

What do We Know about Facial Cognition? What Should We do with this Knowledge?*

Michael J Wenger and James T. Townsend, **Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges**. Mahwah, NJ: Lawrence Erlbaum Associates, 514 pages, ISBN 0-8058-3234-3, \$150.00

Review By:
Geoffrey R. Loftus

■ *Michael J. Wenger is an Assistant Professor at the University of Notre Dame and is co-author of Cognitive psychology. • James T. Townsend is Rudy Professor of Psychology at Indiana University, is a former editor of the Journal of Mathematical Psychology and is co-author of The Stochastic Modeling of Elementary Psychological Processes. • Geoffrey R. Loftus is Professor of Psychology at the University of Washington.*

Face processing, and face recognition in particular, has acquired quite a bit of cachet over the past few years. Two main categories of newsworthy events have caused this to happen. First, there has been a series of advances in computer-aided face recognition. For example, a New York Times, article headlined, "The Face of Security Technology" that appeared on January 20, 2002, is summarized as, "Joseph J Atick, chairman and chief executive of Visionics, most visible spokesman for technologies that identify individuals by their physical characteristics, field known as biometrics; he drew criticism after Sept 11 for asserting confidently that at least two of hijackers could have been intercepted at Logan Airport in Boston had facial recognition technology been deployed there; Visionics also gained notoriety in summer of 2001 when Tampa, Fla, police used its technology to search via closed-circuit cameras for wanted criminals among weekend crowds." Although practical applications of automated face recognition have been appreciated for some years, interest in it ratcheted up following the terrorist attacks of September 11, 2001, as the problem of recognizing objects in general (e.g., weapons at airport checkpoints) and faces in particular (e.g., of potential terrorists) was suddenly thrust to the forefront of national security.

The second category of newsworthy face-recognition-related events involves eyewitness testimony, examples of the fallibility of which are provided by numerous news accounts over the past few years. Many people—psychologists and lawyers in particular—have known for years that deficiencies in human face-recognition ability has seriously compromised the criminal justice system: These deficiencies have led to false convictions, years unfairly spent in prison, and executions of innocent people. A recent accelerating agent for interest in face-recognition issues was a landmark book about DNA exonerations of falsely convicted individuals (Scheck, Neufeld, & Dwyer, 2000) where

the authors noted that, "In a study of DNA exonerations by the Innocence Project, 84% of the wrongful convictions rested, at least in part, on mistaken identification by an eyewitness or victim."

So amidst this explosion of popular interest in face recognition comes *Computational, Geometric, and Process Perspectives on Facial Cognition*, edited by two highly respected researchers. The book, despite its somewhat intimidating title, is mostly readable, contains a wealth of valuable information, and certainly belongs on the bookshelf of any serious student of face recognition in particular and facial cognition in general.

Following are two quite distinct sections. In the first, I will provide a straightforward description of the book—what it's about, how it's organized, and how it hangs together. In the second section, I will describe my personal wish list that the book fostered and, as suggested by my allusions to newsworthy aspects of facial cognition, will focus on potential practical applications of facial-cognition work.

Structure and Synopsis

The book includes 12 chapters written by a total of 23 contributors. It is not a handbook—it isn't the source you'd choose if you were seeking a systematic, exhaustive review of, say, face-inversion effects. Although it provides quite a bit of new data, the book's main strength, which is considerable, is that among them, its chapters describe the major cutting-edge quantitative theories of a wide range of face-processing capabilities: face recognition, face discrimination, expression recognition, morphs, caricatures—you name it, it's discussed somewhere in the book.

The book covers far too wide a range of topics to even sketch them all in a brief review. I would advise

the inquisitive reader to read the first and last chapters first. Chapter 1, to be described in more detail below, sets forth the book's major themes, while the last chapter ("Are Reductive [Explanatory] Theories of Face Identification Possible? Some Speculations and Some Findings") by Uttal attempts to describe limitations which the reader would do well to keep in mind while reading everything in between.

Three major themes emerge. The first, and by far the most pervasive, involves face space; the second involves the holistic/feature-by-feature distinction (read "controversy") and the third involves generalizing from few to many viewpoints of a face.

