P value < 0.01 Women I: 1.94 (SE 0.29) versus C: 2.79 (SE 0.22) for vignette where professor is the first author on student's thesis: No difference for effectiveness and consequences for any vignette	+ for men – for women

Table 1. Effects on behavioural change (Continued)

				3 - student's dissertation):		
				a) talking to a dean		
				b) filing a complaint		
				c) contacting a journal		
Actual change - research integrity						
Marušić 2006 (RCT)	Journal authors	I1: Categorical form for declaring authorship contribution I2: Instructional form for declaring authorship contribution	C (I3): Open-ended form for declaring authorship contribution	Number of authors not satisfying authorship criteria of the International Committee of Medical Journal Editors (ICMJE)	I1: 68.7 versus I2: 32.5 versus I3: 62.6; I3 versus I1 z = -3.034, P value = 0.002; I3 versus I2 z = -2.884, P value = 0.004, no diff. I2 versus I3 (z = 0.3315, P value = 0.74)	+ I2
Ivaniš 2008 (RCT)	Journal authors	I1: Ordinal rating of authorship contributions (on a scale from 0 = none to 4 = full)	C (I2): Binary rating of authorship contribution	Per cent of authors satisfying authorship criteria of the International Committee of Medical Journal Editors (ICMJE)	I1: 76 versus I2: 39, P value < 0.05	+ for I1
Actual change - plagiarism prevention						
Ballard 2013 (RCT)	Undergraduate students	I1: Academic integrity module + Turnitin submission I2: Academic integrity module + no Turnitin submission I3: No academic integrity module + Turnitin submission	C: No instruction, no Turnitin submission	Mean Turnitin similarity score (form 0 to 100)	I1: 11.32 (SD 10.48) versus I2: 12.24 (SD 15.09) versus I3: 15.42 (SD 15.85) versus C: 14.78 (SD 14.40); F(1,92) = 0.072, MSE = 193.26, P value = 0.789, η ² = 0.001	-
Landau 2002 (RCT)	Undergraduate students	I1: Instruction about plagiarism with feedback I2: Instruction about plagiarism with examples I3: Instruction about plagiarism	C: No instruction	Frequency (mean no.) of using: a) overlapping words b) 2-word strings c) 3-word strings in assignments	a) I1: 21.4 versus I2: 17.2 versus I3: 17.5 versus C: 19.0, F(1,90) = 6.86, P value = 0.01, MSE = 28.12, η ² = 0.07 b) I1: 8.3 versus I2: 4.1 versus I3: 4.9 versus C 8.3, F(1,90) = 12.39, P value	+ for I2

Table 1. Effects on behavioural change (Continued)

		with examples and feedback			= 0.01, MSE = 27.14, $\eta^2 = 0.12$ c) I1: 4.3 versus I2: 1.4, I3: 1.6 versus C: 4.1, $F(1,90) = 11.13$, P value = 0.01, MSE = 15.27, $\eta^2 = 0.11$	
Newton 2014 (RCT)	Undergraduate students	I: Training session on paraphrasing, patch-writing and plagiarism	C: No instruction	Quality of paraphrasing in a written assignment (average score, marked by 2 assessors, score range not provided): a) referencing b) patch-writing c) plagiarism	a) I: 3.49 (SE 0.19) versus C: 2.54 (SE 0.20), $F(1,116) = 11.7$, P value < 0.01, $\eta^2 = 0.09$ b) I: 4.30 (SE 0.16) versus C: 3.77 (SE 0.17), $F(1,116) = 5.63$, P value < 0.05, $\eta^2 = 0.05$ c) I: 4.93 (SE 0.10) versus C: 4.55 (SE 0.11), $F(1,116) = 6.97$, P value < 0.01, $\eta^2 = 0.06$	+
Risquez 2011 (RCT)	Undergraduate students	I: 1-hour in-class tutorial on plagiarism prevention	C: No intervention – regular class	Reported engagement (mean score, scale from 1 = completely disagree to 5 = completely agree) in: a) copying text and using it without citations b) using internet to copy text and using it without citations c) using internet to purchase a paper and present as own	a) I: 1.38 versus C: 1.53, $t = -1.610$, P value > 0.05 b) I: 1.39 versus C: 1.37, $t = 0.187$, P value > 0.05 c) I: 1.10 versus C: 1.00, $t = 2.080$, P value < 0.05	+ for c (worse behaviour)
Schuetze 2004 (RCT)	Undergraduate students	I: Homework assignment to reduce citation problems	C: Standard teaching	Mean no. citation problems in students' term papers	I: 2.57 (SD 1.22) versus C: 3.87 (SD 1.11), $t(73) = -4.57$, P value < 0.01	+
Youmans 2011	Undergraduate students	I1: Requirement to use citation in writing assignment + warning that Turnitin will be used I2: Requirement to use citation in writing assignment + no warning that Turnitin will be used I3: No requirement to use citation in writ-	C: No requirement to use citation in writing assignment + no warning that Turnitin will be used	Mean per cent text overlap in Turnitin report	I1 versus I2 versus I3 versus C $F(1,85) = 0.96$, not significant Requirement to use citations versus no requirement: 10.26% (SD 5.66%) versus 4.76% (SD 7.30%), $F(1,85) = 8.35$, P value < 0.001, $\eta^2 = 0.17$ Warning versus no warning on use of Turnitin: 7.59% (SD 7.17%) versus 7.29% (SD 7.10), $F(1,85) = 2.08$, not significant	-

Table 1. Effects on behavioural change (Continued)

		ing assignment + warning that Turnitin will be used				
Estow 2011 (CBA)	Undergraduate students	I: Plagiarism themes and assignments embedded in a methods and statistics course	C: Non-intervention course	Quality of paraphrasing in course assignments (mean score, scale from 1 – direct copying of a significant proportion of original without quotation marks to 4 – good paraphrasing)	I: pretest 2.54 (SD 0.86) post-test 3.32 (SD 0.86) versus C: pretest 3.12 (SD 0.78) post-test 2.83 (SD 0.81); $t(41) = 1.55$, P value = 0.13	–
Walker 2008 (CBA)	Undergraduate students	I: Paraphrasing training to avoid plagiarism	C: No training	Post-test mean score of extent of plagiarism in written course assignments (lower score, less plagiarism)	I: 1.75 (SD 7.49) versus C: 4.56 (SD 6.52), P value < 0.01	+
Chao 2009 (non-equivalent CBA)	Undergraduate students	Training in plagiarism prevention, involving: I1: instructions and examples I2: discussion and student practice work	C: Warning about plagiarism, no training	Percentage of students with assignments containing some amount of plagiarism Mean percentage plagiarized text in the course assignments	I1: 29% versus I2: 36% versus C: 55% $\chi^2_2 = 5.139$, P value = 0.077 I1: 2.29% versus I2: 1.9% versus C: 5.45%, $F(2,113) = 3.399$, P value = 0.037, post hoc Tukey test for I2 versus C P value = 0.46	– + for I2
Kose 2011 (non-equivalent CBA)	Undergraduate students	I: Use of Turnitin in practising avoiding plagiarism in academic writing	C: No intervention	Range of percentage plagiarism level in submitted essays	I: pretest 8% to 22% post-test 0% to 12% versus C: post-test 2% to 22%	+
Belter 2009 (non-equivalent CBA)	Undergraduate students	C: Online module on research integrity within a course	C: No module on research integrity	Percentage plagiarized course papers	I: 6.5 versus C: 25.8, $\chi^2_1 = 18.39$, P value < 0.001	+
Bilić-Zulle 2008 (non-equivalent CBA)	Undergraduate students	I1: Warning that plagiarism is forbidden I2: Warning that student essays will be examined by plagiarism detection software and that plagiarism will be penalized	C: No intervention	Median (5 to 95 percentile) proportion of plagiarized text in essays	I1: 21 (0 to 87) versus I2: 2 (0 to 20) versus C: 17 (0 to 89); $H = 84.64$, P value < 0.001; post hoc test P value < 0.05 for I2 versus I1 and C	+ for I2
Marshall 2011	Students	I1: Warning about plagiarism (year 2)	C: No intervention (year 1)	Number of plagiarism occurrences	I2: 0.3% versus I1 and C: 1.9%, P value = 0.013	+

