9

How to Read a Scientific Article

The QDAFI Method of Structured Relevant Gist

Pascal Wallisch

Most academic fields document and communicate the knowledge gained from research in the form of articles published in professional journals. These publications are full of the latest information as to where the field is at the current moment and where it is likely to go next. They also constitute a historical record of how the empirical edifice of any given field came to be. Thus being able to gainfully read original research papers is an indispensable skill. Unless we are able to do that, we will have to rely on a summary by someone else, who might well have a limited understanding of the research and its implications, or have interests misaligned with those of the researchers, or a combination of both.

Being able to access this information ourselves is important. But there are roadblocks. Lots of them.

Expert and Non-Expert Readers

Written by experts, for experts, scientific articles tend to be full of specialized jargon while also leaving many important things unsaid, as the authors can rely on the shared assumptions and tacit knowledge of the intended audience. However, there are times when non-experts want or need to read

Critical Reading Across the Curriculum, Volume 2: Social and Natural Sciences, First Edition.
Edited by Anton Borst and Robert DiYanni.

a research paper, as in the case of students, journalists, or simply interested laypeople. This presents a problem for these readers, as their lack of common ground will make much of what is in the paper seem like gibberish (Clark and Brennan, 1991). Even worse – especially for students – is that because we cannot remember what we cannot understand (Bartlett, 1932), even after heroically struggling through a paper, we will retain very little of this hard-won information long term.

Moreover, authors attempt to convey to other researchers exactly what took place in the study – in principle one should be able to replicate the research solely from the information provided in the article itself, something that is an increasingly critical consideration (OSC, 2015). Consequently, research papers tend to be chock full of technical details, not all of which are equally important to the overall point of a paper. There is probably no correlation between how easy something is to understand in a paper and how important it is. Keeping in mind that people tend to remember what they are able to understand, and that non-experts won't necessarily know what is important and what is not, some readers might remember particulars that are easy to understand, like that there were 57 participants in the second experiment, but which are of negligible significance. This issue is exacerbated by the fact that the retention of minutiae likely comes at the expense of more relevant information. People generally tend not to retain an accurate and comprehensive memory of the information actually presented. Rather, our long-term memory performs a kind of compression operation, a compression that happens in a semantic space: we retain whatever meaning we are able to extract, but not much else. Known as memory for "gist," what is remembered long term is often surprisingly sparse and a caricature of the original information (Reyna and Brainerd, 1995). In other words, most of the information encountered is discarded, including most – if not all – details of how it was presented, such as syntax, particular phrasing, font types, and the like.

These concerns illustrate how reading a research paper poses a formidable challenge to non-experts such as students or journalists. The first step in meeting and overcoming this challenge is to acknowledge its tremendous magnitude. Unless someone is a scientist in the same field as the author, one should not expect to be able to read a research paper as effortlessly as something written for a general audience, such as a piece in a newspaper or magazine. Non-expert readers should not expect to be able to get anything useful out of reading research papers willy-nilly. The good news is that being able to read a research paper is a skill, and skills can be acquired. For this purpose, I have developed the "QDAFI" method. The QDAFI method enables non-expert readers to gainfully read research papers with a good deal of understanding. It was designed to exploit the typical structure of

research papers, as well as to leverage the cognitive aspects of reading research papers, some of which were outlined above.

For instance, if we know that readers have a tendency to retain only the gist of some passage, we can anticipate this and get ahead of it by trying to make sure that the gist that is being retained is the *relevant* gist – what matters most about the article in question from the perspective of someone trying to remember it in the future. There are other considerations as well. Research in cognitive psychology suggests that the format of information, which shapes how that information is encoded, is of critical importance for retrieval (Craik and Tulving, 1975; Roediger and Karpicke, 2006). As we usually want to retrieve information from a paper in the form of a targeted question (such as "what did the study find?"), it makes sense to encode the information in the form of answers to the handful of questions most likely to be asked. Every research paper, regardless of field or content, is more or less structured in a way conducive to this approach. The QDAFI method amounts to a mining operation: it will require a bit of work, but we're essentially trying to extract nuggets of relevant information from the paper by digging for them in places where we know that the authors put them, sometimes buried under jargon or masses of extraneous information that is of interest only to experts.