Chapters 1-6: Face Space

Chapter 1 ("Quantitative Models of Perceiving and Remembering Faces" by O'Toole, Wenger, and Townsend) gets right down to business, describing face space—a concept which then wends its way through many of the book's subsequent chapters and which is really the main theme around which most of the book is organized. A face space begins with the observation that all faces differ from one another along various dimensions—both with respect to the physical characteristics of the faces themselves, and with respect to the way in which the faces are represented in people's minds. After noting a variety of ways in which people must process and use face appearances (e.g., remembering faces, judging various aspects of faces such as attractiveness or typicality) the authors point out that any of these tasks, "requires a system capable of comparing individual faces with information structures representing other individual faces and with groups of (presumably) known faces in memory. In both the computational and psychological face literatures, the most common theoretical framework for doing this relies on the abstract notion of a face space."

The authors then go on to list three kinds of face spaces that have been proposed. These are:

Abstract Face Space. In an abstract face space, an individual face is conceptualized as a point in a multidimensional space. The difference between two faces (obviously an important concept for much of face processing) is then represented by the distance between the two points corresponding to the two faces. Although not explicitly stated as such, the authors appear to provide two instantiations of an abstract face space which are the following.

Psychological Face Space. A psychological face space is one in which the face representations are based on some form of behavioral data such as a set of similarity ratings among a set of faces—which can then be transformed into an actual face space via multidimensional scaling procedures. Such a process yields direct representations of perceptual similarity among faces.

Physical Face Space. A physical face space is one whose dimensions are based on physical facial features or relations among facial features (which might be, for instance, ratio of face height to width, interocular distance, and so on). Physical face spaces begin with a set of images and use as fundamental data measures of physical attributes. In principle, one could use any set of physical attributes as the dimensions (e.g., one dimension could be nose length; another could be distance between the centers of the two pupils, and so on). In practice, however, the features are usually derived more abstractly using various forms of linear-systems analyses, the simplest of which is principal components analysis.

Face Space Expanded in Subsequent Chapters. Obviously, the concept of a face space is a broad one. The reader interested in a more detailed analysis of different face-space types might be advised after reading Chapter 1 to take a brief detour to Chapter 9 ("Is All Face Processing Holistic?" by Cottrell, Dailey, Padgett, & Aldophs) where a kind of "metaspace of face spaces"—actually a space of potential face-space features—is provided in their Figure 1. In Chapters 2-6, numerous authors describe numerous uses of face space which provides a front end upon which other information-processing models can operate. In what follows, I'll attempt to hit the high points of the relevant chapters.

Chapter 2 ("The Perfect Gestalt: Infinite Dimensional Riemannian Face Spaces and Other Aspects of Face Perception" by Townsend, Soloman, & Smith) is probably the most mathematically challenging treatment in the book. Essentially, it describes the usefulness of a kind of dimensional anchor point: conceptualizing each face in terms of a complete description of the face (hence the "infinite dimensions"). I will leave it to the ambitious reader to digest it (slowly) because I did not fully understand it, and I do not want to misrepresent it. (Indeed, Jim Townsend admitted to me that even *he* had some difficulty following this chapter after he hadn't seen it in awhile.)

Chapter 3 ("Face-Space Models of Face Recognition" by Valentine) brings the reader back to Planet Earth with elaborations of some of the basic face-space issues raised in Chapter 1 (one wonders why Chapter 3 didn't follow Chapter 1 directly. Such a scheme would, to this reviewer anyway, have made more organizational sense). Among many other topics, Valentine here introduces the notion of "distinctiveness" which, when viewed within the context of face space, forms another major theme to which subsequent chapter writers routinely return: For instance in Chapter 5, Busey shows how the cross-racial recognition effect (people recognize members of their own race better than members of other races; e.g., Malpass & Kravitz, 1969) can be couched within the notion of distinctive regions of face space. More generally, Valentine describes recognition-memory effects both within the context of face space

theories and, to some degree, within the context of alternative non face-space theories.

Chapter 4 ("Predicting Similarity Ratings to Faces Using Physical Descriptions" by Steyvers & Busey) describes three methods for defining face features: geometric attributes (e.g., eye separation, mouth width), principal components, and the output of a system of gabor filters ("gabor jets") applied to various parts of the face. These features are then used as input to a connectionist model whose purpose is to reduce the number of face-space dimensions and ultimately to account for empirical similarity ratings among pairs of faces.

Chapter 5 ("Formal Models of Familiarity and Memorability in Face Recognition" by Busey) takes up the sub-themes of face recognizability and distinctiveness by connecting face spaces to a specific recognition sampling model (the SimSample model). As noted earlier, Busey is one of the few authors in this volume to actually tackle an important practical issue (the cross-racial effect) within the context of a quantitative theory.