Table 1. Effects on behavioural change (Continued)

(non-equivalent CBA)		I2: Training in avoiding plagiarism (year 3 and 4)		Overall percentage text match in year 1 to 4 for assignments:	a) I2: 20.8 and 18.1 versus I1: 37.4 versus C: 28.5	+ for I2
				a) Critical appraisal	b) I2: 23.6 and 20.8 versus I1: 28.0 versus C: 30.3	
				b) Health information	c) I2: 19.5 and 11.7 versus I1: 21.6	
				c) Principles and practice of HTA		
Rolfe 2011	Undergraduate students	I: Instruction and feedback based on a plagiarism detection software	C: No intervention	Percentage of students who plagiarized	I: 80 versus C: 72	-
(non-equivalent CBA)				Percentage of submissions with plagiarism due to poor paraphrasing	I: 9 versus C: 28, P value < 0.05	+
				Percentage of submissions with plagiarism due to poor paraphrasing and failure to acknowledge source:	a) I: 30 versus C: 46 b) I: 16 versus C: 8 c) I: 33 versus C: 16	-
				a) plagiarism without citation		
				b) plagiarism without reference		
				c) plagiarism with no citation or reference		
				Percentage of submission with lack of citations	59 versus 32, P value < 0.05	+ (intervention worse)

CBA: controlled before-and-after study

ICMJE: International Committee of Medical Journal Editors

HTA: health technology assesment

MSE: mean standard error

RCT: randomized controlled trial

SD: standard deviation

SE: standard error

Table 2. Effects on knowledge and/or skills related to research ethics/integrity

Study (design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results	Significant difference
Aggarwal 2011 (RCT)	Researchers	I1: Research ethics course delivered online	I2 (C): Research ethics course delivered on site	Knowledge (percentage of correct responses; median, range): a) pre-course b) immediately after the course	a) I1: 62 (40 to 89) versus I2: 69 (36 to 39), P value = 0.07 b) I1: 77 (43 to 95) versus I2: 82 (44 to 93), P value = 0.02 c) I1: 80 (-15 to 34) versus I2: 83 (54 to 98), P value = 0.69	+ for I2 in b)

Table 2. Effects on knowledge and/or skills related to research ethics/integrity (Continued)

				c) 3 months after the course	P value < 0.005 versus baseline for I1 and I2	
				Difference in knowledge gain from baseline to 3 months (percentage of correct responses; median, range)	I1: 13% (-15 to 34%) versus I2: 17% (0 to 41%) P value = 0.14	-
Arnott 2008 (CBA)	Undergraduate students	I: Dialogue-based computer tutoring system on research methodology	C: Standard teaching	Post-test-pretest gain (main) in the score on a written knowledge test (30 questions)	I: 10.9% (SD 11.8) versus C: 3% (SD 9.4%); effect size 0.75 SDs; F(1,94) = 5.99, P value = 0.016, $\eta^2 = 0.06$	+
Fisher 1997 (CBA)	Undergraduate students	I: Enhanced ethics instruction	C: Standard instruction	Knowledge of research ethics procedures and ability to weigh scientific responsibility and participant rights and welfare (mean score (0 to 5) on a written essay)	I: Pretest 1.69 (SD 1.00), post-test 2.13 (SD 1.03) versus C: 1.72 (SD 0.92), post-test 1.79 (SD 1.20); critical difference 0.28, P value < 0.01, effect size for pretest-post-test difference $d = 0.44$ for intervention and $d = 0.05$ for control	+

CBA: controlled before-and-after study

RCT: randomized controlled trial

SD: standard deviation

Table 3. Effects on knowledge and/or skills related to plagiarism

Study (design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results	Significant difference
Landau 2002 (RCT)	Undergraduate students	I1: Instruction about plagiarism with feedback I2: Instruction about plagiarism with examples I3: Instruction about plagiarism with examples and feedback	C: No instruction	Plagiarism Knowledge Survey mean score* (lower scores indicated better knowledge)	Pretest I1: 8.00 versus I2: 7.84 versus I3: 7.60 versus C: 8.28, post-test I1: 6.75 versus I2: 7.09 versus I3: 6.95 versus C: 7.99; F(1,90) = 4.41, P value = 0.04, MSE = 2.17, $\eta^2 = 0.05$	+
Moniz 2008 (RCT)	Undergraduate students	Enhanced instruction on plagiarism: I1: Direct	C: Standard teaching on plagiarism	Knowledge and understanding of plagiarism	No effect, raw data or statistical analysis not presented	-

Table 3. Effects on knowledge and/or skills related to plagiarism (Continued)

I2: Student-centred

Newton 2014 (RCT)	Undergraduate students	I: Training session on paraphrasing, patch-writing and plagiarism	C: No intervention	Knowledge about in-text referencing (average score from 0 to 5)	I: 3.66 (SE 0.13) versus C: 3.22 (SE 0.13), $F(1,116) = 5.75$, P value < 0.05 , $\eta^2 = 0.05$	+
Schuetze 2004 (RCT)	Undergraduate students	I: Homework assignment to reduce citation problems	C: Standard teaching	Score on knowledge test on plagiarism (mean score, max 5)	I: 3.51 (SD 1.39) versus C: 3.85 (SD 1.01); $t(74) = 1.20$, P value = 0.24	-
Chertok 2014 (CBA)	Undergraduate students	I: Course on academic integrity	C: No intervention	Score on 13 true/false questions on knowledge about plagiarism (mean no. correct answers)	Pretest I: 10.8 (SD 1.7) versus C: 10.5 (SD 1.6) Post-test I: 10.8 (SD 1.9) versus C: 10.1 (SD 1.8) No difference in change pre-test to post-test (P value = 0.09).	-
Estow 2011 (CBA)	Undergraduate students	I: Plagiarism themes and assignments embedded in a methods and statistics course	C: Non-intervention course	Pretest-post-test mean difference in identifying plagiarism in test examples (no. sentences correctly identified as plagiarism)	1.52 (SD 0.98) to 2.59 (SD 0.75) versus 1.94 (SD 0.43) to 2.06 (SD 0.66); $F(1,42) = 9.19$, P value = 0.004, $\eta^2 = 0.1.8$	+
				Pretest-post-test mean difference in number of identified strategies to avoid plagiarism (no. listed strategies)	2.44 (SD 0.93) to 2.69 (SD 0.60) versus 2.63 (SD 1.18) to 2.00 (SD 0.73); $F(1,41) = 5.86$, P value = 0.02, $\eta^2 = 0.13$	+
Barry 2006 (CS)	Undergraduate students	I: Assignments in paraphrasing quotes from publications	C: No intervention	Pretest-post-test mean difference in score on plagiarism definition test (scale not provided)	2.14 (SD 0.81) to 3.23 (SD 0.99) versus (pretest not measured) to 2.33 (SD 1.08); $t(62) = 3.45$, P value = 0.001 for post-test-control comparison; pretest-post-test comparison $d = 1.26$, post-test-control comparison $d = 0.86$	+

CBA: controlled before-and-after study

CS: controlled study

MSE: mean standard error

RCT: randomized controlled trial

SD: standard deviation

SE: standard error

Table 4. Effects on attitudes and/or perceptions related to research ethics/integrity

