Psychological Bulletin
1995, Vol. 118, No. 2, 178-182

0033-2909/95/$3.00

How to Critique a Published Article

Ray Hyman
University of Oregon

Few replies get published. Many that do get published could be more effective. A reply can be effective only if it is published, readable, and credible. This article suggests guidelines for writing good replies. The guidelines begin with things to keep in mind—for example, that you might be wrong—and things to do before you begin to write. The latter include considering not responding, considering alternative viewpoints, and distancing yourself from the task. Things to do while writing include using the principle of charity, avoiding ad hominem and emotionally laden arguments, and keeping the rebuttal short. After completing the first draft, you should put it aside for awhile as well as get a second opinion. A good reply complements rather than discredits the argument of the target author.

In the 1985 *Journal of Parapsychology,* I published a detailed critique of a series of parapsychology experiments known as the *ganzfeld studies* (Hyman, 1985). The late Charles Honorton, a leading parapsychologist and one of the major contributors to the ganzfeld database, wrote a lengthy reply to my critique that appeared in the same issue (Honorton, 1985). Although I disagreed with most of Honorton's points, I felt that much of his reply was based on legitimate points of disagreement.

Some of his other points, however, irritated me. I felt that they not only misrepresented what I had done and written but also were petty and intended to imply that I was irresponsible and incompetent. I prepared a detailed rejoinder, almost as long as my original critique, and submitted it to the *Journal of Parapsychology.* The editor sent it to Honorton so that he could submit a counterreply to accompany my rejoinder.

Soon afterward, I encountered Honorton at a conference. We discussed our latest round of articles on the ganzfeld debate. I was surprised when Honorton informed me that the tone of my rejoinder had both shocked and hurt him. He felt compelled to reply in kind. I told him that he should have expected such a rejoinder given what I saw as deliberate, petty, and mean-spirited attempts in his reply to smear my character. Honorton saw the situation differently. He insisted that I was the one responsible for any provocations in our debate. At the time, I found his position incredible. With hindsight, I can now believe that we both sincerely believed the other was at fault.

As we discussed our controversy further, I realized some obvious things that had somehow previously escaped me. Our differences depended on specific details of methodological or statistical interpretations that were so minute that only Honorton and I would ever have mastered them sufficiently to be in a position to decide who was correct. Honorton's supporters believed Honorton was correct, whereas my supporters believed that I was correct. None of these third-party participants had the time or inclination to devote to the enormous task required to check these matters first hand. Instead they were content to accept Honorton's evaluations or my evaluations on faith. Even the few investigators who attempted their own analyses of the ganzfeld experiments did so by starting with either Honorton's or my quality evaluations as given. No one, as far as I could tell, other than Honorton or myself actually scrutinized the data from the individual experiments in the detail needed to decide whose quality judgments were the most plausible.

I realized that the continuing debate between Honorton and me was focused on the minutiae that only he and I could understand or would care about. Even worse, the haggling over these details was obscuring the more general issues. As we talked further, I was surprised to find that Honorton and I agreed on a number of central points. I proposed that I withdraw my reply and, instead of continuing what could have become a never-ending exchange of replies and rejoinders, that we attempt to write a joint article on our agreements and disagreements, with the emphasis on each of our agreements. Honorton accepted my proposal. We required four iterations to produce a manuscript on which both of us could agree (Hyman & Honorton, 1986).

I describe this personal experience not to propose it as a model for how disputes in the scientific literature should be resolved. Rather, I want to use it to bring up some considerations to keep in mind when critiquing a published article. An important consideration is how difficult it is to keep emotions from getting in the way of rational exchange. I believed that my initial critique of the ganzfeld research was objective and fair. Honorton clearly felt otherwise. I certainly was irritated and angered by what I perceived to be excesses in Honorton's counterattack. I believed that my reply, however, was again objective and well within the bounds of accepted practice. Honorton was upset at the "tone" of my reaction and was ready to respond in kind.