Chapter 6 (Characterizing Perceptual Interactions in Face Identification Using Multidimensional Signal Detection Theory" by Thomas) provides a useful summary of Thomas's previous work on multidimensional signal-detection theory in the process of applying it to face perception. As one might expect, the application of multidimensional signal-detection theory to anything-space and to face-space in particular makes for a harmonious fit.

Chapters 7-9: The Holistic-Feature-by-Feature Issue

Chapters 7 ("Faces as Gestalt Stimuli: Process Characteristics" by Wenger and Townsend), 8 ("Face Perception: An Information Processing Perspective" by Campbell, Schwarzer & Massaro), and 9 (Is All Face Processing Holistic? The View from UCSD by Cottrell, Dailey, Padgett, and Adolphs) introduce a new major theme: whether face processing is done holistically or feature-by-feature. Importantly, these chapters also introduce definitions of "holistic" couched within specific quantitative models which is welcome given the pervasive vagueness with which the term has been tossed about in the past. These definitions will not end the debates about this particular face-processing dichotomy as one can, at the very least, continue to argue about (1) whether the definitions implied by the various models are appropriate, (2) which aspects of face processing may be considered (by any definition) to be done in one way or the other, and (3) what is the relation between "holistic" and "gestalt"?

Chapters 10-11: Viewpoint Independence

Chapter 10 ("Viewpoint Generalization in Face Recognition: The Role of Category-Specific Processes"

by Edelman and O'Toole) and 11 ("2D or Not 2D? That is the Question: What can we Learn from Computational Models Operating on Two-Dimensional Representations of Faces" by Valentin, Abdi, Adelman, and Posamentier) address viewpoint independence. The major question addressed in both these chapters is: What kinds of two-dimensional information is required such that the observer (human or computer) have sufficient knowledge to carry out tasks requiring three-dimensional information? Edelman and O'Toole focus mostly on data involving such generalization, whereas Valentin et al. focus more on a particular, autoassociative model of the process.

What's Missing Here?

The book is terrific in terms of bringing the reader up to date on extant quantitative and computational models of face processing. One of its (ironic) virtues is that it left me somewhat frustrated, pondering the kinds of information I'd still like to see addressed. Lack of space prevents me from enumerating them all; however the one that leaps most quickly to mind is more connection to the real world or real problems.

I began this review by mentioning some newsworthy applications of facial cognition. These applications are by no means exhaustive. Knowledge of facial cognition is critical in terms of automated recognition (e.g., of terrorists or missing children); of how to "artificially age" faces (e.g., so that a child kidnapped and last photographed at age 3 might be recognizable 13 years later at age 16); of how to construct a lineup so that a witness's memory for some perpetrator is tested in a well-defined optimal manner (more on this below), and so on. In reading the book, I was struck at how almost entirely lacking it was in acknowledging such real-world problems and in providing any clues as to how relevant face-processing research might illuminate them. There are some exceptions; for example Busey (Chapter 6) does at least take note of the cross-racial identification effect, which he then describes within the context of distinctiveness in face space. However, this nod to practical problems is the exception rather than the rule in this book (and, I note, nothing about the cross-racial effect even appears in the index.)

Cognitive psychology, while not exactly overflowing with quantitative application of theory to real-world problems does include at least some good examples of such applications including, for example, applying mathematical theories of learning and memory to optimizing learning in many educational domains (e.g., Atkinson, 1972) and applying human-factors theories to optimize instrument-panel design (e.g., Green, 1989). It is reasonable to suppose that carefully crafted mathematical and computational models of facial cognition could serve similar purposes.

An optimistic reader might have hoped to find in this book such real-world connections as (1) copious

references to the eyewitness testimony literature (e.g., Wells, 1993) and (2) names and addresses of companies manufacturing race-recognition software, along with some indication of what such companies are trying to accomplish, how they are accomplishing it, and what, if anything, their efforts have to do with the facial-cognition data and theory that the book so artfully addresses. I do not by any means feel that all scientific research must be driven entirely by ecological considerations (see Loftus, 1983) but I do believe that when there are practical aspects of some research area staring us in the face, they should be acknowledged. By contrast to this book, another new book about facial cognition (Rakover & Cahlon, 2001; see also Rakover & Cahlon, 1989) includes eyewitness testimony as a major real-world problem as a central theme.

I now pose a challenge to facial-cognition researchers. Solving the problems entailed in this challenge would provide strong benefits to society while at the same time, I believe, providing validation of the facial-cognition work itself. I have chosen the particular challenge because it involves issues with which I am familiar. It's certainly not the only problem whose solution could be based on set of algorithms issuing from facial cognition theory.