Study (design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results	Significant difference
Roberts 2007 (RCT)	Medical students	I1: Criteria-oriented (analytic-focused) instruction I2: Participant (empathy)-oriented instruction	C: No instruction	Significance of presented ethical problems (overall mean score for 10 vignettes, scale from 1 – ethically acceptable to 9 – not at all ethically acceptable)	I1: 5.76 (SD 0.96) versus I2: 5.09 (SD 1.01) versus C: 5.08 (SD 1.22); maximum d = 0.34, intervention group main effect $F(2,71) = 4.16$, P value < 0.02	+ for I1
				Attitudes related to research participant decision-making (individual mean score for 10 attitudes, scale from 1 – strongly disagree to 9 – strongly agree)	Mean d = 0.63 for I2 versus C for 8 out of 10 attitudes (all simple effects P value < 0.05, one P value < 0.08); mean d = 0.24 for I1 versus C for all attitudes (all P value > 0.20, one P value < 0.07); intervention group attitude interaction $F(18,122) = 2.50$, P value < 0.01	+ for I2 – for I1
				Perception of clinical research participants' attitudes related to decisional capacity (overall mean score for 2 attitudes, scale from 1 – strongly disagree to 9 – strongly agree)	I1: 4.63 (SD 1.49), d = 0.36 versus I2: 2.68 (SD 1.56), d = -0.83 versus C: 4.04 (SD 1.85), main effect of intervention group P value < 0.0001	+ for I2 – for I1
				Perception of clinical research participants' attitudes related to vulnerability (overall mean scores on 2 attitudes, scale from 1 – strongly disagree to 9 – strongly agree)	I1: 4.62 (SD 1.72), d = 0.12 versus I2: 3.88 (SD 1.80), d = -0.27 versus C: 4.39 (SD 2.14), no significant effect	–
				Perception of clinical research participants' attitudes related to surrogate decision-makers (overall mean score for 4 attitudes, scale from 1 – strongly disagree to 9 – strongly agree)	I1: 6.02 (SD 1.73), d = -0.20 versus I2: 6.43 (SD 1.80), d = 0.01 versus C 6.41 (SD 2.14), main effect of intervention group P value < 0.65	–
				Ethically relevant attitudes related to clinical research (overall mean score for 10 attitudes, scale from 1 – strongly disagree to 9 – strongly agree)	I1: 7.02 (SD 0.70), d = 0.44 versus I2: 6.68 (SD 0.73), d = 0.00 versus C: 6.68 (SD 0.88); no main effect of intervention group, P value < 0.15	–
				Importance of ethical duties of investigators	I1: 7.50 (SD 1.08), d = 0.39 versus I2: 7.10 (SD 1.13), d = 0.06	–

Table 4. Effects on attitudes and/or perceptions related to research ethics/integrity (Continued)

				(overall mean score for 13 duties, scale from 1 – not at all important to 9 – very important)	versus C: 7.03 (SD 1.37); no main effect of intervention group, P value < 0.28	
Rose 1998 (RCT)	Graduate students	I: Presence of authorship policy when judging authorship disputes in student-faculty collaboration	C: No authorship policy	Attitude about ethicality of professor's first authorship place from work in collaboration with a student (overall mean score on 3 vignettes, scale from 1 – highly unethical to 7 – highly ethical)	Men I: 3.15 (SE 0.17) versus C: 3.22 (SE 0.18); F(1,156) = 0.12, P value > 0.12	– for men + for women
					Women I: 2.36 (SE 0.19) versus C: 3.10 (SE 0.23); F(1,116) = 7.81, P value < 0.01	
				Attitude about ethicality or professor submitting the manuscript to a journal without discussing authorship with student	No effect of intervention (data not presented in published article)	–
Clarkeburn 2002 (CBA)	Undergraduate students	I: Ethics and research integrity discussion within a science course	C: No intervention – regular science course	Ethical sensitivity (mean score on Test of Ethical Sensitivity in Science, TESS; max score 15)	I: Pretest 4.68 (SD 2.27) versus post-test 5.30 (SD 2.25); P value < 0.05	+
					C: pretest 4.89 (SD 2.18) versus post-test 4.67 (SD 1.95); P value > 0.05	
					More students progressing in I than in C ($\chi^2 = 24.941$, P value < 0.0001)	
				Moral reasoning (Defining Issues Test, DIT; mean P-score pretest, % students with or without change post-test)	I: Pretest 32.89 (SD 15.83); students post-test progressing 51.8%, regressing 46.5%, no change 1.7%	–
					C: Pretest 30.44 (SD 14.18); students post-test progressing 59.3%, regressing 39.5%, no difference 1.2, P value > 0.05	
				Meta-ethical understanding (Perry Questionnaire, progression to Perry types)	P value = 0.461 (data not presented for individual groups)	–
Hull 1994 (CBA)	Graduate students	I: Research ethics course	C: No intervention	Socio-moral reasoning (Sociomoral Reflection Objective Measure – Short Form, SROM-SF; scale not provided in the published article; score change pre-post intervention)	I: +3.20 versus C: -10.33; main effects F(1,37) = 5.037, P value < 0.05	+

Table 4. Effects on attitudes and/or perceptions related to research ethics/integrity (Continued)

May 2013 (CBA)	Undergraduate students	I1: Stand-alone ethics training	C: No intervention	Perspective taking (mean score on 6-item measure, scale from 1 –strongly disagree to 7 – strongly agree)	I1: 5.30 (SD 0.50) versus I2: 5.34 (SD 0.51) versus C: 4.96 (SD 0.52); effect of intervention $F(2,83) = 3.85$, P value < 0.05, $\eta^2_p = 0.09$	+ for both I1 and I2
		I2: Course-embedded ethics modules		Moral efficacy (mean score on 14-item measure, scale from 1 –not confident at all to 5 – very confident)	I1: 3.90 (SD 0.46) versus I2: 3.90 (SD 0.46) versus C: 3.56 (SD 0.48); effect of intervention $F(2,74) = 3.79$, P value < 0.05, $\eta^2_p = 0.09$	+ for both I1 and I2
				Moral courage (mean score on 6-item measure, scale from 1 –strongly disagree to 7 – strongly agree)	I1: 5.46 (SD 0.54) versus I2: 5.58 (SD 0.55) versus C: 5.13 (SD 0.56); $F(2,83) = 3.81$, P value < 0.05, $\eta^2_p = 0.08$	+ for both I1 and I2
				Moral judgement (Defining Issues Test, DIT; mean N2 index)	I1: 40.84 (SD 12.19) versus I2: 42.94 (SD 11.86) versus C: 42.51 (SD 12.19); no effect of intervention ($F(2,123) = 0.35$, P value > 0.10 $\eta^2_p = 0.01$	–
				Moral meaningfulness (score on 4-item measure, scale from 1 –strongly disagree to 7 – strongly agree)	I1: 5.83 (SD 0.87) versus I2: 6.13 (SD 0.87) versus C: 5.62 (SD 0.90); no effect of intervention ($F(2,71) = 3.81$, P value > 0.10 $\eta^2_p = 0.05$	–
Strohmetz 1992 (CBA)	Undergraduate students	I: Role-play instruction in research ethics	C: Standard research ethics course	Perceptions of ethical utility of a published study (pre-post change in score on a scale from 0 – none to 100 – highest)	I: 3.13 to 10.76 for 6 individual courses; P value range 0.069 to 0.0002 versus C: -1.18 P value = 0.39	+ (increase)
				Perceptions of ethical percost of a published study (pre-post change in score on a scale from 0 – none to 100 – highest)	I: Pre-post change range perception change from -0.22 to -8.98 for 6 individual courses; P value range 0.50 to 0.00001 versus C: pre-post change 1.14, P value = 0.41	+ (decrease)

