

Critically appraising research



Copyright information:

Critically appraising research: is published by the British Association for Counselling and Psychotherapy, BACP House, 15 St John's Business Park, Lutterworth, Leicestershire, LE17 4HB.

T: 01455 883300 **E:** bacp@bacp.co.uk **www.bacp.co.uk**

BACP is the largest professional organisation for counselling and psychotherapy in the UK, is a company limited by guarantee 2175320 in England and Wales, and a registered charity, 298361.

Copyright © 2024 British Association for Counselling and Psychotherapy.
All rights reserved.

Permission is granted to reproduce for personal and educational use only.
Commercial copying, hiring and lending are prohibited.

BACP and the BACP logo are registered trade marks of BACP.

Design by BACP

It should be noted that this resource offers broad guidance, which sets out professional good practice, but should not be substituted for legal and other professional advice, including supervision, applicable to your particular circumstances

About the authors

Written by

Chris Evans and Jo-anne Carlyle

Chris Evans

I was science focused as a teenager in a family of teachers and humanities experts and trained in medicine moving into Psychiatry after 18 months of hospital medicine. I went on to trainings in group and individual analytic therapies and systemic psychotherapies. From 1986 to 2016 I worked 50/50 in clinical and research in the NHS, from community to high secure care. Earlier publications with Jo-anne looked at forensic psychotherapy research and also described "containing containers": working in high secure settings.

I long ago renounced my teenage "quant" rebellion and have very diverse interests resulting in 157 peer-reviewed publications to date. My biggest research programme has been CORE Clinical Outcomes in Routine Evaluation system: a set of instruments and philosophy for measuring change in therapies. I co-led translations of CORE measures into over 30 languages since CORE launched in 1998. Most of my research has been quantitative but and I am passionate about the importance of qualitative data and methods.

The book Outcome measures and evaluation in counselling and psychotherapy (2021) was another collaboration with Jo-anne, extending CORE. We argue for routine change measurement but challenge the overselling of questionnaire change data hoping to give practitioners and managers tools to use measures wisely. Since 1998 my main research collaboration has been with Professor Clara Paz from Ecuador. Clara and I work to create useful evidence about therapy change outside the global north "factory model", as well as rethinking measure translation and adaptation. I also work on novel psychometric methods for repeated measures and individual change and on "rigorous idiography": exploring validity in purely idiographic data. I have created free tools around our book and therapy measurement: a glossary of over 260 terms (and increasing), a collection of more detailed explanatory articles: Rblog, and a growing collection of online apps.

About the authors (continued)

Jo-anne Carlyle

I am a Clinical and Forensic Psychologist, Organisational Consultant and Psychoanalytic Psychotherapist and director of [PSYCTC.COM](https://www.psyctc.com). I have 35 years of experience of clinical practice, and drawing on my work in mental health and humanitarianism I am now focusing on the restrictions that a Western Global North clinical approach and pedagogy place on more culturally aligned work for mental health cross-culturally. Linked to this, I hold a conviction that research and evaluation are essential, but that when they become reductionistic they fail the patient, society, the treatment journey and our genuine capacity to make significant shifts in practice.

I previously worked at the Tavistock and Portman NHS Trust, Exeter University, South London and Maudsley NHS Trust, and at Broadmoor and Rampton Hospitals. My work has been driven by a commitment to social inclusion and social change especially in the light of global inequities. My current work often uses experimental and group approaches to think about the impact of colonialism and patriarchy on organisational structures and leadership and works to explore other models of leadership including shared, lateral and distributed leadership.

I have worked in India, Ecuador, Canada and Lithuania as well as online globally. I sit as a Non-Executive Director on education and mental health boards. I was co-creator of the Tavistock Adult Depression Study, and my publications span research, methodology and clinical interests and I am co-author of a Sage Commissioned book (with Chris Evans) "Outcome Measures and Evaluation in Counselling and Psychotherapy" (2021).

Contents

About the authors	3
Chris Evans	3
Jo-anne Carlyle	4
What is critical appraisal applied to research?	6
Why is it important to read research with a critical eye?	7
Read the paper in context	7
Locate yourself and the paper	8
Applying critical appraisal through the anatomy of a paper	10
Journal papers and other formats	10
Varieties of papers	12
Title	13
Authors	14
Abstract	15
Keywords	16
Citations and references	16
Introduction	17
Methods	18
Results/Findings	20
Discussion [& Conclusion]	22
Collateral information about a paper	24
Summary	25
Other resources	26
Possibly useful further reading	27

What is critical appraisal applied to research?

