Theory, Logic and Data: Paths to a More Coherent Eyewitness Science

GARY L. WELLS*
Iowa State University, USA

SUMMARY
Building a better scientific literature on eyewitness reliability is a continuous process. The articles in this special issue of *Applied Cognitive Psychology* are timely and well-reasoned reflections on the need to learn from the history of the area, recognize the unique ways in which the legal system operates, and develop better theory. The current article comments on these articles and embellishes on some of these points. Copyright © 2008 John Wiley & Sons, Ltd.

*No, a thousand times no; there does not exist a category of science to which one can give the name applied science. There are science and the applications of science, bound together as the fruit to the tree which bears it* (Louis Pasteur, 1871).

Pasteur’s quote was the first thing that came to mind when I was asked to write a commentary on the articles appearing in this special issue of *Applied Cognitive Psychology*. At some level I can agree with Pasteur’s point. But to adopt this fruit–tree metaphor would gloss over some very important points that the authors have made in this special issue devoted to the basic versus applied memory research debate. The authors of these articles, each highly accomplished, leading thinkers in the eyewitness area, provide valuable insights that can help guide the continued development of a useful science of eyewitness behaviour.

THEMES

History
There are three main themes that run through these articles. One is a historical theme. Bornstein and Penrod (2008) take the reader back to G.F. Arnold and compare and contrast his early writings with Hugo Munsterberg. Brewer and Weber (2008) take the reader back to early research on the accuracy of sensory judgements and the early use of reaction time measures as indicators of psychological processing. Clark (2008) notes the historical

*Correspondence to: Gary L. Wells, Department of Psychology, West 117 Lago Hell, Iowa State University, IA 50011, USA. E-mail: glwells@iastate.edu

Copyright © 2008 John Wiley & Sons, Ltd.
context in which his current mathematical modelling of eyewitness identification is grounded. Sporer (2008) describes the prominence of European contributions to eyewitness issues that preceded developments in North America. The authors were not asked to provide historical perspectives, but it is clear that they all recognize that there are lessons in these histories. And, it seems to me that one of the main lessons is that the research and theory problems that modern eyewitness research has identified are remarkably similar to the problems of focus 80, 100, or even 120 years ago. I will not comment more on the history theme except to note that students and those newer to the eyewitness area perhaps should be required to read this history.

Theory

A second theme running though these articles is that basic research and theory are essential to the eyewitness area. Each article tends to lament the way that eyewitness theory development has tended to take a back seat to data. Each author tends to have their own type of response to this theory–data gap. Brewer and Weber (2008) and Deffenbacher (2008), for instance, argue that existing theories from basic areas of cognitive psychology can provide a framework. Clark (2008) articulates the value of mathematical modelling as a critical way to clarify and test theoretical claims. Lane and Meissner (2008) note how theory, along with converging operations and field studies, can bridge the applied–basic divide. Memon, Mastroberardino, and Fraser (2008) nicely describe the importance of the causal versus associative relation distinction. Sporer (2008) introduces a model of eyewitness testimony for organizing research. Turtle, Read, Lindsay, and Brimacombe (2008) argue effectively for the use of basic theory and principles to consolidate the wide range of phenomena in the eyewitness area and argue for more emphasis on process rather than phenomena.

Eyewitness researchers would do well to pay attention to these carefully-reasoned calls for theory. But, it seems to me that Clark (2008) made a very interesting observation when he wrote that eyewitness theories have in fact been proposed, but they have not been well-developed. In some cases, such as with the automatic–deliberative conceptualization of Brewer, Weber, and Semmler (2007), the dual processing theory of Meissner, Tredoux, Parker, and MacLin (2005), or the competition/corroboration theory of Charman and Wells (2007), the theories have not had much time to mature and develop. But, Clark cites the relative judgement conceptualization, originally proposed nearly 25 years ago (Wells, 1984), and notes that it remained largely undeveloped, until very recently, since its original conceptualization. It is not my intent to defend my writings from nearly 25 years ago, but it seems to me that the relative-judgment conceptualization of eyewitness identification illustrates how useful a simple theory can be, even when it is not fully developed. The relative-judgment conceptualization brought together, in a simple explanatory framework, various divergent observations that preceded it, such as the surprising propensity for eyewitnesses to select someone from target-absent lineups, the reason for the pre-lineup warning that the culprit might not be present (to try to reduce the witness’s reliance on mere