CBA: controlled before-and-after study

RCT: randomized controlled trial

SD: standard deviation

SE: standard error

Table 5. Effects on attitudes and/or perceptions related to plagiarism

Study (design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results	Significant difference
Brown 2001 (RCT)	Undergraduate students	I1: Educational statement	C: No intervention	Score (mean) for severity of behaviour in examples of plagiarized text* (scale 0 – none to 100)	a) I1: 81.01 (SD 2.32) versus I2: 70.72 (SD 2.61) versus C: 70.72 (SD 1.89); main ef-	+ for I1

Table 5. Effects on attitudes and/or perceptions related to plagiarism (Continued)

	on plagiarism			– extremely serious/everybody does it):	fect of intervention F(2,195) = 5.45, P value < 0.01	
	I2: Warning statement on plagiarism			a) respondents' views	b) I1: 85.22 (SD 1.88) versus I2: 75.51 (SD 2.64) versus C: 76.81 (SD 1.89); main effect of intervention F(2,195) = 4.1, P value < 0.02	
				b) perceived staff view		
				c) perceived frequency among colleagues	c) I1: 68.12 (SD 1.59) versus I2: 75.51 (SD 1.30) versus C: 74.20 (SD 1.02); main effect of intervention F(2,195) = 5.2, P value < 0.001	
				Perception of the necessity of citing sources when writing (composite score for verbatim copy or uncited paraphrasing; scale from 0 – not necessary to 100 – absolutely necessary)	I1: 87.1 (SD 22.4) versus I2: 82.4 (SD 26.3) versus C: 83.4 (SD 23.5); no main effect of intervention F(1,195) = 0.94, P value > 0.05	–
				Perceived likelihood that their colleagues would cite a source (composite score for verbatim copy or uncited paraphrasing; scale from 0 – very unlikely to 100 – very likely)	I1: 55.6 (SD 34.4) versus I2: 62.7 (SD 29.5) versus C: 59.1 (SD 30.4); no main effect of intervention F(2,195) = 2.1, P value > 0.005	–
Compton 2008 (RCT)	Undergraduate students	Message about plagiarism:	C: No intervention	Attitudes towards plagiarism (assessed by 5 measures on a scale from 0 = negative to 6 = positive):	Overall no effect of intervention (results of ANCOVA not presented)	–
		I1: Fear-based		a) threat generated by inoculation messages		
		I2: Guilt-based		b) attitude towards plagiarism		
		I3: Rational (cognitive-) based		c) importance or salience of the topic of plagiarism (involvement in the issue)		
				d) how often plagiarism is discussed		
				e) vested interest in plagiarism		
				Attitudes towards justification of plagiarism (attack messages in phase 3), assessed by 2 measures on scale from 0 = more resistance to 6 = less resistance towards attack message):	Overall no effect of intervention (results of ANCOVA not presented)	–

Table 5. Effects on attitudes and/or perceptions related to plagiarism *(Continued)*

				a) attitude towards the justification of plagiarism		
				b) perceived credibility of the source of messages justifying plagiarism in relation to its character, competence and sociability		
Dee 2012 (RCT)	Undergraduate students	I: On-line-based tutorial on plagiarism	C: No intervention	Perception of having a good understanding of plagiarism in academic writing (score difference I to C, 5-point scale from strongly disagree to strongly agree)	0.075 (SE 0.032), P value < 0.05	+
				Perception of:		
				a) participants' actual behaviour in avoiding plagiarism	a) 0.006 (SE 0.136), P value > 0.05	-
				b) confidence in knowledge of how to avoid it	b) 0.034 (SE 0.336), P value > 0.05	
				c) whether the instructor would notice or address plagiarism in writing assignments	c) 0.061 (SE 0.407), P value > 0.05	
				(score difference I to C, 5-point scale from strongly disagree to strongly agree)		
Newton 2014 (RCT)	Undergraduate students	I: Training session on paraphrasing, patch-writing and plagiarism	C: No intervention	Perception (scale from -3: strongly disagree to +3: strongly agree) of:	Multiple regression analysis presented:	+ for a) and c)
				a) note-taking ability	a) b = 0.81, SE b = 0.23, β = -0.27, P value < 0.001	- for b)
				b) confidence in paraphrasing source material	b) b = -0.16, SE b = 0.27, β = -0.05, P value > 0.05	
				c) confidence in referencing sources in assignments	c) b = -0.57, SE b = 0.27, β = 0.17, P value < 0.05	
Risquez 2011 (RCT)	Undergraduate students	I: 1-hour in-class tutorial on plagiarism prevention	C: No intervention	Perceived seriousness of breaching academic guidelines in examples (mean score, scale from 0 – no breach to 100 – extremely serious breach):	a) I: 66.99 versus C: 55.47; t = 2.903; P value < 0.01	+ for a) and b)
				a) as participants' own view	b) I: 75.47 versus C: 63.54; t = 3.084, P value < 0.01	
				b) as participants' assessment of lecturers' view	c) I: 86.59 versus C: 81.13, t = 1.546, P value > 0.05	
				c) expectation from the student in example		
				Attitudes towards plagiarism (mean score, scale from 1 – completely disagree to 5 – completely agree) for copying text and using	Copying text without citation:	-

Table 5. Effects on attitudes and/or perceptions related to plagiarism (Continued)

				internet to copy text without citation:	a) I: 1.58 versus C: 1.63, t = -0.317	
				a) wrong in my view	b) I: 1.30 versus C: 1.29, t = 0.089	
				b) wrong in lecturer's view		
				c) strictly punished in college	c) I: 1.52 versus C: 1.49, t = -0.249	
				Using internet to copy text without citation:		
					a) I: 1.49 versus C: 1.58, t = -0.677	
					b) I: 1.31 versus C: 1.29, t = 0.214	
					c) I: 1.49 versus C: 1.48, t = -0.085	
					all P values > 0.05	
Schuetze 2004 (RCT)	Undergraduate students	I: Homework assignment to reduce citation problems	C: Standard teaching	Confidence in one's ability to avoid plagiarism in future (mean score, from 1 – considerably worse to 5 – considerably improved)	I: 4.17 (SD 0.67) versus C: 2.75 (SD 0.84); t = -7.50, P value = < 0.001	+
				Perception of own understanding of plagiarism (mean score, from 1 – considerably worse to 5 – considerably improved)	I: 4.06 (SD 0.64) versus C: 3.12 (SD 0.97); t = 73, P value < 0.001	+
				Perception of own understanding of importance of proper citations (mean score, from 1 – considerably worse to 5 – considerably improved)	I: 4.10 (SD 0.82) versus C: 3.90 (SD 0.90); t(73) = -9.6, P value = 0.12	-
Chertok 2014 (CBA)	Undergraduate students	I: Course on academic integrity	C: No intervention	Pre-post change in attitude score on 17-question test on (range 0 - appropriate ethical attitude) to 51 - inappropriate attitude)	I: 6.3 (SD 7.1) versus C: 3.3 (SD 5.4), P value < 0.001	+

CBA: controlled before-and-after study

RCT: randomized controlled trial

SD: standard deviation

SE: standard error

Table 6. Participants' reaction to intervention (RCT - randomized controlled trial; CBA - controlled before-and-after)

Study (Design)	Participants	Intervention (I)	Control (C)	Outcomes (O)	Results	Significant difference
Research integrity:						

Table 6. Participants' reaction to intervention (RCT - randomized controlled trial; CBA - controlled before-and-after) (Continued)