“Research is formalized curiosity. It is poking and prying with a purpose.”

Zora Neale Hurston

We start with this quote because it humanises “research” and locates it as something that we all do! The quote also contains information that is both explicit and implicit. At the explicit level it is a definition of research – linking curiosity, poking and prying with the energy of exploration and discovery and the importance of systematic observation (‘formalised’). At the contextual level, there are additional data points. It is a quote by an African-American woman which alerts us to remember that research is not an objective truth within a positivist framework, but something shaped by social and cultural norms and values that means that the “formalized curiosity” of our research bank is, inevitably, biased, for example, underrepresenting women and people of colour. Zora Neale Hurston's output included anthropology, folkloric studies, fiction and journalism: crossing traditional academic disciplines but always drawing on research.

The reader of research requires as much of a critically engaged and inquiring mind as the authors of a paper or project. Research is relational: in its communication, dissemination, and digestion.

In this BACP fact sheet, we introduce some key concepts about critical appraisal and then use the anatomy of a traditional research paper in the therapy/counselling field to illustrate how those ideas can be worked with in practice. The relatively formalised structure of a traditional paper is not so strictly applied to other research reports in our area (see below) but using it allows us to examine, in context, key aspects of critical appraisal: identifying the authors' questions, and yours; checking ethics; looking for narrow focus, bias, reflexivity; and reaching an overall appraisal through all this.

Why is it important to read research with a critical eye?

The art of good research is to be open to the unexpected, what is not immediately obvious, what cannot immediately be seen. As we go through the process of “reading” a paper, we will try and weave in some of these more lateral or “associative” frameworks that are necessary for knowledge development; these are vital to enable us to be critical: to notice and challenge the author’s assumptions but also, *our own*.

So, a critical eye should be open to surprise, and it must try to weigh the evidential claims of the work. Just because some work has been published, it does not mean the findings are definitive or generalisable, in fact, in the therapeutic field they can never simply be either of these things. The art of critical appraisal is all about what to look for in the paper, how to interpret the results, how to shape your own view of the claims the authors are making.

Read the paper in context

No work is unbiased, and all authors transmit some of their own beliefs, whether consciously or tacitly. We might expect the conscious ones to be stated in the epistemological position and choice of methodology, but it is the job of the reader/reviewer to be curious about elements that may not have been reflectively explored. Good critical appraisal extends the authors' own examination of their work. It is not about setting out to attack or destroy what the authors say, but about constructive criticism of what was said in a manner that aims to add value, whether just for yourself or also for others.

The biases we all bring are not just personal, they also always come from our societal, historical location as our opening quote underlines. The context of any work will have had an important influence on the justification for the research, how it was resourced, how the research question has been defined, and how it is communicated. We explore this further in the “anatomy of a paper” below. Any reader is also encountering the report within a context and should consider that.

Locate yourself and the paper

When starting to read a paper, pause for a moment and ask yourself why you are reading it at this point in time and what it is that you want to get from it. A student reader can have a very general wish to learn, but sharpen this by asking the questions "why has this paper been chosen as required reading?" and "how does the authors' context relate to mine?". A practitioner might read for CPD or a desire to keep abreast of new developments in the field. Other interests might be wanting to learn about a new clinical development, or about work with a particular group of clients, or exploring a new theoretical or clinical orientation. Locating yourself in relation to the paper means trying to understand something of the history of the concepts in the paper, how they were first written about and developed and is helped by imagining the social and cultural context in which the paper was born and developed and where you site yourself in relation to that. At any point in history there are many factors determining how ideas are translated into published research.

Once you have some idea what you want from the paper ask yourself how this relates to what you think the authors are trying to achieve? Some of their intentions may be explicit in the title though that is necessarily a very condensed statement of aims and sometimes some aims really only emerge in the discussion section of a paper. Aims may be very tight, for example a single binary question ("is this approach more effective than that one?") or they may be very open and exploratory. For many reasons authors generally aim high and often claim more than is actually able to be determined from the information available. Understanding authors' aims enables us to judge whether those aims are met.

Critical appraisal recognises that research papers are NOT simple statements of fact but that they are communications that utilise interpretations of findings in other literature and themselves require interpretation. Papers, like all human communications, have at least three levels:

- The overt content.
- The claims about the evidential value of that content, generally a mix of overt and covert communication.
- What is not said – this refers both to the inevitable choices a researcher has to make to exclude information and lines of enquiry, but also to their equally inevitable unstated motivations and choices – for example, where did the sample come from? A student sample in Scotland will not generalise simply to one in Cape Town, Quito or Vilnius.