1I would state this point somewhat differently. Instead of ‘focus on processes rather than phenomena’, I would say ‘focus on the processes of phenomena’. Phenomena are essential because demonstrable, replicable phenomena give us something to theorize about. And, some phenomena, such as the post-identification feedback effect, own-race bias, pre-lineup instruction effect, removal-without replacement effect, and weapon focus, to name just a few, are so striking and central that they naturally lead researchers to seek process interpretations. And, the proper test of a process interpretation is that it explains already-established phenomena and predicts new phenomena.
relative judgements), the reason that witnesses sometimes identify an innocent suspect who
does not look very much like the target (he looks more like the target than did the other
lineup members), and the reason the fillers in a lineup need to be carefully selected (so that
an innocent suspect is not automatically the best ‘relative’ choice). But the relative-
judgement conceptualization also led to predictions about things not yet observed, such as
(a) the greater reluctance to select someone from a show-up than from a lineup, (b) the
different processes that witnesses report when they identify the correct target (‘pop out’
and non-comparison language) than when they identify an innocent suspect (comparison
language), (c) the potential superiority of the sequential procedure over the simultaneous
procedure and (d) the removal-without-replacement effect. For an underdeveloped theory,
the relative judgment idea has had both a long life and considerable explanatory and
predictive power. My point is not to defend the relative-judgement conceptualization, but
instead to note that simple theories, even if they are underdeveloped, can help organize
divergent empirical observations and make testable predictions.2

Clark’s (2008) idea that computational modelling can be a key tool in testing and
refining theory is, in my opinion, one of the most important developments in the last decade
of eyewitness research. Computational modelling is not for everyone. I have long held the
view that the size of a reading audience in psychology and law is cut in half for every
formula or equation that is placed in an article. In our 1980 Psychological Bulletin
article on eyewitness identification (Wells & Lindsay, 1980), we had 12 main equations plus
numerous proofs. It is among the most important articles that I have ever written, but
almost no one has ever read it. I hope we are entering a new era in which our students
become comfortable with computational modelling and the power of mathematics to guide
our thinking on these problems.

The basic–applied fit

Another theme noted in many of the articles is that there are some important differences
between the nature of basic research on memory, judgement and decision making and the
legal setting to which the work might be applied. Lane and Meissner (2008) discuss the
importance of the interaction between field and lab research. Students and new eyewitness
researchers will find the ‘study space’ idea of Malpass et al. (2008) a fruitful one for
identifying critical gaps in the empirical work as it maps on to real-world conditions
outside the lab.

I believe that it is important to note two ways in which the applied eyewitness problem
does not mesh perfectly with the traditional basic research. First, the basic methods of
scientific psychology, which yield group probabilistic statements, are not naturally suited
to the individualistic fact-finding mandate that characterizes the courtroom trial. As Doyle
(2005) notes, a researcher might be able to say that X happens 80% of the time and Y
happens 20% of the time, but the fact finder at trial wants to know if this particular witness
is one of the eight or one of the two. Along these lines, Turtle et al. (2008) note that the legal
system wants to know ‘How reliable is this particular witness?’, but researchers ask ‘What
factors affect eyewitness evidence?’, It is as if an agronomy scientist were asked ‘What is
the exact size of the fruit that a specific bud from this tree will bear?’, whereas the scientists

2Indeed, a theory can even be wrong and yet be useful because it is better than any other available theory at the
time. The value of a theory is its ability to explain and organize past findings and make new predictions.

DOI: 10.1002/acp
I have argued that this individualistic versus probabilistic problem of eyewitness application tends to go away when the emphasis of the research is on system variables, at least when they are applied to improving methods for collecting eyewitness evidence, rather than on estimator variables, when estimators are used to postdict errors. The reasoning is simple: The purpose of system–variable research is to minimize the probability of eyewitness error across cases rather than to decide post hoc whether a particular eyewitness instance was an error or not. Hence, in some ways the fit between the methodological nature of our science and the needs of the legal system is a better fit for system (prevention) variables than it is for estimator (postdiction) variables (see Seelau & Wells, 1985 for a fuller development of this point).