Aggarwal 2011 (RCT)	Re-searchers	I (I1): Re-search ethics course delivered on-site	C (I2): Re-search ethics course delivered on-line	Satisfaction with the course, with regard to: a) value of group discussions (from 1 – strongly disagree to 5 – strongly agree) b) value for work in research (from 1 – strongly disagree to 5 – strongly agree) c) speed of covering the course material (too slow, just right, too fast)	a) I1: 15% score 5 versus I2: 24% score 5, P value = 0.003 b) I1: 17% score 5 versus I2: 25% score 5, P value = 0.005 c) I1: 10% too fast versus I2: 0% too fast, P value < 0.001	+ for I2
Fisher 1997 (CBA)	Undergraduate students	I: Enhanced - ethics instruction	C: Standard instruction	Students' and staff' satisfaction with the course: a) difficulty (3 items, on a scale from 1 – elementary to 5 – very difficult, b) value (6 items, on a scale from 1 – excellent to 5 – poor) c) additional topics addressed (6 items on a scale from 1 – strongly agree to 4 – strongly disagree)	Mean score range: a) students 3.08 to 3.19, staff 2.86 to 3.43 b) students 2.43 to 2.82, staff 2.00 to 2.57 c) students 2.13 to 2.36, staff 1.00 to 2.00	+
Plagiarism:						
Dee 2012 (RCT)	Students	I: On-line-based tutorial on plagiarism	C: No intervention	Reaction to the course (score on a 5-point scale from strongly disagree to strongly agree): a) enjoyed the class b) found it academically difficult c) found it or stressful d) procrastinated on writing assignments	Score difference: a) I: versus C: 0.301 (SE 0.236), P value > 0.05 b) I versus C: -0.030 (SE 0.122), P value > 0.05 c) I versus C: -0.176 (SE 0.129), P value > 0.05 d) I versus C: 0.050 (SE 0.161), P value > 0.05	-
Moniz 2008 (RCT)	Undergraduate students	Enhanced instruction on plagiarism: I1: Direct I2: Student-centred	C: Standard teaching on plagiarism (PowerPoint)	Reaction to the course (scale not specified, higher score greater satisfaction): a) highly interactive b) principles of plagiarism taught c) mostly lecturing	a) Mean I1: 2.63 (SD 1.13) versus I2: 2.04 (SD 0.80) versus C: 2.61 (SD 0.95) F(2,197) = 7.21, P value < 0.01 b) Mean I1: 1.88 (SD 1.30) versus I2: 1.61 (SD 0.80) versus C: 1.88 (SD 1.01) F(2,197) = 1.01, P value = 0.37	+ for a), c), d), f)

Table 6. Participants' reaction to intervention (RCT - randomized controlled trial; CBA - controlled before-and-after) (Continued)

				d) seatwork exercises used	c) Mean I1: 2.38 (SD 1.06) versus I2: 3.09 (SD 0.89) versus C: 2.30 (SD 1.04) F(2,197) = 11.43, P value < 0.01	
				e) plagiarism cases discussed		
				f) group work	d) Mean I1: 2.28 (SD 1.27) versus I2: 1.82 (SD 0.76) versus C: 2.24 (SD 1.11) F(2,197) = 3.55, P value = 0.03	
					e) Mean I1: 2.02 (SD 1.24) versus I2: 1.74 (SD 0.84) versus C: 1.84 (SD 1.06) F(2,197) = 1.09, P value = 0.34	
					a) Mean I1: 3.38 (SD 1.35) versus I2: 1.58 (SD 0.94) versus C: 3.78 (SD 1.23) F(2,197) = 60.45, P value < 0.01	
Kose 2011	Undergraduate students	I: Use of Turnitin in practising avoiding plagiarism in academic writing	C: No use of Turnitin software	Percentage of participants in the intervention group reporting that the software was easy to use	81% (13 of 16)	+
				Percentage of participants in the intervention group who judged the software as useful	100% (all 16)	+
Rolfe 2011	Undergraduate students	I: Instruction and feedback based on a plagiarism detection software	C: No intervention	Participants reporting:	Per cent positive response:	+
				a) positive experience with the feedback	a) 100% (52 out of 52)	
				b) enough training	b) 83% (43 out of 52)	
				c) help in improving work	c) 96% (50 out of 52)	
Walker 2008 (CBA)	Undergraduate students	I: Paraphrasing training to avoid plagiarism	C: No training	Participants judging the training (scale from 1 – not at all to 7 – very much):	Mean score:	+
				a) useful	a) I: 5.30 (SD 1.26)	
				b) helpful	b) I: 5.35 (SD 1.30)	

CBA: controlled before-and-after study

RCT: randomized controlled trial

SD: standard deviation

APPENDICES

Appendix 1. Appendix 1: search strategies

The searches were based on the strategy developed in MEDLINE and adapted as appropriate to the specifications of each database and website. The strategy was deliberately designed to capture a broad range of references and the 'explode' feature was used wherever this was applicable to the database. There were no language restrictions.

All information sources were searched most recently in April 2015 for publications from January 1990 to December 2014.

MEDLINE (OvidSP) search strategy

1. Scientific Misconduct/
2. Fraud/
3. exp Ethics, Research/
4. (research adj3 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent\$ or bias\$ or fraud\$ or honest\$ or reliab?l\$ or fair\$ or impartial\$ or selective\$)).tw.
5. ((scientific or academic) adj3 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)).tw.
6. ((researcher\$ or scientist\$) adj3 (integrity or honest\$)).tw.
7. Plagiarism/
8. (plagiari\$ or falsif\$).tw.
9. Publication Bias/
10. Duplicate Publication as Topic/
11. Retraction of Publication as Topic/
12. Peer Review, Research/
13. (data adj3 (interpretat\$ or inaccura\$ or inadequa\$ or deceptive or deceit or bias\$ or impartial or manipulat\$ or misus\$ or misleading or mispresent\$ or mistreat\$ or selective or suppress\$ or fabricat\$ or fraud\$ or falsif\$ or false)).tw.
14. Research Report/
15. (report\$ adj3 (selective or deceptive or deceit or misleading or inadequate or independent)).tw.
16. (research adj3 (underreport\$ or under-report\$)).tw.
17. ((publication\$ or publishing) adj3 ethics).tw.
18. (bias adj3 (publication\$ or publishing or analys#s or design)).tw.
19. (publication\$ adj3 (redundant or duplicate or multiple or salami or undeserving)).tw.
20. (inaccura\$ adj3 citation\$).tw.
21. Authorship/
22. ((author\$ or contribut\$) adj3 (undeserv\$ or ghost or guest or gift\$)).tw.
23. Conflict of Interest/
24. (interest adj3 (conflict or competing)).tw.
25. or/1-24
26. exp Education, Professional/
27. exp Teaching/
28. exp Curriculum/
29. Mentors/
30. (educat\$ or teach\$ or train\$ or motivat\$ or instruct\$ or interven\$ or promot\$ or supervis\$ or mentor\$).tw.
31. (course\$ or seminar\$ or workshop\$).tw.
32. exp Policy Making/

33. Program Development/
34. ((program\$ or plan\$ or policy or rule\$ or procedure\$ or standard\$ or code\$) adj3 (formulat\$ or develop\$ or improve\$ or expand\$)).tw.
35. or/26-34
36. randomized controlled trial.pt.
37. controlled clinical trial.pt.
38. intervention studies/
39. experiment\$.tw.
40. time series.tw.
41. (pre test or pretest or posttest or post test).tw.
42. random allocation/
43. impact.tw.
44. intervention?.tw.
45. chang\$.tw.
46. evaluation studies.pt.
47. evaluat\$.tw.
48. effect\$.tw.
49. comparative study.pt.
50. or/36-49
51. Humans/
52. Animals/
53. 52 not (51 and 52)
54. 50 not 53
55. 25 and 35 and 54
56. limit 55 to yr="1990 - Current"

Academic Search Complete (EBSCOhost) search strategy

S21 S10 AND S14 AND S19 Limiters - Published Date: 19900101- 20141231

S20 S10 AND S14 AND S19

S19 S15 OR S16 OR S17 OR S18

S18 TI (evaluat*) or AB (evaluat*)

S17 TI (intervention*) or AB (intervention*)

S16 TI (control*) or AB (control*)

S15 TI (random*) or AB (random*)