It can be helpful to read a paper through twice, the first time to get a grasp of the overt structure, the style and the content; then reading it again, skimming to get distance, or digging deeper, and perhaps turning to other resources to enrich your understanding. This second read should pick up not just the overt, but also the covert claims of evidential value and help you identify your opinion of the paper and its reported findings.

Authors need readers and readers need authors! Writing is active, performative; readers receptive, seemingly passive. However, critical appraisal is very active work and readers' collective critical appraisals determine the paper's impact. Authors' ideas do not come out of nowhere: they will have come from discussions with colleagues, experiences with clients and reading other research: to write a paper is to be a node between that collective generativity and the readership. Authors assert ownership by presenting the work to the community. We hope this fact sheet helps you own your active part and encourages you to be a constructively critical reader and hence bringing your contribution to the development of therapies and of research processes.

Applying critical appraisal through the anatomy of a paper

We started with the quote "*Research is formalized curiosity. ...*" from Zora Neale Hurston. She has written pure fiction and journalism as well as more academic work and that should not be surprising as a good research report should tell a story, and tell it well, honestly and clearly. A research story is constrained by the evidence, the data, and shaped by a traditional structure which, if used well, makes for clarity in both writing up research, and in appraising reports. The archetypal structure is that of a research paper but first we look at the wider variety of research reports.

Journal papers and other formats

Typically, research is published as papers in journals but there are other formats and other locations with a very rough hierarchy of likely evidential value.

1.	Peer-reviewed, indexed journals. "Peer-reviewed" means published papers have been assessed by at least one, generally two or more "peers", usually other researchers. The reviewers recommend actions to the editor of the journal, typical options are: reject, recommend minor or major changes, or accept. Thus peer-reviewed papers have already had some critical appraisal, but it is no guarantee of evidential value and certainly the reviewers' interests in the paper may not be yours. "Indexed" means that the contents of the journal are indexed in one or more reference databases which improves accessibility. Whereas publication in an indexed, peer-reviewed journal used to be a possible indicator of quality, that world is controlled to a great extent by a very few large publishers and is now acknowledging issues with "predatory" journals that invite submissions of minor rewrites of other papers, using essentially toothless peer-review.
2.	Other journals, newsletters and house publications. These are generally neither peer-reviewed nor indexed. However, as indicated above, this does not necessarily mean that the content is of low quality.
3.	Previews available on the internet. Open sharing pre-publication or pre-submission versions of papers has been common in the "hard sciences" for decades and they are now being used in our field. Be wary of any that lack a clear statement of their developmental stage, e.g. "early draft for discussion" or "final version accepted for publication in ...". Such work includes excellent online resources in some blogs and in repositories guaranteed to remain publicly accessible for the foreseeable future.

4.	Books or book sections/chapters. These vary greatly in quality. Some include ground-breaking conceptual and theoretical developments and others are, well, rubbish! With the move to E-books and publishing, the costs of producing books have dropped and increasingly there may be no quality control over what gets published.
5.	Student theses: research doctorates ("PhD" or "DPhil"), professional doctorates (e.g. clinical or counselling psychology and psychotherapy and counselling trainings) and Masters. These vary enormously in quality ranging from those that are as good as peer-reviewed papers to much weaker ones. Sadly, few of the good theses are converted to indexed, accessible papers.
6.	Other documents ("grey literature"). This might include reports by governmental or non-governmental bodies, as well as resources available on the internet, often freely available. These encompass the hugely historically important archives such as census data and repositories of personal documents and include vital histories of groups disenfranchised from more formal and common sharing. Increasingly important are blogs and other text and now video presentations, vlogs, dialogues, tutorials and commentaries. Whilst the historical archives are generally of clear quality, the ever-expanding volume of other material ranges from really excellent pieces of work to fake news. It is probably safest to regard the best of these, such as Wikipedia or TED talks, as useful adjuncts (rather than core sources) to the main literature of peer-reviewed papers.
7.	"Bottom drawer/filing cabinet" reports (these terms are used in "systematic reviewing" not just inventions of ours). Much research, perhaps the majority, sits in bottom drawers (or now in electronic storage somewhere never accessed) and never makes it into the public domain. Some remains there due to poor quality, however, publication biases mean that valuable and even important work often stays in bottom drawers. Work that does not produce "significant" findings has been less likely to get accepted into journals and this publication bias skews the literature and means that key information is lost to the research database. Student theses and dissertations that do not make it into the indexing of university outputs (partly determined by a hierarchy of universities and countries) is hidden to us. Data collected by clinicians and other practitioners and clinical services data collection, including audits rarely get to be shared and accessible. Even professional researchers can find that they do not have the time, resources, confidence or knowledge to transform completed work into publications or other outputs. This is even more the case where work is cross disciplinary. One particular challenge is the huge volume of work in languages other than English that does not get read by most Global North researchers.