The QDAFI method yields, deliberately, only gist, and gist of the kind that we'll need if we are to remember what matters and discuss the paper in the future – in other words, structured and relevant gist. The method consists of answering five questions that are the same for every paper – and in my view the most important to ask of any paper. These questions structure the relevant gist and tend to correspond to the structure of the article itself, which provides an important guide for the non-expert reader. The answers must be brief, between one (ideally) and three (maximally) sentences each, together fitting on a single page (without fiddling with the margins or fonts). This constraint necessarily enforces the discipline needed to focus on the most relevant information: the gist, which is all that readers remember anyway. Putting in the necessary work to condense our understanding of a paper to a few sentences does wonders for our grasp of its purpose and scientific relevance.

The QDAFI Method: An Overview

Specifically, the QDAFI is a brief summary that consists of the answers to the following five questions:

Q: What was the **question** that the authors tried to answer?
D: What did the authors **do** to answer the question?

A: What was the **authors'** rationale?
F: What did the authors **find?**
I: What is the authors' **interpretation** of these findings with regard to the initial question?

Let us now look at each item of the QDAFI (and how to address it) in detail.

Q: What *question* did the authors try to answer in this paper?

Every good research paper starts with a question it is trying to answer. It is worth restating this question explicitly and as succinctly as possible in the first section of the QDAFI, as without it none of the other sections will make any sense. Whatever the authors did, they presumably did it to answer this question.

The Q portion of the QDAFI method can usually be addressed by looking at the *Introduction* section of papers. If we are lucky, the authors state their question outright; if not, and much more commonly, we will have to read between the lines to reconstruct what we think the authors had in mind. One complication is that there might actually be two questions explicitly or implicitly stated in the paper. The one question that authors will always state explicitly is the "specific research question" the study was designed and executed to answer. This question could be anything – for example, "What is the influence of changing font color on memory retention?" The second question, and the one that usually pervades the spirit of the introduction, is the "theoretical question." This question is typically the reason why the study was done in the first place. A study is typically not done arbitrarily, but touches on some larger theoretical issue or controversy in the field – for example, "What is the influence of context on memory?"

So section Q should ideally consist of one sentence stating the specific research question, or two sentences if we manage to spot the theoretical question as well. No more. Consider (for now, not when doing the QDAFI) all the things that are typically in an introduction, but won't make it into the QDAFI because they are irrelevant for that purpose. For instance, introductions often spend a lot of effort on describing why the question is important in order to make it more likely for the paper to be published in a prestigious journal. But for our purposes, this is irrelevant. Presumably readers are interested in a particular research paper and know why it is important to them. Also, introductions tend to dwell on all the other things that have already been done to answer the theoretical question,

highlighting why the authors think more research needed to be done (i.e. their paper). None of this academic self-justification is necessary for purposes of the QDAFI. Realizing this can be liberating, as non-experts in particular can be intimidated by the wall-to-wall citations common in many introductions, none of which the non-expert has read or needs to be concerned with.

D: What did the authors *do* to answer the question as reported in this paper?

In section D of the QDAFI, we want to state as succinctly as possible what the authors actually did in order to answer their research question. This information can be found in the *Method* section of the paper. However, finding it is tricky, as this section tends, along with *Results,* to be the densest of the entire paper, full of technical details of interest to experts. It is very easy to get lost in irrelevant minutiae here. How are we supposed to summarize all of this information in a single sentence, or at most two sentences? It is not easy, but it is doable. For studies that report experiments, one sentence can present the parameters that were varied (the independent variables), and the other what was measured (the dependent variables). If the study is observational, we can devote one sentence to what was measured and one to the conditions under which the measurements were performed.