S14 S11 OR S12 OR S13

S13 TI ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*)) or AB ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*))

S12 TI (course* or seminar* or workshop*) or AB (course* or seminar* or workshop*)

S11 TI (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*) or AB (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*)

S10 (S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9)

S9 TI (interest N2 (conflict or competing)) or AB (interest N2 (conflict or competing))

S8 TI ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*)) or AB ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*))

S7 TI (publication* N2 (redundant or duplicate or multiple or salami or undeserving)) or AB (publication* N2 (redundant or duplicate or multiple or salami or undeserving))

S6 TI ((publication* or publishing) N2 ethics) or AB ((publication* or publishing) N2 ethics)

S5 TI (peer review*) or AB (peer review*)

S4 TI (plagiari* or falsif*) or AB (plagiari* or falsif*)

S3 TI ((researcher* or scientist*) N2 (integrity or honest*)) or AB ((researcher* or scientist*) N2 (integrity or honest*))

S2 TI ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)) or AB ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty))

S1 TI (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#l* or fair* or objective* or impartial* or selective*)) or AB (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#l* or fair* or objective* or impartial* or selective*))

AGRICOLA (OvidSP) search strategy

1. Fraud/

2. (research adj3 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent\$ or bias\$ or fraud\$ or honest\$ or reliab?l\$ or fair\$ or objective\$ or impartial\$ or selective\$)).ti,ab.

3. ((scientific or academic) adj3 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)).ti,ab.

4. ((researcher\$ or scientist\$) adj3 (integrity or honest\$)).ti,ab.

5. (plagiari\$ or falsif\$).ti,ab.

6. peer review\$.ti,ab.

7. (data adj3 (interpretat\$ or inaccura\$ or inadequa\$ or deceptive or deceit or bias\$ or impartial or manipul\$ or misus\$ or misleading or mispresent\$ or mistreat\$ or selective or suppress\$ or fabricat\$ or fraud\$ or falsif\$ or false)).ti,ab.

8. (report\$ adj3 (selective or deceptive or deceit or misleading or inadequate or independent)).ti,ab.

9. ((publication\$ or publishing) adj3 ethics).ti,ab.

10. (bias adj3 (publication\$ or publishing or analys#s or design)).ti,ab.

11. (publication\$ adj3 (redundant or duplicate or multiple or salami or undeserving)).ti,ab.

12. (inaccura\$ adj3 citation\$).ti,ab.

13. ((author\$ or contribut\$) adj3 (undeserv\$ or ghost or guest or gift\$)).ti,ab.

14. (interest adj3 (conflict or competing)).ti,ab.

15. or/1-14

16. Professional Education/

17. Science Education/

18. Curriculum/
19. Mentoring/
20. (educat\$ or teach\$ or train\$ or motivat\$ or instruct\$ or interven\$ or promot\$ or supervis\$ or mentor\$).ti,ab.
21. (course\$ or seminar\$ or workshop\$).ti,ab.
22. Educational Policy/
23. Program Planning/
24. ((program\$ or plan\$ or policy or rule\$ or procedure\$ or standard\$ or code\$) adj3 (formulat\$ or develop\$ or improve\$ or expand\$)).ti,ab.
25. or/16-24
26. random\$.tw.
27. control\$.tw.
28. intervention\$.tw.
29. evaluat\$.tw.
30. or/26-29
31. Humans/
32. Animals/
33. 32 not (31 and 32)
34. 30 not 33
35. 15 and 25 and 34
36. limit 35 to yr="1990 - Current"

CENTRAL (OvidSP) search strategy:

1. Scientific Misconduct.kw,sh.
2. Fraud.kw,sh.
3. Ethics, Research.kw,sh.
4. (research adj3 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent\$ or bias\$ or fraud\$ or honest\$ or reliab?l\$ or fair\$ or objective\$ or impartial\$ or selective\$)).tw.
5. ((scientific or academic) adj3 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)).tw.
6. Plagiarism.kw,sh.
7. (plagiari\$ or falsif\$).tw.
8. Publication Bias.kw,sh.
9. Duplicate Publication as Topic.kw,sh.
10. Peer Review, Research.kw,sh.
11. (data adj3 (interpretat\$ or inaccura\$ or inadequa\$ or deceptive or deceit or bias\$ or impartial or manipul\$ or misus\$ or misleading or mispresent\$ or mistreat\$ or selective or suppress\$ or fabricat\$ or fraud\$ or falsif\$ or false)).tw.
12. Research Report.kw,sh.
13. (report\$ adj3 (selective or deceptive or deceit or misleading or inadequate or independent)).tw.
14. (research adj3 (underreport\$ or under-report\$)).tw.

15. ((publication\$ or publishing) adj3 ethics).tw.
16. (bias adj3 (publication\$ or publishing or analys#s or design)).tw.
17. (publication\$ adj3 (redundant or duplicate or multiple or salami or undeserving)).tw.
18. Authorship.kw,sh.
19. ((author\$ or contribut\$) adj3 (undeserv\$ or ghost or guest or gift\$)).tw.
20. Conflict of Interest.kw,sh.
21. (interest adj3 (conflict or competing)).tw.
22. or/1-21
23. Education, Professional.kw,sh.
24. Professional Training.kw,sh.
25. Teaching.kw,sh.
26. Curriculum.kw,sh.
27. Mentors.kw,sh.
28. (educat\$ or teach\$ or train\$ or motivat\$ or instruct\$ or interven\$ or promot\$ or supervis\$ or mentor\$).tw.
29. (course\$ or seminar\$ or workshop\$).tw.
30. Policy Making.kw,sh.
31. Program Development.kw,sh.
32. ((program\$ or plan\$ or policy or rule\$ or procedure\$ or standard\$ or code\$) adj3 (formulat\$ or develop\$ or improve\$ or expand\$)).tw.
33. or/23-32
34. 22 and 33
35. limit 34 to yr="1990 - Current"

CINAHL with Full Text (EBSCOhost) search strategy

- S43 (S21 AND S31 AND S41) Limiters - Published Date: 19900101- 20141231
- S42 (S21 AND S31 AND S41)
- S41 S37 NOT S40
- S40 S39 NOT ((S38 and S39))
- S39 (MH "Animals+")
- S38 (MH "Human")
- S37 S32 or S33 or S34 or S35 or S36
- S36 TI (evaluat*) or AB (evaluat*)
- S35 TI (intervention*) or AB (intervention*)
- S34 TI (control*) or AB (control*)
- S33 TI (random*) or AB (random*)
- S32 PT randomized controlled trial
- S31 S26 or S27 or S28 or S29 or S30

S30 TI ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*)) or AB ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*))

S29 (MH "Program Development+")

S28 (MH "Policy Making")

S27 TI (course* or seminar* or workshop*) or AB (course* or seminar* or workshop*)

S26 TI (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*) or AB (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*)

S25 (MH "Mentorship")

S24 (MH "Curriculum+")

S23 (MH "Teaching+")

S22 (MH "Education, Health Sciences+")

S21 S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20

S20 TI (plagiar* or falsif*) or AB (plagiar* or falsif*)

S19 TI (interest N2 (conflict or competing)) or AB (interest N2 (conflict or competing))

S18 TI ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*)) or AB ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*))

S17 (MH "Authorship")

S16 TI (innacura* N3 citation*) or AB (innacura* N3 citation*)

S15 TI (publication* N2 (redundant or duplicate or multiple or salami or undeserving)) or AB (publication* N2 (redundant or duplicate or multiple or salami or undeserving))

S14 TI (bias N2 (publication* or publishing or analys?s or design)) or AB (bias N2 (publication* or publishing or analys?s or design))

S13 TI ((publication* or publishing) N2 ethics) or AB ((publication* or publishing) N2 ethics)