Varieties of papers

The archetypal paper reports empirical data, usually this is new data but sometimes a paper is a re-analysis. However, as well as these kinds of traditional research reports, there are probably four further categories to know about as they can help critical reading: [pre-]registrations, protocol papers, different review papers and the "response literature".

Pre-registrations. These involve the researchers registering a plan of what they are going to do before they start collecting data (hence "pre-"). This started as a corrective to the bottom drawer problem and the publication bias toward significant findings: now respectable (non-predatory) journals simply will not accept papers about controlled trials without prior registration in a publicly accessible register such as clinicaltrials.gov. Registration may involve no quality checks at all or some minimal checking but has no formal peer-review. Some journals now offer a guarantee, subject to some quality checking, that any study with pre-registration will be published regardless of the significance of the findings. Pre-registration records help critical appraisal making it easy to see if the paper stuck to the original plan. The principle that pre-registration is a useful indication of evidential value is stronger, and currently more common, in quantitative work than for qualitative but, with suitable adaptation, could be useful in the qualitative world too.

Protocol papers. These are like pre-registrations but are more detailed with the formal anatomy of a paper and published in journals rather than online registers. Protocol papers give the whole protocol for a piece of work and are peer-reviewed. A protocol paper can be useful where a single research programme may lead to a number of separate papers reporting different facets of, or projects within, the programme and can save a lot of duplication in introductions and methods sections of the eventual papers. That a paper was preceded by a protocol paper is some indicator of evidential value though again it is more common for quantitative than qualitative work.

Review papers. Traditional literature review papers could be extremely selective and subjective. That led to the development of "systematic reviews" which have formalised ways aiming to identify and summarise all the existing literature about a topic. This has gone hand in hand with the development of meta-analysis: statistical analyses of findings across many individual papers and reports, which has become a branch of statistics in its own right. Although still very much a field built on the quantitative literature there are also developments of new methods of reviewing qualitative work.

"Response literature". This includes editorials, letters to journals (though many journals do not have correspondence), online comments, editorial pieces accompanying a paper and full papers which are responses to earlier work or re-analyses of the data. When reading a paper, it can be useful to search on the internet (see "Collateral information ..." below) to see if the paper has elicited any specific responses and to see how much it has been cited in later papers.

Traditional anatomy of a paper

Critical appraisal can be usefully organised by the typical components of a paper

- Title
- Author names, and varying amounts of additional information about them
- Abstract/summary
- Introduction
- Methods
- Results
- [Reflexivity (almost always present in qualitative papers though sometimes within the discussion section)]
- Discussion
- [Conclusion (not always added to a discussion)]
- References
- Other information

There are many variations on this with sections sometimes omitted or others included. Quantitative work is more likely to stick tightly to that sequence, whereas something like Methods may not be a distinct section in some qualitative papers.

Title

A good title says what the paper is trying to do and captures your attention and imagination. There are inevitable limitations to this: describing a complex study or idea in one sentence, occasionally two, is a challenge. However, when you have read a paper go back and see if you think the title was well chosen. If you think it was not, you may see how the authors' views of their work differs from yours.

Authors

Details about authors remind us that papers are human, personal products and that the people who did the work and created the paper do so in organisational, sociopolitical contexts that may be very important in helping us understand the backstory behind the paper.

As well as authors' names a paper will usually tell you about their qualifications and institutional locations. Some journals encourage authors to give a paragraph about themselves and their history. Any good paper will have contact details for at least one author (the "corresponding author", not always the first). Increasingly journals expect authors to give their ORCID ID: a persistent digital identifier which points to publicly accessible information about the person (ours are <https://orcid.org/0000-0003-4981-4438> and <https://orcid.org/0000-0002-4197-4202>). Those are useful as Email addresses and organisational locations change. ORCID allows you to find out about an author and, though some of the information is chosen by the author, much, e.g. the publications listed, is curated by ORCID and will be correct (if not necessarily complete).

Some journals ask authors for photos. Although this can give a human face to the research, photos – as with institutional locations and other information require us to notice our assumptions and biases and be curious as to how this information may change our reading of the paper. Does this reinforce or help us to challenge some of our stereotypes and prejudices? Should we have photos at all?