Unless a detail is absolutely critical to the research question and the interpretation of results, and unless that detail makes a difference for remembering the study in a couple of years, it does not belong under D. We need to be extremely stingy in terms of what is included here. If the paper reports on multiple experiments, a third sentence can be used to outline how the conditions and measurements changed in each (e.g. "they controlled for potential confounding variables x, y and z"). We need to be only as specific as capturing the essence of the experiment requires. For instance, when reporting on what was varied, we may need only record the independent variables – e.g., "they varied luminance levels." But if a greater level of detail is the point of the experiment, because the authors were the first to achieve, for example, the specific luminance levels of "5, 10, 15, and 20 foot-lamberts," then we would note that information. The goal is to not get lost in the details, yet at the same time not leave out anything crucial. Unless it is critical to the question that the study asked, details such as the precise numbers of participants, their gender, etc., do not belong here. Again, as in section Q, recognizing this distinction between what is essential and what is not can be a great relief.

A: What was the *authors'* rationale?

This section is perhaps the trickiest of all, yet it is at the heart of the QDAFI method. Stating the rationale necessitates that readers understand what the authors were trying to do and paraphrase it in their own words. If we can state the rationale, we probably have a good grasp of the paper as a whole. While the other sections Q, D, F, and I can sometimes be completed by more or less copying from the paper, A often requires inferring an unstated purpose. So the rationale is crucial, indeed. It answers central questions to the meaning of the authors' work, such as "What was the idea behind the study?" and "What is its logic?" "What allows the authors to infer anything useful about the question from what they did?" And: "Why did the authors do what they did in D and not something else, out of the infinite number of things they could have done to answer their research question?"

The rationale or idea links the specific research question to the actual methods the authors used. Done properly, the rationale sets up what the authors can conclude from any possible outcome, given their methods and question. Without a valid rationale, nothing can be concluded from any given result. Contrary to popular belief, the data do not speak for themselves. Rationales are rarely spelled out explicitly, as either the experts in the intended audience will grasp them intuitively, or a given rationale might be part of the shared culture of a field.

Rationales can be tricky to track down, but the information necessary to make an informed guess about them is usually distributed in the *Introduction* and *Method* sections. For instance, the authors of a study might have a theory that more intelligent people have faster neurotransmission, which they discuss in their introduction. If this were the case, they could predict that such people are quicker to respond to a stimulus, and so the authors have designed an experiment to test the relationship between intelligence and response-time to a stimulus, which they explain in their *Method* section. Our A section of the QDAFI might therefore read: *It is believed that intelligent people have faster neurotransmission. Faster neurotransmission would be evident in faster response-times to stimulus*. With this rationale, we can now meaningfully interpret the results of our experiments: if a significant number of intelligent people are slower, this falsifies the theory, whereas if intelligent people were found to be consistently faster, it would provide empirical support in favor of the theory. Without a rationale, not much can be concluded from measuring reaction times and intelligence – it might even seem downright arbitrary to do so. The stronger the rationale, the more conclusive the results will be, whether for or against the

theoretical issue at stake. Depending on the complexity of the rationale of the study, this section might require anywhere from one to three sentences, but not more.

F: What did the authors *find*?

This is perhaps the most straightforward section of the entire QDAFI. What the authors found as the results of their study is usually reported in the *Results* section of the paper, and quite explicitly so. The biggest challenge here is to identify the key findings relevant to the question the authors asked initially and to understand them. Those findings are usually expressed in extremely technical language, often involving statistical notation. Again, the discipline imposed on us by the QDAFI method comes to the rescue. It forces us to identify the core of the matter. Which of the many results reported in a paper (the average paper in psychology contains no less than 11 p-values [Nuijten et al., 2015]) are actually relevant to the original question? Some authors like to build up to their main result; others hedge it with many side-considerations. Both approaches may leave readers with lots of information to sift through before identifying the most important findings.