With hindsight, the details on which we were venting our mutual antagonisms involved minutiae that would not matter to most readers. Honorton and I considered these details to be vitally important, but readers probably had little patience or interest in such particulars. Our supporters were more than willing to trust the handling of these details to Honorton and me. They were more interested in the most general questions of what we agreed and disagreed on. When Honorton and I focused on these general points of agreement and disagreement, we were able to forge an article that not only succinctly captured

Correspondence concerning this article should be addressed to Ray Hyman, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

the essence of the debate but also pointed toward an agreed on procedure for resolving it.

The Problem With Replies

Good replies of published articles play an important role in the advancement of science. Rarely does a published contribution stand as the last or best commentary on its argument. The original argument, or the data on which it is based, often requires correction, supplementation, or alternate explications. The scientific knowledge base grows as much from these extensions or revisions as it does from original contributions. The reply can contribute to the growth of science, however, only if it is read, understood, and accepted as credible. At the very least, it has to get published.

An example of what I consider to be a constructive exchange that avoids the pitfalls I discuss is an exchange between Over and Evans (1994) on the one hand and K. N. Kirby (1994a, 1994b) on the other. Over and Evans, whereas finding some faults with Kirby's experiments and interpretations, applauded Kirby's application of signal detection theory to the Wason selection task. Kirby, in his rejoinder, defended both his experiments and interpretations but at the same time admitted a basis for Over and Evans' concerns. He also commended their use of the idea of epistemic utility as complementing his own approach and as being potentially very important. The exchange was not only civil but also appeared to have advanced the prospects for eventually understanding what is occurring in a task that has baffled psychologists for such a long time.

Debates and discussions about claims occur in a variety of formats in the psychological literature. Some publications have a letters to the editor section, but most do not. Some have special sections for short commentaries. More often, a response or commentary to a published argument has to undergo the same refereeing process as the original article and must satisfy the editor and the referees in meeting the same standards as an original contribution. Sometimes, the editor solicits or invites commentary from relevant individuals.

In writing the present article, the prototype I have in mind is as follows: An article is published in which the author argues for a conclusion that may or may not challenge the prevailing position on an issue. Another author, usually someone identified with a view or position different from that of the target article, writes a reply to the arguments of the target article. If the reply is published, it typically is accompanied by a short rejoinder by the target author.

Over and Evans's (1994) commentary on Kirby (1994a) and Kirby's rejoinder (1994b) closely fit the prototype. The category I have in mind, however, is a fuzzy one, and what I write is intended to apply to exchanges that depart in various ways from the prototype. Although I can point to examples, I find it difficult to provide an explicit definition that would unambiguously classify exchanges as either belonging or not belonging to the set I have in mind. For example, I would tend to exclude from this discussion book reviews and replies by disgruntled authors. Yet, I would cite Price's (1955) article as an example of a misguided reply, even though it can be construed as a long book review. The foci of Price's reply are the experiments and conclusions in Soal and Bateman's *Modern Experiments in Telepathy* (1954). As a book review, this took up an unprecedented nine pages in *Science*. The ensuing reaction resulted in all the articles of the January 6, 1956, issue of *Science* being devoted to replies to Price (Bridgeman, 1956; Meehl & Scriven, 1956; Rhine, 1956a, 1956b; Soal, 1956; Wolfle, 1956) as well as Price's (1956) rejoinder.

Most replies submitted to professional journals usually do not get published. The editor of *Psychological Bulletin* turns down most reply submissions because they are too minute in the points they make, too biased, too limited in scope, and too ad hominem. Even the few replies that do succeed in making it through the review procedure often fail to get their point across.

Guidelines for Writing Successful Replies

No handbooks exist for writing replies. Nor does it make sense to attempt to generate a list of specific rules that apply to all situations. The reasons for responding to a published article are varied, and the specific procedures for rebutting the arguments in an article vary correspondingly. Nevertheless, it is possible to suggest some general guidelines that cover most situations. These guidelines contain no surprises or secrets. Anyone, with a moment's reflection, could probably write a list of such guidelines that will almost certainly duplicate most, if not all, of the ones I provide. Although these guidelines are available to all, they frequently become inaccessible at the time one feels compelled to right a grievous wrong in a published article.