The institutional locations are useful and should be where the authors were when they did the work. As you read, notice whether all the authors are from one institution or widely spread? Are the authors from one country or many? There may be a statement of who did what in the paper and sometimes, particularly in quantitative papers a formal statement of which authors take responsibility, "are guarantors for", what parts of the research. This can be helpful to understand more about the gestation of the paper and it is a small corrective to the tendency for senior people to be authors on many papers despite having contributed very little to them.

Linked to authors, sometimes, often at the end of the paper, there are often "declarations of interests", a phrase increasingly replacing "conflicts of interest" to put the stress of transparency about interests. There may be acknowledgements to people or organisations who helped in the preparation of the finding but who are not authors, and there may be a formal statement of how the work was funded. Somewhere in the paper, often in the methods section but sometimes at the end, there should be a clear statement of the ethical position and of having obtained approval from an ethics committee, identifying the committee and often the actual application number and date of approval.

Abstract

Abstracts (sometimes "summary") should capture the essence of all aspects of the paper and often determine whether or not readers will read the rest of the paper. A misleading abstract is an indicator of poor quality.

With internet access a paper's title, keywords, authors' names and the abstract are usually freely available. Paradoxically this wider availability of abstracts has both positive and negative outcomes: a huge gain that more people can access research literature but the need to sell the paper to readers probably increases the risk that abstracts may oversell the work. Increasingly journals also freely provide the bibliography of a paper which may influence whether you choose to read it. We will come back to references below.

Abstracts may be unstructured or structured, i.e. broken down by headings that are required by the journal (usually the same headings of the paper itself). It is increasingly common for papers to have short bullet points either with the abstract or at the end of the paper. Typical bullet points, usually set by the journal are:

- "What was already known"; "what this paper adds/changes"; "implications" (a useful and challenging way to think about any paper).
- "Key findings"; "limitations".
- "Implications for practitioners"; "implications for further research".

Some journals, particularly in the more "medical" edges of our field, may also require a lay readable abstract or summary and some journals encourage short video abstracts.

Structured or unstructured, textual or video, an abstract has to summarise the rest of the paper in a fraction of the words: typically 5%. A good abstract summarises fairly but many will overstate what was found. If the topic of the paper is important to you there is no substitute for reading the whole paper and, when you finish, reviewing whether you think the abstract oversold the findings is/as one way to judge the evidential quality of the paper.

Keywords

These help the indexing of papers into the literature databases. Like abstracts, they can underline what the authors think is important in the paper.

Citations and references

Citations to papers can come up in any of the main sections of a paper. They should always map correctly to entries in the reference list at the end of the paper. Different disciplines cite with different styles: numbers, footnotes and various ways of putting authors' names and the year of the paper in the text so people who read across disciplines will need to get used to different styles.

What authors choose to cite can be useful indicators of quality and bias, though appraising this take practice and experience of the field. The key is pay attention to citations while trying not to get distracted by them! Good citing/referencing acknowledges influences, creates a network of connections between papers, ideas and findings, and it locates the paper in the wider field around it, enabling the reader to explore that network. A key indicator of the quality of a paper is what it contributes to the wider network of work and how appropriately and transparently it positions itself in the wider network.

Authors can use citation for poor reasons: to try and impress the reader with how much they have read and perhaps how esoteric their reading is; to show they are "in the right club" by acknowledging powerful actors in the field; or adding references that are not very relevant in the hope of shoring up a weak point in a paper. A clue to citing for kudos rather than network location is giving a generic reference despite making a specific claim. With time it gets quite easy to see such inappropriate citing. A slightly more complex issue is where, as a reader, you feel there should have been a citation to a specific claim but none was offered. Reasons to cite references differ across the parts of a typical paper so we will touch on that below.

Introduction

This is the opening narrative about the paper. It should locate the paper in relation to existing ideas, findings and theory. It should give the aims of the paper and discuss whether these are tight or broad.

Papers vary in how extensively they summarise the existing literature, but an introduction should not become a systematic review. It should provide the background to the work being presented, and what the authors considered the main contributory theory and practice in the area. Too extensive a review can distract from what is potentially new or innovative in the paper at hand, or can be used to cover up when the substance of the paper is thin. We believe that all papers, qualitative, quantitative, mixed, hybrid or blended, should have a brief epistemological position declared in the introduction, i.e. that the authors should say how they evaluate the evidential claims in their work. A simple statement of the authors' position helps us know whether we will be reading the paper from a different position to that of the authors. Sometimes introductions are used to dive extensively into the epistemological position of the authors and take pages that might have been used more effectively for the findings or discussion, or for other work! Sadly, in the overall majority of quantitative papers there is no epistemological position statement and this creates a rather naïve assumption of an "empirical positivist" or "the facts speak for themselves" stance, whether or not the authors really hold that.