S12 TI (research N2 (underreport* or under-report*)) or AB (research N2 (underreport* or under-report*))

S11 TI (report* N2 (selective or deceptive or deceit or misleading or inadequate or independent)) or AB (report* N2 (selective or deceptive or deceit or misleading or inadequate or independent))

S10 TI (data N2 (interpretat* or inaccura* or inadequa* or deceptive or deceit or bias* or impartial or manipul* or misus* or misleading or mispresent* or mistreat* or selective or suppress* or fabricat* or fraud* or falsif* or false)) or AB (data N2 (interpretat* or inaccura* or inadequa* or deceptive or deceit or bias* or impartial or manipul* or misus* or misleading or mispresent* or mistreat* or selective or suppress* or fabricat* or fraud* or falsif* or false))

S9 (MH "Peer Review")

S8 (MH "Publication Bias")

S7 (MH "Plagiarism")

S6 TI ((researcher* or scientist*) N2 (integrity or honest*)) or AB ((researcher* or scientist*) N2 (integrity or honest*))

S5 TI ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)) or AB ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty))

S4 TI (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#!* or fair* or objective* or impartial* or selective*)) or AB (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#!* or fair* or objective* or impartial* or selective*))

S3 (MH "Research Ethics+")

S2 (MH "Fraud")

S1 (MH "Scientific Misconduct")

ERIC (OvidSP) search strategy:

1. Cheating/

2. Deception/

3. (research adj3 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent\$ or bias\$ or fraud\$ or honest\$ or reliab?l\$ or fair\$ or objective\$ or impartial\$ or selective\$)).tw.

4. ((scientific or academic) adj3 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)).tw.

5. ((researcher\$ or scientist\$) adj3 (integrity or honest\$)).tw.

6. Plagiarism/

7. (plagiari\$ or falsif\$).tw.

8. Peer Evaluation/

9. (data adj3 (inaccura\$ or inadequa\$ or deceptive or deceit or bias\$ or impartial or manipul\$ or misus\$ or misleading or mispresent\$ or mistreat\$ or selective or suppress\$ or fabricat\$ or fraud\$ or falsif\$ or false)).tw.

10. Research Reports/

11. (report\$ adj3 (selective or deceptive or deceit or misleading or inadequate or independent)).tw.

12. (research adj3 (underreport\$ or under-report\$)).tw.

13. ((publication\$ or publishing) adj3 ethics).tw.

14. (bias adj3 (publication\$ or publishing or analys#s or design)).tw.

15. (publication\$ adj3 (redundant or duplicate or multiple or salami or undeserving)).tw.

16. (inaccura\$ adj3 citation\$).tw.

17. ((author\$ or contribut\$) adj3 (undeserv\$ or ghost or guest or gift\$)).tw.

18. Conflict of Interest/

19. (interest adj3 (conflict or competing)).tw.

20. or/1-19

21. exp Professional Education/

22. exp Teaching Methods/

23. exp Curriculum/

24. Mentors/

25. (educat\$ or teach\$ or train\$ or motivat\$ or instruct\$ or interven\$ or promot\$ or supervis\$ or mentor\$).tw.

26. (course\$ or seminar\$ or workshop\$).tw.

27. Educational Policy/

28. Program Development/

29. ((program\$ or plan\$ or policy or rule\$ or procedure\$ or standard\$ or code\$) adj3 (formulat\$ or develop\$ or improve\$ or expand\$)).tw.

30. or/21-29

31. random\$.tw.

Interventions to prevent misconduct and promote integrity in research and publication (Review)

32. control\$.tw.
 33. intervention\$.tw.
 34. evaluat\$.tw.
 35. or/31-34
 36. 20 and 30 and 35
 37. limit 36 to yr="1990 - Current"

GeoRef (EBSCOhost) search strategy

S20 S9 AND S13 AND S18 Limiters - Published Date: 19900101- 20141231

S19 S9 AND S13 AND S18

S18 S14 OR S15 OR S16 OR S17

S17 TI (evaluat*) or AB (evaluat*)

S16 TI (intervention*) or AB (intervention*)

S15 TI (control*) or AB (control*)

S14 TI (random*) or AB (random*)

S13 S10 OR S11 OR S12

S12 TI ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*)) or AB ((program* or plan* or policy or rule* or procedure* or standard* or code*) N2 (formulat* or develop* or improve* or expand*))

S11 TI (course* or seminar* or workshop*) or AB (course* or seminar* or workshop*)

S10 TI (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*) or AB (educat* or teach* or train* or motivat* or instruct* or interven* or promot* or supervis* or mentor*)

S9 (S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8)

S8 TI (interest N2 (conflict or competing)) or AB (interest N2 (conflict or competing))

S7 TI ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*)) or AB ((author* or contribut*) N2 (undeserv* or ghost or guest or gift*))

S6 TI (publication* N2 (redundant or duplicate or multiple or salami or undeserving)) or AB (publication* N2 (redundant or duplicate or multiple or salami or undeserving))

S5 TI (peer review*) or AB (peer review*)

S4 TI (plagiari* or falsif*) or AB (plagiari* or falsif*)

S3 TI ((researcher* or scientist*) N2 (integrity or honest*)) or AB ((researcher* or scientist*) N2 (integrity or honest*))

S2 TI ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)) or AB ((scientific or academic) N2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty))

S1 TI (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#l* or fair* or objective* or impartial* or selective*)) or AB (research N2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent* or bias* or fraud* or honest* or reliab#l* or fair* or objective* or impartial* or selective*))

LILACS (BIREME) search strategy

(MH:"Scientific Misconduct" OR "Mala Conducta Científica" OR "Má Conduta Científica" OR "Ethics in Publishing" OR "scientific fraud" OR "Fraudulent Data" OR "Research Misconduct" OR "Ética en Publicación" OR "Fraude Científica" OR "Dados Fraudulentos" OR "Negligencia en la Investigación" OR "Má Conduta Investigativa" OR "Ética em Publicação" OR "Fraude Científica" OR "Dados Fraudulentos" OR "Má Conduta em Pesquisa" OR MH:"Ethics, Research" OR "Ética en Investigación" OR "Ética em Pesquisa" OR "research ethics" OR "Ética en la

Investigación" OR "Ética na Pesquisa" OR "research integrity" OR "scientific ethics" OR "scientific integrity" OR MH:Authorship OR Autoria OR MH:"Authorship and Co-Authorship in Scientific Publications" OR "Autoría y Coautoría en la Publicación Científica" OR "Autoria e Co-Autoria na Publicação Científica") AND (MH:"Education, Professional" OR "Educación Profesional" OR "Educação Profissionalizante" OR Teaching OR Ensenanza OR Ensino OR MH:Curriculum OR Currículo OR MH:Mentors OR Mentores OR educat\$ OR educación OR educação OR teach\$ OR ensenanza OR ensino OR trai\$ OR motivat\$ OR instruct\$ OR interven\$ OR promot\$ OR supervis\$ OR mentor\$ OR course \$ OR seminar\$ OR workshop\$ OR MH:"Policy Making" OR "Formulación de Políticas" OR "Formulação de Políticas" OR MH:"Program Development" OR "Desarrollo de Programa" OR "Desenvolvimento de Programas") AND (PT:"Randomized Controlled Trial" OR "Ensaio Clínico Controlado Aleatorio" OR "Ensaio Clínico Controlado Aleatório" OR random\$ OR control\$ OR intervention\$ OR intervención OR intervençao OR evaluat\$ OR evaluación OR avaliação)