Things to Keep in Mind

Before you begin your reply, you should consider the possibility that you, not your target, might be wrong. Remember, a published article has survived scrutiny by referees who presumably have some expertise in the relevant area. Also, in setting yourself up as an expert who can provide the *correct* account, keep in mind that you are implicitly suggesting you are somehow smarter, better informed, or otherwise superior to the author of the target article. Even when you explicitly disavow that you are setting yourself above the author, the subtext of a reply almost always implies that the original author is inadequate in some way. You should use tact and skill to frame your reply in such a way that you do not appear to be casting aspersions on the competency of your antagonist.

Before You Begin Writing

Consider not responding. Despite your indignation or sincere desire to set the record straight, more often than not the best course is to do nothing. If, indeed, the target article has erred in some way, subsequent research and developments in the field will often suffice to remedy the error. Both the reply and subsequent rejoinder by the original author will consume much of the author's and your time and resources. Ask yourself whether you and your opponent might better contribute to science and humanity by devoting your time to other matters.

Perhaps this suggestion is unrealistic. The urge to set the record straight and to rectify what you believe to be a wrong conclusion on the basis of misrepresentations and biased interpretations does not succumb to advice to do nothing. Getting

advice from a neutral party on this might help. Unfortunately, often it is only in hindsight that you realize it might have been better not to respond.

Price's (1955) critique of Soal and Bateman (1954) presents an interesting, if unusual, example. Price argued that if you accept Rhine's (Pratt, Rhine, Smith, Stuart, & Greenwood, 1940) and Soal's (Soal & Bateman, 1954) experiments on ESP and telepathy *as reported,* then you are forced to conclude that they have demonstrated the reality of ESP. Price asserted that they have convincingly proved their case by accepted scientific standards. However, to accept their conclusions, Price maintained, is tantamount to denying just about all the gains of modern science. Price then invoked David Hume and Thomas Paine to the effect that if one or more people vouch for phenomena which would be miraculous in the framework of contemporary science, then you are justified in calling these claimants liars. Price, admitting he had no factual basis for making his accusations, asserted that Rhine and Soal must have cheated to get their results. He described several possible, but elaborate, scenarios of how such cheating might have occurred in Soal's (Soal & Bateman, 1954) experiments.

I believe that Price's (1955) critique should never have been published. Price's scenarios were unrealistic and, in my opinion, he created a false dichotomy by insisting that the parapsychologists either had proved their claims or had cheated. Ironically, both parapsychologists and critics now agree that Soal (Soal & Bateman, 1954), in fact, did cheat after all. However, Price's scenarios were wrong. Price, it turned out, was right (about Soal cheating) but for the wrong reasons. The eventual discovery that Soal cheated and how he did it was made by Markwick (1978). She did so by examining his data with the intention of clearing Soal of the suspicion of having manipulated his data.

Price's (1955) willingness to accuse investigators of cheating purely on the basis of a priori ideas of what is scientifically possible carried criticism beyond reasonable bounds. Boring (1955) also critiqued Soal and Bateman's (1954) article by relying on a priori notions of what is possible and what is not. Boring, who otherwise was a careful and reliable scholar, asserted that "in a good experiment you would turn telepathy on and note the number of hits. Then you would turn it off—the control experiment—and note the number" (p. 112). I was astounded to encounter this statement from the usually reliable Boring. I can only conclude that Boring was so focused on his a priori argument that he completely missed the fact that Soal and Bateman clearly emphasized carrying out only those controls that Boring says they did not!

Consider alternative viewpoints. Try putting yourself, in turn, in the shoes of the original author, the editor of the relevant journal, and the typical reader of that journal. Imagine you are the author or the author's advocate. You might be surprised how much support you can find for the author's position.

Distance yourself from your reply. Do not rush into your rebuttal. Emotion and vitriol are more dominant in initial reactions to published articles with which you disagree. By putting such an article aside and returning to it at a later time and in a calmer frame of mind, you can gain a new perspective and often will realize aspects about the original article that you previously overlooked. To put the issue in perspective, ask yourself whether all this will matter a year from now. How about 10 years from now?