As well as an epistemological stance all papers should also have a methodological position which provides a bridge from the aims and the epistemological position to the actual methods. Methodology is an overall approach that generates the procedures and methods; methods are the actual tools used within that methodology. The methodology should be clear in the introduction of the paper and give the rationale for what is going to be done, and why. In quantitative papers the methodology is typically a statistical position: descriptive, exploratory, estimating or hypothesis testing. In qualitative work the range of methodological positions is far wider and probably more linked to the epistemological position, for example therapy dialogue may be inspected using content, thematic, conversation or discourse analyses, to name just a few. Papers should cite enough introductory or overview work to enable the reader to know what else to read if not familiar with the particular approach.

Methods

One task of the methods section is to restate the aims from the introduction more precisely, specifically to set out the map to the method. In the "hard science" tradition, the only tasks of the methods section were to describe the specific aims of the study and to spell out the method/s used, in order to allow replication of the work. Therapy research remains influenced by that tradition with the idea of a "scientist-practitioner" prevalent in clinical psychology and associated professions, such that the "enough for replication" model of the methods section is still dominant. That model has utility for much quantitative work where the ideas of replicability and generalisability make sense. However, that model can become reductive in therapy research where replicability across clients may be meaningless for things that matter. Rather than just enabling replication, a good methods section should build, in a methodologically appropriate way, on the "why" in the introduction so the reader can understand what was done and how, and what information to expect in the results section.

Facts never simply speak for themselves in any empirical therapy research and the methods section is about how the authors draw inferences from their data. However, the word "inference" is used differently across the dichotomisation of research into quantitative and qualitative. A good qualitative methods section will explain in some detail who did what and how, what level of inference was used and will generally have a "reflexivity", "personal reflection" component. Authors should acknowledge their influence on the process of creating and interpreting findings, and thus help us appraise their work and hold onto the recognition that simple replicability may not be a sensible aspiration for some methods. Sadly, this does not stop some qualitative research being criticised for non-replicability.

By contrast, a quantitative hypothesis testing paper should reframe the aims from introduction into "null hypotheses" and "alternative hypotheses". In this model, statistical "inference testing" is used to choose between those hypotheses and deem things "statistically significant" or not. Be wary of papers that use the language of statistical significance in the results section if they have not at least given null hypotheses in the methods section: the work may have been done well, but the authors did not share part of the logic of their approach and have gone beyond what their findings really mean. As noted above, hypothesis testing is only one of a number of statistical approaches, however the idea that the only function of research is to answer binary options with statistical significance persists.

A curious part of the methods section for many papers is that the participants are described here rather than in the results section. This is an American Psychology Association (APA) requirement and stems from the laboratory experiment tradition where the number of participants, be they rats, pigeons or humans, is all really that matters and their individuality, even their demographics, assumed to be unimportant. APA rules also require that if questionnaires were used, observed internal reliability (the extent to which they are or are not subject to random influences, see Evans & Carlyle, 2021, chapter 5), should be given even though, as with the demographic and other descriptors of the participants, they are clearly results.

When reading the methods section of any paper, watch out for "blinding with science": attempts to impress rather than to explain. Sometimes a methods section can be so technical as to baffle someone not already familiar with the methods. Any profession will have technical terms, but a good methods section should try to describe these in clear and straightforward terms that promote inclusivity and encourage understanding. Citations in the methods section should be to the methods used and should enable the reader to find out more about them if they are not absolutely clear from the text.

Results/Findings

This section presents the findings that emerged from the use of the methods. A good results section does not go back to explain the methods nor reach into the implications of the findings. We prefer "findings" to "results" as the latter implies more of a cause/effect revelation of truths by the methods; "findings" recognises that what emerges is part of a complex system and needs to be embedded in its context to be understood and interpreted.

As noted, different methods involve different forms of inference and interpretation. Generally qualitative research involves overt inference and it is given in the results section, whereas quantitative research, beyond using formal statistical "tests", tends to move interpretation to the discussion section. Sometimes something happened to make the original design impossible to follow or execute as planned, for example, a service reorganisation or therapist sickness. In such situations the results section should contain a narrative about this though it will certainly be returned to in the discussion. The results section of a paper generally requires few if any citations.

Data are analysed into findings: that is, they are always summarised and simplified, and analysis involves various procedures introduced in the methods section. It can be helpful to think of this in the active verb form: "analysing" not "analyses", "tabulating" not "tables" and "plotting" not "plots". This reminds us that presentation of findings came from decisions made by the authors and our critical appraisal should consider those choices: were they the best ways to present things? What does the presentation amplify? What does it hide?