A second way in which the eyewitness problem does not perfectly mesh with traditional basic research concerns what I have called the directional-condition problem. Basic research, traditional statistics and training in scientific psychology almost universally focus on estimating the probability of a particular behavioural response to the presence or absence of a particular stimulus. For example, a recognition memory study assesses the probability of hits, false alarms, correct rejections and false rejections. In each case, the direction of the calculated conditional probability can be represented as the probability of a response given the stimulus or \( p(R|S) \). In face recognition, for instance, the hit rate is calculated as the probability of saying ‘yes’ given that the stimulus was the previously-viewed face and a false alarm is the probability of saying ‘yes’ given that the stimulus was not a previously-viewed face. But the problem for the trier of fact (e.g. judge or jury) is represented as the opposite direction of this conditional. Specifically, the trier of fact needs to know the probability of the stimulus given a particular response or \( p(S|R) \). These are not parallel problems. In the lab, we know whether the witness was shown an old face or a new face and the interest is in assessing the probability of a particular behavioural response. But in the courtroom, we know the behavioural response of the witness and the task is to estimate the probability that the face was old or new. The distinction is important because, except under rare circumstances (where relevant base rates happen to be exactly 50/50), the value of any conditional probability \( p(A|B) \) is not the same as \( p(B|A) \). We first articulated this problem in the eyewitness area nearly 30 years ago (Wells & Lindsay, 1980) and have revisited it a number of times since. Mathematically, the solutions are best represented with Bayesian statistics, which make it clear that a major factor, and perhaps the major factor, governing the chances of mistaken identification is the base rate for culprit-present versus culprit-absent lineups. Turtle et al. (2008) make this point and argue that, in the absence of target-present versus target-absent base rate information, an expert would not be able to testify about the probability that an identified suspect was mistakenly identified no matter how well the lab conditions mimicked the specific case. Brewer and Weber (2008) also make this same point when they note that ‘the basic researcher asks—given that the stimulus is old/new—how likely is the participant to recognize it as such’,

\[ Consider, for example, the probability that someone is a male given that they are taller than 6 feet 6 inch versus the probability that someone is taller than 6 feet 6 inch given that they are male. The former probability is over 99% and the latter probability is less than 1%. In this example, the very low base rate for anyone (male or female) being over 6 feet 6 inch tall creates these asymmetric probability estimates. For the same reason, a laboratory-obtained probability that an eyewitness will identify the suspect given he is innocent might be 5%, but the probability that the suspect is innocent if identified by an eyewitness might be 80% (or any of a family of possible probabilities, depending on the base rate for culprit-absent lineups). In this example, a very low base rate for the lineup containing the actual culprit creates these asymmetric probability estimates. \]
which does not fit the problem presented to the trier of fact. Indeed, the entire Signal Detection Theory approach is founded on this particular direction of the conditional [i.e. $p(R|S)$], which is the reverse of the conditional probability question confronting the trier of fact in an eyewitness case. Pressing Pasteur’s metaphor a bit further, it seems to me that the basic research tree cannot bear the fruit of application without appreciating fully that real-world postdiction is Bayesian in nature and this requires some deviation and extrapolations from Signal Detection statistics and other standard inferential and descriptive statistics that dominate basic psychological science.

THE APPLIED VALUE OF BASIC SCIENTIFIC LOGIC

Although all the articles in this special issue touted the value of theory development, there was a conspicuous absence of promoting the value of scientific logic itself as an applied tool. Consider, for example, the lineups-as-experiments analogy that was articulated nearly 20 years ago (Wells & Luus, 1990). The idea simply posits that a police lineup is like a psychology experiment. The detective is like an experimenter who has a hypothesis (that the suspect is the culprit). The detective tests the hypothesis with a design (embed suspect among fillers) and a procedure (e.g. pre-lineup instructions) using participants (wit...
purpose of the hearing was to determine whether I could testify as an expert for the defence on the eyewitness identification procedure. A critical issue in the case was that the suspect’s photo in the photo-lineup was a colour photo and all the filler photos were black and white. I proffered an opinion in the hearing that the lineup was suggestive based on the broader principle that the suspect’s photo should not stand out. Ebbe Ebbesen then testified that I should not be allowed to proffer such an opinion because there were no empirical lineup studies that had been conducted examining the effect of having one colour photograph embedded among five black and white photographs. In Malpass et al.’s (2008) terminology, this was an empty cell in the study space. I explained that the opinion was based on a general principle that the suspect’s photo should not stand out, not on a specific study, and that the absence of such a study in the empirical literature probably reflected a belief by researchers that law enforcement would never create such a lineup and, hence, such a study would likely be criticized by journal reviewers for being absurdly unrealistic. The judge ruled that my testimony was admissible and in his written opinion noted:

‘There is good reason that social scientists have not done studies using photographic lineups where the only colour picture used is that of the suspect’s... because the probability of police using such a lineup would be too low. In any event, Dr Wells can safely extrapolate principles from studies... The “fit” requirement of Daubert is not intended to limit all applications of scientific theories to the data which produced them. In some situations, a principle is helpful even if there is not a study directly on point’.

(pp. 24–25, footnote omitted).

This judge’s reasoning indicates that the legal system (or at least US Magistrate Judge Robert P. Murrian) is on the same page as the authors in this special issue of *Applied Cognitive Psychology*. The case represents a good example of what Lane and Meissner (2008) meant when they stated that ‘theory allows one to make reasonable predictions about eyewitness memory even for specific situations that have never been encountered in the laboratory or in field research’. Empirical data must be the foundation of our expertise, but conceptual models, process understanding and theory are the keys to a useful application of psychological science to the eyewitness problem.

**REFERENCES**