Answer some preliminary questions. You might find it helpful to answer some obvious, but potentially critical, questions before you begin. Some of these questions are implicit in what I discussed in preceding sections. Some important questions to ask yourself are the following: Who cares? Does it really matter and to whom? What are the reasons, causes, explanations, data, and things to which you are objecting? Who is your target audience? Have you made a list of all the issues and points on which you and your opponent can agree? What aspects of the target article do you consider praiseworthy?

Take pains to ensure that you correctly understand the target claim. This suggestion is related to the principle of charity that I discuss in the next section. In reviewing a number of exchanges, I was struck by the almost universal complaint of target authors that their position had been misunderstood, misrepresented, or otherwise distorted. Many target authors accuse their antagonists of attacking straw men. Of course, it is possible that the target author is responsible for the misunderstanding. Perhaps he or she was not entirely clear in presenting the original argument. In many cases, the position being attacked has developed over time, and the claims and theoretical arguments may have undergone several revisions. Sometimes, what the target author has claimed and what his or her followers believe differ in significant details.

A recent personal experience in trying to rebut a claim was the exchange between Bem and Honorton and myself (Bem, 1994; Bem & Honorton, 1994; Hyman, 1994). Several things conspired to make it difficult for me to focus on Bem and Honorton's specific claim as it appeared in the published version. I was familiar with the earlier article by Honorton et al. (1990) that first presented the findings that were central to Bem and Honorton article. The earlier article contained a much stronger claim. Before I saw the manuscript that Bem and Honorton submitted to *Psychological Bulletin,* I had read press reports about it. These press reports also made very strong claims. The initial manuscript also contained stronger claims than the final version. By the time the original manuscript had gone through its revisions, the claims being made were quite mild and rather uncontroversial. The only remnant of the initially strong claim was in the title that implied that replicable evidence already existed. As the exchange revealed, Bem and I both came to the innocuous conclusion that the fate of the ganzfeld-psi hypothesis depends on forthcoming attempts at independent replication. We differed only on how optimistic we were about the outcome.

The referees faulted the first commentary I submitted because, as they correctly observed, I was rebutting the claims of the original Honorton et al. (1990) article rather than the somewhat milder and different version of the revised article. I eventually managed, I hope, to confine my commentary to the claims as they appeared in the published Bem and Honorton (1994) article rather than to the earlier and stronger variants.

I use this personal example to show how difficult it is to pin down and focus on the correct formulation of the claim to rebut. Despite this difficulty, the attempt to respond to the most accurate and latest version of the target's claim is very important. The quality and value of the exchange improve dramati-

cally when both sides believe that their opponent has correctly understood their position.

Take pains to avoid the appearance of self-serving data selection. Almost as common in the exchanges I have reviewed is the complaint that one's antagonist has overlooked important studies (or, in some cases, has included irrelevant or badly flawed studies). Because antagonists in these exchanges often communicate with each other before the exchange is published, it may be possible to avoid both this and the preceding problem of alleged misrepresentation. The person preparing the reply could supply the target author with (a) a formulation of the target author's position to see if this, indeed, is correct; and (b) a list of references used in evaluating the claim to see if they are acceptable and if any key studies have been overlooked.

Writing Your Reply

Use the principle of charity. By *charity* I do not intend the everyday sense of giving or helping those less fortunate than you. The sense of *leniency in judging others* is somewhat closer to what I have in mind. Even closer is the use in contemporary philosophy that asserts you should assume your adversaries are rational. If their arguments appear weak, silly, or patently wrong, then, according to this principle, you should look for the apparent weakness in how the argument was expressed rather than in the argument itself. For my purposes, the point is to give your adversary the benefit of the doubt. You should not only state your opponent's case in a fair and objective manner, but you also should *try to rephrase the argument* in its strongest possible light. Too often the reply is an attack on a straw person. If your reply is worth taking seriously, you should critique the opposing argument in its strongest presentation. If you do not attack the best argument and evidence for the opposing case, you will fail to convince thoughtful, but previously neutral, readers.