Tabulation takes many forms in different traditions but should always help make the data, the findings, more transparent and meaningful to the reader. In qualitative work there are the different traditions of transcribing recorded conversation, ways to tabulate excerpts of speech, techniques that produce theme maps or word storms for word level content analysis. Tables should have sufficient narrative, often in footnotes as well as accompanying text, that there should be no ambiguity in what they are saying. In quantitative work plotting is a common and again needs some narrative in the titles and subtitles, footnotes and in the accompanying text. In some quantitative work, statistical procedures may have been used, often alongside tabulation and/or plotting, to distil the data.

Statistical procedures range from using a mean, through to extremely complex procedures. A growing challenge in our field is that the newer methods are hugely powerful and can offer real advantages over earlier methods. For example, "multi-level modelling" methods which have exploded over the last forty years improve routine outcome research as they allow us to separate the effects of client, therapist and service differences. Client variables might include gender, age, reluctance to use therapy; therapist variables might include gender again, but also levels of experience, modality, profession and trainings; service variables include size, location and referral networks. However, there is a growing problem that such methods may be incomprehensible to most readers. More seriously still, authors themselves may not fully understand the procedures they have used and not recognise that they may have vulnerabilities. In principle, peer reviewers should catch this, but many reviewers are not themselves sufficiently familiar with the methods to know the vulnerabilities. There are no magic answers to this but if you do not understand something in a paper, keep open the possibility that it may not actually be correct.

Whether you are confident with statistics or not your critical appraisal also looks for some narrative in the paper about possible limitations of the procedures used – this may be in the methods section or the discussion. In assessing a paper, try to distinguish those that genuinely explore limitations and try to assess their possible impact on the findings with those that are more virtue signalling (see below) and do not fully or thoughtfully describe the possible impacts.

Where work is a replication and/or extension of earlier research, it is generally best if the new summary findings can be directly compared with the earlier findings, tabulating or plotting the new findings against the old ones to help comparison. A final note, it is sadly still rather rare in papers we read to see formal attempts to compare say the composition of a participant group with existing referential data: a weakness in our field.

Discussion [& Conclusion]

When we reach this section we have been taken through the background, aims and aspirations, epistemological position and methodology in the introduction; then the procedures used in the methods and the findings in the results. The discussion should now be exactly that: a discussion of the findings. Some journals expect separate discussion and conclusions sections, otherwise the last paragraph or two of the discussion is often concluding.

This section should have at least a reflexive note on how the authors' positions may have affected the findings (though sadly this is almost never seen in quantitative work). It can be particularly informative if authors say if they were surprised by anything.

No study can be perfect in its execution and the more "naturalistic" the study, almost certainly the more the problems and issues. Authors should not pretend the work is perfect and criticising minor or inevitable imperfections is not good critical reflection on the authors' part. Discussion of weaknesses is vital to evaluate the findings and their implications and to improve research methods. However, it is not uncommon for authors to turn reflexive assessment of weaknesses unhelpful virtue signalling. E.g.

"One weakness is the non-random nature of the sample, however, with over 500 participants, findings should be fairly robust to this."

Which is not necessarily true at all and could be parodied as:

"Look we know it's an internet volunteer sample snowballed off our therapy association's Email list and social media and that we have no idea who participated and who didn't, and it could be that most of the participants weren't even therapists but people who found the survey on social media, but we're doing this little declaration of limitations to show you what good, honest people we are so you can put this away between friends and trust us that the findings will generalise to any sample of therapists anywhere in the world."

A worthwhile discussion of limitations does not just list them but attempts to estimate how much impact they might have had on implications to be drawn from the findings. It will go on to suggest what new work, or what reanalyses of the data, might address the issues. Papers that end with an assertion simply that more research is needed without saying what are just wasting words!

As well as exploration of limitations, a good discussion returns to the introduction and looks, in a balanced way, at how far the findings have addressed the aims. Where the work was a very direct replication and/or extension of a previous study the comparison of the new with the older findings will have been in the results section and the discussion will be mainly about the implications of differences or similarities with that original research. Where work was not replication/extension work the discussion should review how the findings sit against the existing literature. Unless everything went completely to plan the discussion should look at what analyses were preplanned and what may have been emergent or reactive.

Unless the procedures were derailed by unexpected events, or findings were quite unexpected, it is generally expected that citations in the discussion will not introduce new theory or aims nor introduce new references to existing work.