Additionally, peer-reviewers who vet papers before publication often request additional analyses of personal interest to them, which authors may then obligingly incorporate in order to ensure their work is published. This adds even more extraneous information. Readers must find the key finding – the finding that matters in terms of the original question. Having completed A, the rationale, will now help us determine which finding or pattern of findings to look for in the *Results* section. Usually, this key finding can be stated in a single sentence. Sometimes, if the findings are particularly surprising or complicated, an additional sentence will be needed to elaborate. Never are more than three sentences necessary. Given the plethora of statistics typically reported by authors, such brevity can only be accomplished by focusing on the big picture – what matters in terms of the question, as spelled out by the rationale. The key is to focus on the facts. What did the authors actually find? This should be stated as dispassionately as possible, without mentioning anything about what these findings might mean. That is for the next, and last, section.

I: What is the *interpretation* of these findings with regard to the initial question?

Put simply, what do the authors think their findings mean? Answering this question is usually straightforward, as the authors will tell you what they

think in the *Discussion* section of the paper. But that's not all there is to the QDAFI's "I," which consists of one to three sentences. The first sentence records what the authors think the findings mean in terms of the original question. The second sentence identifies an issue that threatens this interpretation. The third states what we, the readers, think these findings mean. Thus there are potentially three I's to be addressed in this section: the *interpretation* by the authors, the *issues*, and the *interpretation* by the reader.

Suppose that some authors want to know whether the ability to delay gratification in childhood is a major predictor of life success (Mischel et al., 1989), and imagine that they find that it is. They then conclude that this finding suggests that the ability to forego small but immediate rewards for large but delayed ones is the bedrock of individual success, as well as social functioning. If we agree with this interpretation, we are done with this section in one sentence. However, it is possible that upon examination the authors' conclusions appear problematic. For instance, they might neglect to measure the level of trust of the experimental participants, trust that the delayed reward will in fact materialize, as the experimenter promises. It has been shown that children from lower socio-economic-status (SES) backgrounds have lower amounts of trust in promised future outcomes. In what for them is an unreliable world, going for the safe, immediate outcome might be the most rational strategy. In addition, low childhood SES is linked to poor adult life outcomes. In other words, a major potential confound has not been accounted for by the authors. While interpreted as the ability to delay gratification, what might really be being measured is trust level as a proxy for SES, which is already known to be linked to adult life outcomes.

Thus, having identified this problem with the authors' interpretation of their findings, we may now offer our own: unless ruled out in a future study, low childhood SES is linked to poor adult outcomes, likely mediated by trust. So far, so good. The key problem for this section is to identify an issue, the second "I." To be worth identifying and recording on the QDAFI, an issue must be so serious that it fundamentally threatens the authors' interpretation of their experiment's outcome – not some minor technical limitation.

It should be understood that every research paper suffers from a large number of problems. This is necessarily the case, as we live in a non-ideal world. Life is complicated. Things don't always work out. However, the issues that the QDAFI method leads us to focus on represent only problems that are true showstoppers. A real issue in this sense is usually something like faulty logic (e.g. the authors interpret a correlational study causally,

when they should have done an experiment); a potential confounding variable like the one we outlined above; or a serious technical problem (e.g. the method the authors used does not have the sensitivity to reliably detect the results the authors claim they found). It is not something trivial like the number of participants, as we always would like to have more research participants (unless the study is seriously underpowered [Wallisch, 2015]), or the size of the monitor on which the stimuli were presented (unless it was really too small for the participants to reliably see the stimuli), or other minor technical issues.

All studies have limitations, but not necessarily issues that threaten the conclusions (I) authors draw from their findings (F) with respect to their question (Q). It might not be reasonable or feasible for non-experts to find such fundamental flaws in a study. Instead of manufacturing issues, non-expert readers are advised to stick with the interpretation of the authors, and to succinctly state it in one sentence with a focus on the big-ticket item: what do these results mean with respect to the empirical question at hand? There are often plenty of considerations raised by any given study, but it is best to focus on the biggest one, the one described in section Q.

