Anderson Andrew (Orcid ID: 0000-0001-7015-0061)

GUEST EDITORIAL

Writing an Introduction to a scientific paper

I believe a surprising number of authors of scientific articles write poor Introductions. Indeed, a poorly constructed Introduction is one of the most common reasons for my recommending that a manuscript be rejected. Even if flaws in an Introduction can be wiped away via feedback from the peer review process, this can cause problems of scientific integrity. For example, a peer review noting that the Introduction lacks a hypothesis will commonly lead to the authors providing one in a subsequent revision. This is acceptable if what is being reported truly was the hypothesis that preceded the study design, and previously was unclearly reported or simply omitted in error. More problematic — and, I suspect, more common — is if the hypothesis was manufactured after the fact, a process sometimes called HARKing (Hypothesis After Results Known); Kerr [1] outlines the many problems with this practice.

It is first necessary to understand what role an Introduction serves. Generally, Introductions should provide a compelling argument for why the current study is required. To do this, the Introduction should outline the shortest logical path through the critical literature that can support the study's hypotheses and/or aims.

When crafting an Introduction, it can be useful initially to determine this logical path in short-form as bulleted points, so that the logic can be ruthlessly scrutinized. For example, the following describes the shortest logical path through the relevant literature to support a (fanciful) paper about cats:

- In a random sample of cats sitting around household fireplaces, blue cats are warmer than red cats (Moggie et al., 2007).

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/opo.13071

- In a laboratory study, Purr and Miaux (2012) showed that blue cat fur absorbed more infra-red radiation, and argued that this could explain blue cats' warmth.

- However, Pett et al (2017) have found mammals with the blue pigment gene also have significantly more mitochondria: therefore, might blue cats be intrinsically warmer due to a higher metabolic rate, independent of absorbed infra-red radiation?

- In this study we measured a sample of blue and red cats both near to and away from a fireplace, to test the hypothesis that blue cats will remain warmer than red cats when away from an infra-red source, due to the former having a higher metabolic rate.

The Introduction — either in short or long form — should be written prior to designing a study. If you are not able to articulate a logical path through the critical literature to support a study, then it is likely that the study should not be commenced. Furthermore, having an appropriately articulated hypothesis or aim supported by a compelling Introduction brings clarity to precisely what your experimental design needs to be, how the results need to be analysed and reported and what previous work might need to be discussed in light of these results: in essence, a good Introduction guides the structure of all aspects of your study and the subsequent paper describing it.

Most commonly, an Introduction will end up justifying the need for a study in one of two ways:

- **A significant knowledge-gap in the literature has been identified**, which the current study aims to fill, or

- **A conflict has been identified in the literature**, which the current study aims to reconcile (hopefully in a hypothesis-led way). The fanciful logical path outlined above falls into this category.

Note should be taken of the modifier "significant" applied to knowledge-gaps. Merely noting a gap's existence is not sufficient, as novelty is not inherently a justification for a study. I am unaware of work investigating the association between a person's preference for the colour orange and their risk of vision loss from glaucoma. Therefore, such an investigation would be highly novel. However, there is good reason why no previous studies have investigated this and so why a knowledge-gap exists: there is no logical argument for why colour preference should relate to future glaucomatous vision loss. A good Introduction should provide an argument for why it is important to plug an identified gap in the literature, if it is not immediately obvious.

Note, too, the mention of *reconciling* previous conflicts. Some Introductions successfully identify there is a conflict, but then stumble by failing to hypothesise a cause for the conflict. Most commonly this occurs when a paper attempts merely to replicate previous experimental approaches. For example, an Introduction may successfully identify a conflict by noting that a group of previous studies of a particular phenomenon found one thing, whereas another group of studies found the opposite. The Introduction then concludes along the lines of "because of this identified conflict, here we perform the previous experiment (or a small variant of it) again". Why should the outcome of the current study hold more weight than previous work? Assuming some sort of binary outcome, the findings of the new study will necessarily agree with some of the previous literature and not others. This certainly doesn't reconcile the conflict, and may actually add further confusion. With enough new studies, it may be that some vague consensus emerges, but this is not an efficient way for knowledge to progress. Much better is to hypothesise a plausible reason *why* previous work may be in conflict, and to design an experiment to test this hypothesis. The new study then provides a definite outcome: it allows previous conflicting work to be reconciled, or it allows a plausible reason for why previous work was in conflict to be abandoned so that other reasons can be

sought. As once noted by a wise reviewer, the best papers "simultaneously add and take away from the literature"; they add new knowledge and, based on this knowledge, allow previous conflicting work to be reconciled (a *conflict* is taken away from the literature) or flaws in previous work to be noted (*papers* are effectively taken away because they are found to be flawed).

Logic is the key!

It is worth emphasising the need for a *logical* path through the literature. All too often an Introduction provides a variety of interesting background facts, but little (or no) logical linking of these. Even experienced researchers sometimes fall into the trap of treating the Introduction as some form of task list: "I need to have a section that talks about this, and a section about this other thing, and I also need to mention the work of authors X and Y". What starts life as a list, often continues as one. What is missing is a narrative structure that clearly outlines a logical argument. I would therefore discourage authors from initially thinking "What needs to be in my Introduction", and rather think "What is the logical argument that the Introduction is trying to make?"

One indicator that an Introduction has become a list of background facts, rather than a logical argument, is the absence of linking conjunctions. For example, conjunctions such as "However,..." or "In contrast,…" often highlight where conflicts exist between studies, that the current study might redress. Conjunctions such as "Therefore,", "Because of this,…" or "As such,…" often highlight where a logical extension of an argument — or a hypothesis — is being made, which the study might test. A further indicator that you are failing to outline a logical argument is if your Introduction is still readable when the sentences describing previous studies are omitted. This suggests your discussion of the literature is not being used to shape an argument, but is largely descriptive.