Collateral information about a paper

One key theme in this fact sheet is that papers are not presenting unarguable facts, they are pieces in a network of theories, methods and findings. Your reading of the paper tries to locate it in that network, and, as you are reading it, you are extending your own your own mental network of therapy. Appraisal draws on information in the paper as we have shown, however this can be helped greatly by finding collateral information about the paper.

As noted, most peer-reviewed journals are indexed, that is all papers published in them are entered into one or more of the huge bibliographic databases. Unfortunately, not all these indices are freely accessible to private individuals, an exception, though medically oriented, is PubMed (<https://pub.med.ncbi.nlm.nih.gov/>). Some of the interfaces to these databases allow you to see all the papers, in the same database, that have cited the paper you are reading, and this can be useful and can give a useful map of the academic, Global North, influence of a paper.

We have mentioned that Wikipedia and TED talks can help with background (used carefully), and that ORCID can give you curated lists of authors' publications and a bit more information about the authors. ResearchGate (<https://www.researchgate.net/>) is not curated as ORCID is, so it gives authors more control over what they put there, but it can be very informative and give you the chance to ask the author directly for a copy of the paper if it's not freely available. On the whole, the other similar systems, including Google Scholar are less useful in our view.

A last good way to get collateral information when appraising a paper is by reading one or more papers the authors have cited that seem particularly likely to develop your understanding of their position about the topic area or about methodology or simply the methods they used.

Summary

We started with this quote and we will finish with it:

“Research is formalized curiosity. It is poking and prying with a purpose.”

Critical appraisal is not a mechanical task, it is a process that gets easier with practice. It is not about deciding simply whether a paper is right or wrong, it is about your assessment of the evidential value of the paper: both in relation to your own learning needs but also in terms of the wider professional community. It is not about simply accepting or rejecting what the authors said, but about assessing what they said. It is not about agreeing with or disagreeing with facts presented to you, it is about deciding what you think the strengths and weaknesses are in the story the authors told you about their “formalized curiosity”, their aims, their participants, their data, their analyses, and their conclusions. You complement it by poking and prying at it with your own aims, epistemology and your own curiosity. The more therapists learn to be constructively critical, curious, appraising readers, the better our field will grow!

Other resources:

BACP research resources

Our recent book (Evans, C. & Carlyle, J. (2021). Outcome measures and evaluation in counselling and psychotherapy (1st ed.). SAGE Publishing) is about change/outcome measurement, only one small part of therapy research, however pointers about critical appraisal were central to our book so it and the online glossary, <https://ombook.psyctc.org/glossary> and supporting pages generally, <https://ombook.psyctc.org/book> should be useful. Similarly, Chris's pages about the CORE system, <https://www.coresystemtrust.org.uk/> are about the CORE measures so may be useful if a paper uses a CORE measure, though the principles will apply for any change/outcome measure. Chris's PSYCTC.org pages (<https://www.psyctc.org/psyctc/>) have more general research support.

SCOPUS

<https://www.scopus.com/home.uri>

PubMed

<https://pubmed.ncbi.nlm.nih.gov>

EBMLive site – that looks at the biases that are inherent in research.

<https://ebmlive.org/reasons/>

ResearchGate

<https://www.researchgate.net/>

ORCID

<https://orcid.org/>

Possibly useful further reading:

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures: *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>

Evans, C., & Carlyle, J. A. (2021). Outcome measures and evaluation in counselling and psychotherapy. SAGE. Paperback ISBN: 9781473906730

Greenhalgh, T. (1997). How to read a paper: The Medline database. *British Medical Journal*, 315, 180–183.

Greenhalgh, T., & Taylor, R. (1997). How to read a paper: Papers that go beyond numbers (qualitative research). *BMJ*, 315(7110), 740–743. <https://doi.org/10.1136/bmj.315.7110.740>

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Academic Medicine*, 89(9), 1245–1251. <https://doi.org/10.1097/ACM.0000000000000388>

Sanderson, S., Tatt, I. D., & Higgins, J. P. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography. *International Journal of Epidemiology*, 36(3), 666–676. <https://doi.org/10.1093/ije/dym018>

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, 19(6), 349–357. <https://doi.org/10.1093/intqhc/mzm042>

Van Ommeren, M. (2003). Validity issues in transcultural epidemiology. *British Journal of Psychiatry*, 182(5), 376–378. <https://doi.org/10.1192/bjp.182.5.376>

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & STROBE Initiative. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *BMJ (Clinical Research Ed.)*, 335(7624), 806–808. <https://doi.org/10.1136/bmj.39335.541782.AD>