PsycINFO (OvidSP) search strategy

1. Fraud/
2. Professional Ethics/
3. (research adj2 (integrity or ethics or conduct or misconduct or malpractice or manipulation or misleading or mispresent\$ or bias\$ or fraud\$ or honest\$ or reliab?l\$ or fair\$ or impartial\$ or selective\$)).tw.
4. ((scientific or academic) adj2 (fraud or ethics or integrity or misconduct or malpractice or manipulation or honesty or dishonesty)).tw.
5. ((researcher\$ or scientist\$) adj2 (integrity or honest\$)).tw.
6. (plagiari\$ or falsif\$).tw.
7. Peer Evaluation/
8. peer review\$.tw.
9. (data adj2 (interpretat\$ or inaccura\$ or inadepua\$ or deceptive or deceit or bias\$ or impartial or manipulat\$ or misus\$ or misleading or mispresent\$ or mistreat\$ or selective or suppress\$ or fabricat\$ or fraud\$ or falsif\$ or false)).tw.
10. (report\$ adj2 (selective or deceptive or deceit or misleading or inadequate or independent)).tw.
11. (research adj2 (underreport\$ or under-report\$)).tw.
12. ((publication\$ or publishing) adj2 ethics).tw.
13. (bias adj2 (publication\$ or publishing or analys#s or design)).tw.
14. (publication\$ adj2 (redundant or duplicate or multiple or salami or undeserving)).tw.
15. (inaccura\$ adj2 citation\$).tw.
16. ((author\$ or contribut\$) adj2 (undeserv\$ or ghost or guest or gift\$)).tw.
17. Conflict of Interest/
18. (interest adj2 (conflict or competing)).tw.
19. or/1-18
20. exp Teaching/
21. exp Curriculum/
22. Mentor/
23. (educat\$ or teach\$ or train\$ or motivat\$ or instruct\$ or interven\$ or promot\$ or supervis\$ or mentor\$).tw.
24. (course\$ or seminar\$ or workshop\$).tw.
25. exp Policy Making/
26. exp Program Development/
27. ((program\$ or plan\$ or policy or rule\$ or procedur\$ or standard\$ or code\$) adj2 (formulat\$ or develop\$ or improve\$ or expand\$)).tw.

28. or/20-27
29. random\$.tw.
30. control\$.tw.
31. intervention\$.tw.
32. evaluat\$.tw.
33. or/29-32
34. 19 and 28 and 33
35. limit 34 to yr="1990 - Current"

SCOPUS search strategy

((TITLE-ABS-KEY(research W/2 (integrity OR ethics OR conduct OR misconduct OR malpractice OR manipulation OR fraud* OR honest*))) OR (TITLE-ABS-KEY((scientific OR academic) W/2 (fraud OR ethics OR integrity OR misconduct OR honesty OR dishonesty))) OR (TITLE-ABS-KEY((researcher* OR scientist*) W/2 (integrity OR honest*))) OR (TITLE-ABS-KEY((publication* OR publishing) W/2 ethics OR plagiari* OR falsif*)) OR (TITLE-ABS-KEY((author* OR contribut*) W/2 (undeserv* OR ghost OR guest OR gift*))) AND ((TITLE-ABS-KEY(educat* OR teach* OR train* OR motivat* OR instruct* OR interven* OR promot* OR supervis* OR mentor*)) OR (TITLE-ABS-KEY(course* OR seminar* OR workshop*)) OR (TITLE-ABS-KEY((program* OR plan* OR policy OR rule* OR procedure* OR standard* OR code*) W/2 (formulat* OR develop* OR improve* OR expand*))) AND ((TITLE-ABS-KEY(random*)) OR (TITLE-ABS-KEY(control*)) OR (TITLE-ABS-KEY(intervention*)) OR (TITLE-ABS-KEY(evaluat*)))

LIMIT-TO (PUBYEAR , 1990-2014)

Web of Science search strategy

#16 #15 AND #10 AND #6

#15 #14 OR #13 OR #12 OR #11

#14 TS=(evaluat*)

#13 TS=(intervention*)

#12 TS=(control*)

#11 TS=(random*)

#10 #9 OR #8 OR #7

#9 TS=((program* OR plan* OR policy OR rule* OR procedure* OR standard* OR code*) NEAR/2 (formulat* OR develop* OR improve* OR expand*))

#8 TS=(course* OR seminar* OR workshop*)

#7 TS=(educat* OR teach* OR train* OR motivat* OR instruct* OR interven* OR promot* OR supervis* OR mentor*)

#6 #5 OR #4 OR #3 OR #2 OR #1

#5 TS=((author* OR contribut*) NEAR/2 (undeserv* OR ghost OR guest OR gift*))

#4 TS=((publication* OR publishing) NEAR/2 (ethics OR plagiari* OR falsif*))

#3 TS=((researcher* OR scientist*) NEAR/2 (integrity OR honest*))

#2 TS=((scientific OR academic) NEAR/2 (fraud OR ethics OR integrity OR misconduct OR honesty OR dishonesty))

#1 TS=(research NEAR/2 (integrity OR ethics OR conduct OR misconduct OR malpractice OR manipulation OR fraud* OR honest*))

Indexes=SCI-EXPANDED, SSCI, A&HCI

Timespan=1990-2013, 2013-2014

CONTRIBUTIONS OF AUTHORS

- Conceiving, designing and co-ordinating the review: Ana Marusic (AM), Elizabeth Wager (EW), Hannah Rothstein (HRR), Dario Sambunjak (DS).
- Designing search strategies and undertaking searches: AM, AU, EW.
- Screening search results and retrieved papers against inclusion criteria: AM, EW.
- Appraising quality of papers: AM, EW, DS, HRR, AU.
- Extracting data from papers: AM, EW.
- Writing to authors of papers for additional information: AM, EW.
- Data management for the review and entering data into RevMan ([RevMan 2014](#)): AM, DS.
- Analysis and interpretation of data: AM, EW, DS, HRR, AU.
- Providing research perspective: AM, EW.
- Writing the review: AM, EW, DS, HRR, AU.
- Providing general advice on the review: AM, EW, DS.
- Performing previous work that was the foundation of the current review: AM, EW.

DECLARATIONS OF INTEREST

This review will be used by some of the authors as part of other research projects. AM reports that she was a co-author on two studies included in the review. EW does not report conflict of interest. HRR is one of the authors of the Comprehensive Meta Analysis (CMA) software. AU does not report conflict of interest. DS reports that he was a co-author on one of the studies included in the review. Measures were taken within the review team to prevent biased handling of their own studies by the review authors.

SOURCES OF SUPPORT

Internal sources

- Ana Marusic, Croatia.
Employed by the University of Split, Croatia
- Ana Utrobicic, Croatia.
Employed by the University of Split, Croatia
- Hannah R. Rothstein, USA.
Employed by Baruch College, New York, USA.

External sources

- Croatian Science Foundation, Croatia.
This review was supported in part by the grant No. IP-2014-09-7672 ('Professionalism in Health') to A. Marušić.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

There were no major changes in relation to the published protocol ([Marusic 2013](#)), except that quantitative synthesis was not possible due to the heterogeneity of interventions and outcomes. The protocol planned for the inclusion of controlled before-and-after studies, where we included not only the studies with before-and-after measurement for the control and intervention groups but also the studies that had before-and-after measurement for the intervention group and a single ('postintervention') measurement for the control group. This was addressed as issue of concern in our 'Risk of bias' assessment. We excluded controlled studies that reported outcomes measured only after the intervention, without 'before-intervention' measurement in the intervention groups.

There were also some changes in the search strategy, which we adapted to better capture the topic of the search and decrease the number of retrieved but irrelevant items.

INDEX TERMS

Medical Subject Headings (MeSH)

*Plagiarism; Attitude; Biomedical Research [*ethics]; Controlled Before-After Studies [ethics] [standards]; Controlled Clinical Trials as Topic [ethics] [standards]; Publishing [ethics] [standards]; Randomized Controlled Trials as Topic [ethics] [standards]; Research Personnel [*ethics] [standards]; Scientific Misconduct [*ethics]

MeSH check words

Humans