This is not to say that facts cannot appear in the Introduction. Often concepts need to be defined, a particular unfamiliar technique briefly outlined or the general importance of a field of study established. However, such statements should be added to a well-thought out shortest logical path (see above) and be kept as brief as possible so that the logical flow of this path is not interrupted.

The best Introductions perform a magic trick

By clearly outlining the logical argument that leads to a hypothesis and/or aim, the best Introductions allow the reader to predict what the experiment is going to be. Ideally, a reader should finish reading the Introduction's penultimate paragraph and be able to say (for example) "Based on what I have just read, someone really needs to do an experiment to see if X causes Y, which would explain why previous work is in conflict…". As if by magic, the final paragraph then precisely mirrors the reader's prediction: "Here, we undertook an experiment to determine if X causes Y, and so can explain…".

A particularly ruthless test for this is to give your Introduction to a naïve reader, but with the final statement summarising your experimental aim or hypothesis removed. You can then ask the reader to repeat the Introduction's logical argument in their own words, and to predict what is the experiment you are about to perform. Given that the result is sometimes disastrous (the reader has no idea of what the logical argument is and, consequently, could not predict what experiment needs to be done), it is important that you inform them that this is very much a test of your abilities, and not theirs!

But the "Instructions for Authors" say do not provide an extensive literature review!

Instructions for Authors sometimes suggest that it is not the role of the Introduction to extensively review the literature. This should *never* be taken as license to move mention of

key studies to the Discussion. Yet many times this is exactly what happens. Indeed, particularly for some clinically focussed journals, it is almost as if the authors have learnt this as a *sine qua non* for constructing papers. A reviewer reads the Introduction from such papers and gets the impression that no other study has previously undertaken a similar investigation. When finally arriving at the Discussion, the reviewer is then confronted with statements like "Our findings confirm those of X" or "Our study is in agreement with the work of Y" along with a list of previously unmentioned papers that the current study largely replicates (except, commonly, with less well-controlled methods). Whilst I have no direct evidence for this, I suspect this arises from investigators thinking more about *what* they can do ("We have the resources to run a group of glaucoma patients on test A and test B, and compare the results") without spending sufficient time about *why* they should do it. This commonly leads to performing unoriginal work, which researchers then try to hide by pushing all the relevant work — which is often much more robustly investigated in previous studies — into the Discussion. I suspect that the existence of such studies only became known to the authors after they completed their work, which highlights the danger of treating the Introduction as something constructed in the writing-up phase after the study is complete, rather than an important first step that directs the study design. Such studies are unacceptable and border on the deceptive, as they initially present a study as being original, when often it is a replication. To satisfactorily address the issue of why a study needs to be performed, similar studies must first be considered in the Introduction so that it is clear the new study is required. Without such a justification, it is difficult to argue why a paper should be published. This is not to say that replication itself cannot be a justification, but an argument for why a replication study is useful needs to be made in the Introduction. Most commonly the argument will be a variant of the 'filling a knowledge-gap' type: previous work supporting a position has limitations

(e.g., small sample sizes or a limited diversity of observers) and so our knowledge would be strengthened with the addition of the current study.

Indicators you may have gone wrong

Below is a "checklist" that may help identify when an Introduction is at risk of being poor:

1. **You wrote your Introduction after the study was completed**. Note that you may not have written the Introduction in full, but at least you should have bulleted the shortest logical path through the relevant literature that supports your arguments, ahead of designing and conducting the study.

2. **There is no hypothesis**

3. **There is a hypothesis, but a reader cannot predict what it would be** from the preceding contents of the Introduction

4. **You have described a series of facts in the Introduction**, but have not used these to build an argument that justifies your study. Key indicators are the notable absence of appropriate conjunctions, and when your Introduction remains readable when the sentences describing previous studies are omitted.

5. **Your Discussion is filled with statements saying your work is "consistent with that of author X" or "agrees with the findings of author Y"**. Is your paper simply a replication of something already known? To prevent this, the work of X and Y should have been discussed in the Introduction so that limitations and/or conflicts can be highlighted, which your paper should aim to address, ideally in a hypothesis-led way.

Andrew J Anderson

Department of Optometry and Vision Sciences, The University of Melbourne, Parkville 3010, Australia

E-mail: aaj@unimelb.edu.au

Corresponding author: Andrew Anderson, Department of Optometry & Vision Sciences, The University of Melbourne 3010, Australia. Ph (+613) 9035 9916. E-mail:

References

1. Kerr NL. HARKing: Hypothesizing After the Results are Known. Personality and Social Psychology Review. 1998;2:196-217.

Andrew Anderson is a member of the *Ophthalmic and Physiological Optics* Editorial Board. He qualified as an optometrist from The University of Melbourne in 1993, after which he completed Masters (1997) and PhD (2000) training at the same university. He subsequently undertook post-doctoral positions at the Devers Eye Institute, Portland, Oregon, USA (2000-2002) and the Department of Physiology, University of Cambridge, UK (2002-2004). In 2005 he commenced a tenured lectureship within the Department of Optometry & Vision Sciences, The University of Melbourne, and in 2012 he received a 4-year *Future Fellowship* from Australian Research Council. His research work uses a range of psychophysical methods to determine how the human visual systems processes information, both in normal and diseased eyes, and how this information is subsequently used to make decisions.

Picture of Dr Anderson about here

OPO_13071_AJA (2022).jpg