Avoid ad hominem attacks and emotional language. Although this guideline is obvious, it is consistently violated. One reason might be that the writer of a reply is too involved and close to the task to notice violations that are obvious to less involved individuals. Another reason might be that if the adversaries come from different traditions or scientific subcultures, they may not realize that seemingly innocent words can trigger strong emotions in others. The major difficulty that Honorton and I experienced in composing an article that would be mutually acceptable was not so much in finding a consensus on the big issues but in finding the right phrase or word that would not be offensive or distasteful to either one of us. I was constantly surprised to find that words or phrases that I considered to be neutral or inoffensive were strongly rejected by Honorton as being distasteful or unacceptable in his framework. In turn, Honorton seemed amazed by some of his wordings that I insisted had to be changed.

Ideally, a good reply should allow the facts and arguments to speak for themselves. If you have a strong case, then you should not have to buttress it with valuative judgments about its superiority over its rivals.

I would like to think that a careful application of the principle of charity in conjunction with the desire to avoid ad hominem arguments could have prevented some of the bitterness and hostility that have characterized some published exchanges. A good reply should not only avoid personal attacks but also should take pains to avoid even the possibility that its comments might be wrongly construed as a personal attack.

Russell's (1994) critique of the universality hypothesis of emotional recognition from facial expressions contained a historical reconstruction that brought a spirited defense from Ekman (1994). Whether or not Russell intended to devalue Ekman's contributions, Ekman obviously believed that

> Russell implied not only that Izard and I were poor scholars but also that we did not contribute very much to the research or theory on facial expression. To create such an impression, Russell's account misrepresented who did what, selectively reporting and quoting many of those he cited. (p. 286)

Nonetheless, Russell's critique and Ekman's reply made it to publication without a prior resolution of this emotionally laden issue. This is all the more regrettable because the more substantive aspects of the disagreements involve a huge, complicated database whose direct inspection is beyond the resources of most readers, especially researchers outside the field of emotions and facial recognition. A problem that plagued this dispute, as well as many others, was what studies to include and exclude from the argument. Both sides accused the other of omitting crucial data. One suggestion that might work in cases such as these is for the editor and the referees to insist that the disputants agree on the list of studies to use in the argument. This, however, would not solve all the problems because even the studies that both sides of the dispute used evoked charges of misrepresentation and selective reporting.

Keep it short. A good rule of thumb is that a reply should be less than half the length of the target article. If the issues are so complicated and technical that, in your opinion, they require more space, you should take this as a warning sign. The probability of getting a longer reply accepted is much lower than getting a shorter one accepted.[1] Given that most replies get rejected, submitting a long one is a bad strategy. A long and technical reply is beyond what most readers are willing to work through. The audience for such detailed replies is reduced to your adversary and a few strongly motivated partisans. Such a situation is not appealing to an editor.

As an alternative strategy, consider discussing your objections in a general and polite way. Offer to readers who are interested in your more detailed reply the opportunity to get the longer, unpublished version directly from you.

After Writing the Reply

When you have finished your first version, put it aside for a day or so before reviewing it. This is another way to put some distance between you and the reply. This also allows you time to get a second opinion. Ask a colleague who has no vested interest in the controversy to read your article. In particular, ask

[1] Exceptions can obviously occur. Harris's (1993a, 1993b) reply of Coren and Halpern's (1991) article on the shorter life expectancy for left-handers was about twice as long as the target article. Perhaps this was allowed because of the obvious interest of the topic even to nonpsychologists.

him or her to look for ad hominem attacks and emotionally charged phrases. If possible, also try to get an opinion from someone who you know is sympathetic toward your adversary's position.

What Makes a Good Reply

Implicit in the preceding discussion is the proposition that the quality of the reply depends on the writer's goals. Replies lose credibility and are less likely to get published when they are seen as vehicles for personal attacks, defending one's honor, or both. Replies gain credibility and are more likely to be published when they are neutral in tone and focus on constructive criticism. In most cases, a reply is motivated by several goals, some explicit and others implicit.

As you work at your reply, ask yourself what you are trying to accomplish. Are you trying to humiliate your adversary? To salvage your pride? To vent your anger? To display your superior grasp of the issues? To promote your viewpoint? You may not acknowledge any of these goals as your own, but could your choice of language, illustrations, and mode of argument convey the impression of such goals to your readers? Here, again, is where a second opinion from a neutral colleague can sense what you may not.