And that's it – that's the QDAFI in all its purposeful simplicity. Done right, it will easily fit on a single page and be memorable for years to come. In this case, less is definitely more. Each of the five sections of the QDAFI should be tweetable (using Twitter's original limit of 140 characters), ideally, but this is not a strict requirement.

Benefits of the QDAFI Method

Once completed a few times, the QDAFIs will reveal their usefulness. Being able to do a QDAFI is a beneficial skill for mainly three reasons. First, research papers are not written in a fashion that allows them to be read like any other writing. This is a problem, as they look very much like other written works, but treating them as such won't provide much value in the reading experience. Second, long-term memory is radically semantically reductive over time. Short-term memory can be deceptive because of this. After just having carefully read a paper, all the details will be readily available to our minds (assuming we have understood it). But this memory won't last. Most, if not all, of the details will be forgotten over the course of a few days, weeks, and months. The only thing remembered a couple of years later is what most caught our attention when we read it. This is a problem, especially if our brain at the time thought that the most salient or remarkable thing about the paper is that all the authors' last names started with a "G."

Expertise in any field consists of building up a database of highly semantic knowledge over time. This is, by the way, not a downside of memory, but rather a benefit: The brain needs to make sure that it is not getting too cluttered with irrelevant information. The QDAFI anticipates this housecleaning business and makes sure that the most important points about the paper are also the ones focused on when "reading" the paper. And that is why it is critical that the QDAFI be so short – the shorter the better.

Third, being able to complete a QDAFI has benefits beyond reading and remembering individual papers, providing tools for conducting research, synthesizing disciplinary knowledge, and producing our own research articles. The method offers a very efficient technique for developing literature reviews and summaries, which are essential for writing review papers and introductions to research. Because the QDAFI maximizes information transmission with the minimum amount of text in a clear and cogent way, it provides an effective template for structuring abstracts of original research. Finally, it can begin to inform how we write our own studies and papers, to make them more memorable to readers who may be applying, consciously or not, their own version of QDAFI to our work.

A Demonstration

So far, our discussion of the QDAFI method has been fairly abstract (though with occasional examples). Here I present sample solutions to two papers that are publicly and freely available online and that are of general interest. The first paper is technically straightforward, and investigates whether taking notes by hand versus taking them on a laptop matters for retention of the presented material. The second paper is much more technical, and explores whether individuals with autism exhibit any perceptual benefits. Thus, the choice of these papers illustrates the versatility of the QDAFI method, which works effectively regardless of a paper's specific content or technical sophistication. If you are interested in how the relevant information in long and complicated papers can be extracted and end up in QDAFI form, have a look at the original papers and then the sample QDAFIs below.

> Paper 1: Mueller, P.A., and Oppenheimer, D.M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science,* 25 (6), 1159–1168.
>
> Q: Does taking class notes by hand yield better academic performance than taking them on a laptop?

D: Participants were asked to either take notes on a laptop or by hand while watching TED talks. They were then asked to answer both factual and conceptual questions about the material presented in the talks.

A: It is believed that using laptops to take notes yields poor performance because of distractions. However, even without distractions, if laptop use leads to shallower processing, it could detrimentally impact performance on conceptual questions.

F: Students who typed notes on a computer performed worse on conceptual but not factual questions compared to those who wrote them by hand.

I: Taking notes on a laptop negatively affects performance in response to conceptual questions compared to taking notes by hand, which could be due to the fact that taking notes on a laptop seems to encourage shallow processing, such as copying the material verbatim.

Paper 2: Foss-Feig, J.H., Tadin, D., Schauder, K.B., and Cascio, C.J. (2013). A substantial and unexpected enhancement of motion perception in autism. *Journal of Neuroscience,* 33 (19), 8243–8249.