Ideally, your reply should attempt, and be seen as attempting, to make a constructive contribution to the topic of the target article. A rough ordinal scale from least (a) to most valuable (d) can be set forth as follows: (a) simply showing that the target article is wrong, (b) showing not only that the target article is wrong but also what is right, (c) acknowledging the good points in the target article and allowing for the possibility that the author might be wrong for good reasons, and (d) showing that the target article is wrong in the context of contributing new insights and integration to the topic under discussion.

To summarize, a good reply is one that amends, elaborates, or otherwise clarifies and expands issues raised by the target article without disparaging the contributions of the original author. A good reply deals with the general issues in a way that informs and educates the reader, and it focuses on implications for an entire field rather than on minutiae that matter to only a few specialists within a narrow area of inquiry. To the extent that the reply appears to complement rather than to discredit its adversary, it increases its chances of contributing to the advancement of scientific inquiry.

References

Bem, D. J. (1994). Response to Hyman. *Psychological Bulletin, 115,* 25–27.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115,* 4–18.

Boring, E. G. (1955). The present status of parapsychology. *American Scientist, 43,* 108–117.

Bridgman, P. W. (1956, January 6). Probability, logic, and ESP. *Science, 123,* 15–17.

Coren, S., & Halpern, D. F. (1991). Left-handedness: A marker for decreased survival fitness. *Psychological Bulletin, 109,* 90–106.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268–287.

Harris, L. J. (1993a). Do left-handers die sooner than right-handers? Commentary on Coren and Halpern's (1991) "Left-handedness: A marker for decreased survival fitness." *Psychological Bulletin, 114,* 203–234.

Harris, L. J. (1993b). Reply to Halpern and Coren. *Psychological Bulletin, 114,* 242–247.

Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology, 49,* 51–91.

Honorton, C., Berger, R. E., Varvogolis, M. P., Quant, M., Derr, P., Schechter, E. I., & Ferrari, D. C. (1990). Psi communication in the ganzfeld: Experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. *Journal of Parapsychology, 54,* 99–139.

Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology, 49,* 3–49.

Hyman, R. (1994). Anomaly or artifact? Comments on Bem and Honorton. *Psychological Bulletin, 115,* 19–24.

Hyman, R., & Honorton, C. H. (1986). A joint communiqué: The psi ganzfeld controversy. *Journal of Parapsychology, 50,* 351–364.

Kirby, K. N. (1994a). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition, 51,* 1–28.

Kirby, K. N. (1994b). False alarm: A reply to Over and Evans. *Cognition, 52,* 245–250.

Markwick, B. (1978). The Soal–Goldney experiments with Basil Shackleton: New evidence of data manipulation. *Journal of the Society for Psychical Research, 56,* 250–277.

Meehl, P. E., & Scriven, M. (1956, January 6). Compatibility of science and ESP. *Science, 123,* 14–15.

Over, D. E., & Evans, J. St. B. T. (1994). Hits and misses: Kirby on the selection task. *Cognition, 52,* 235–243.

Pratt, J. G., Rhine, J. B., Smith, B. M., Stuart, C. E., & Greenwood, J. A. (1940). *Extra-sensory perception after sixty years.* New York: Holt.

Price, G. R. (1955, August 26). Science and the supernatural. *Science, 122,* 359–367.

Price, G. R. (1956, January 6). Where is the definitive experiment? *Science, 123,* 17–18.

Rhine, J. B. (1956a, January 6). Comments on "Science and the supernatural." *Science, 123,* 11–14.

Rhine, J. B. (1956b, January 6). The experiment should fit the hypothesis. *Science, 123,* 19.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115,* 102–141.

Soal, S. G. (1956, January 6). On "Science and the supernatural." *Science, 123,* 9–11.

Soal, S. G., & Bateman, F. (1954). *Modern experiments in telepathy.* New Haven, CT: Yale University Press.

Wolfle, D. (1956, January 6). Extrasensory perception [Editorial]. *Science, 123,* 7.

Received April 4, 1994
Revision received January 18, 1995
Accepted January 23, 1995 ■