Q: Do autistic individuals exhibit enhanced motion perception? Are there reliable psychophysical markers of autism that could transcend subjective diagnostics?

D: Groups of autistic children and controls were shown moving gratings varying in size and contrast. The authors measured how long the stimuli had to be shown until perceived accurately by the study participants.

A: It is believed that autistic individuals might suffer from a deficit in inhibitory neurotransmitters. If this is the case, they should show less spatial suppression, which is believed to rely on inhibition, and therefore be quicker to see large stimuli.

F: Autistic individuals are more sensitive to perceive motion for all stimulus sizes, but they are not better at low contrasts.

I: It is not the case that autistic individuals show less spatial suppression – rather, the low contrast results suggest that there is abnormal gain control in autism.

The authors neither measure nor report reaction times; as autistic individuals are known to exhibit longer reaction times, it cannot be ruled out that the results are simply due to a speed/accuracy tradeoff. This does not mean that duration thresholds cannot be used as a psychophysical marker of autism, but it might not be for the reasons the authors think. (For a more detailed discussion of this issue, see Wallisch and Bornstein, 2013.)

Conclusion

For students, doing a QDAFI is hard – really hard, especially in the beginning. But they should not be discouraged. The best way for students to learn this method is simply to do it. They will improve with practice, which is well worth the effort. If it isn't already clear from the foregoing discussion, rest assured that reports so far suggest that learning how to write a QDAFI is an extremely valuable skill. Students in classes where I taught the QDAFI have emailed me, sometimes years later, to say this was indeed the most valuable thing they learned. I am currently trying to figure out how to ethically study the effectiveness of the QDAFI empirically, which is tricky. As we know how helpful the QDAFI is, it would be unethical to teach one group of students how to do this, but not another, and then test their retention of articles at the end of the class. Conversely, if we required groups of students to do QDAFIs versus free-form summaries of papers in alternating weeks, we would have to teach the QDAFI to both groups, and could not rule out that the free-form group also implicitly applied the QDAFI method in any given week. Thus the comparison would perhaps only be valuable for the very first week, before the free-form group was taught how to do the QDAFI. But that seems like a lot of work for little useful data.

Regardless, I believe that it works, due to the principles laid out above, and because of my experience using the QDAFI myself, as well as my experience teaching it to my students. I hope you will find this method useful as well.

Acknowledgments

This method was originally inspired by Stephen Kosslyn's "QuALMRI" technique.

References

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University.

Clark, H. H., and Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition*. Ed. L.B. Resnick, J.M. Levine, and D.T. Teasley. American Psychological Association.

Craik, F.I., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104 (3), 268–294.

Foss-Feig, J.H., Tadin, D., Schauder, K.B., and Cascio, C.J. (2013). A substantial and unexpected enhancement of motion perception in autism. *Journal of Neuroscience*, 33 (19), 8243–8249.

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1989, May 26). Delay of gratification in children. *Science*, 244 (4907), 933–938.

Mueller, P.A., and Oppenheimer, D.M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, 25 (6), 1159–1168.

Nuijten, M.B., Hartgerink, C.H., Assen, M.A., Epskamp, S., and Wicherts, J.M. (2015). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, 48 (4).

Open Science Collaboration (OSC). (2015, August 28). Estimating the reproducibility of psychological science. *Science*, 349 (6251), aac4716.

Reyna, V.F., and Brainerd, C.J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7 (1), 1–75.

Roediger, H.L., and Karpicke, J.D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*, 17 (3), 249–255.

Wallisch, P., and Bornstein, A.M. (2013). Enhanced motion perception as a psychophysical marker for autism? *Journal of Neuroscience*, 33 (37), 14631–14632.

Wallisch, P. (2015). Brighter than the sun: Powerscape visualizations illustrate power needs in neuroscience and psychology. arXiv preprint arXiv:1512.09368