A common situation in law enforcement is this. A crime takes place (say a convenience store robbery). A witness (say the store clerk) has a good enough look at the perpetrator to possibly be able to identify him later. Within minutes following the crime, the police find some plausible suspect—someone fitting the description provided by the witness—in the vicinity of where the crime occurred.

Suspect in hand and witness hovering in the wings, the police now have two potential courses of action. The first is to create a *lineup*—either a live lineup or a photo lineup. As most readers probably know, a lineup is a collection of people, or photos of people. One member of the lineup is the suspect while the other members, called *fillers*, are individuals known to have nothing to do with the crime. There are usually five fillers, so the lineup usually contains six individuals in all. When a photo lineup is used, the photos are typically face shots only. The witness, after viewing a lineup makes an initial decision about whether to identify anyone and if so, decides which lineup member to identify. In memory-research lingo a lineup procedure is, of course, a six-alternative, forced-choice recognition test.

The second course of action open to the police is to carry out what is called a *showup* procedure. In a showup, the witness is brought to where the suspect is being held and asked to make a yes-this-is-the-perpetrator decision (a positive identification) or no-this-isn't-the-perpetrator decision (no identification). A showup procedure is, of course a form of old-new recognition test.

Readers of this review will have no difficulty realizing that a lineup procedure is better than a showup procedure. A lineup, assuming it is done correctly, is a genuine test of memory: The witness chooses the suspect only if the witness's memory of the perpetrator matches the suspect's appearance better than any of the fillers' appearance. In a showup procedure, by contrast, the witness's inclination to positively identify the suspect may depend in part on the match between the witness's memory and the suspect's appearance but, as is well known from decades of signal-detection research, the decision may depend on many other things as well, such as witness's expectations that the suspect is the perpetrator, the witness's motivation to identify *someone*, pressure on the part of the police officer who is administering the procedure, and the witness's bias to say "yes" or "no".

Courts as well as cognition researchers recognize these difficulties, and showup procedures are frowned on. Nevertheless, showups are by no means forbidden, and police use it routinely in the kind of situation just described. The rationale for doing so is threefold. First, a showup is expedient—let's face it, it's a hassle to get together five reasonable filler photos to create a photo lineup (and don't even talk about finding five live individuals for a live lineup, which propels hassle to a formidable new level). Second, a showup can be done immediately when the witness's forgetting of the perpetrator's appearance is minimal; by contrast, getting the witness to view a lineup of any sort usually takes a day or so, minimum. Third, a non-identification is considered grounds for releasing the suspect from custody, which minimizes the inconvenience suffered by the suspect.

How could one combine the advantages of lineup and showup procedures? I suggest the "Instant-lineup system," which begins with the observation that increasingly, police cars are equipped with a variety of high-tech equipment, including laptop computers, digital cameras and internet connections. Given this and related technology, it would be relatively simple, in principle, to digitally photograph the suspect and then create an on-the-spot photo lineup that would include the suspect's photograph plus five computer-selected or computer-generated fillers.

At this point, we come to the challenge for facial-cognition researchers which is: What should the fillers be? There is a good deal of research and information about how fillers should be selected (e.g., Buckout, 1974; Wells, 1993; Wells & Seelau, 1995). For example, it is agreed that (1) the fillers should match the witness's description of the perpetrator and (2) the suspect's face should not be perceptually or conceptually different from the fillers, considered collectively: For instance, it would be bad to have the suspect wearing prison garb and all the fillers wearing normal clothing or vice-versa; it would likewise be bad to have the sus-

pect's photo be darker or lighter or larger or smaller or have a different background from the fillers' photos. However, this advice, while valuable as far as it goes, stops well short of providing formal algorithms for how create fillers.

In the proposed instant-lineup system, there are two potential sources of fillers, both of which require initially downloading the suspect's photo, along with the witness's description and generating an analysis of both the suspect's features (however defined) and the description. The first source of fillers entails access to a database of faces and associated face-feature lists from which appropriate fillers would be chosen. The second possible source is to *generate* fillers and feature lists from scratch, on the fly, again based on the suspect's face and the witness's description.

It is quite clear that, whichever of these schemes (or any other plausible scheme) were used, quite a bit of facial-cognition theory and theory-implementation technology would be needed which would minimally include the following.

1. A specific system for defining facial features, along with the technology for extracting the features from a digitized photo. A variety of very clever such suggestions have been provided by many of the contributors to this book; one test among them would be the degree to which one or the other succeeded in this enterprise.
2. A means of translating a (sometimes vague) verbal description provided by a witness into equivalent-format facial features. This, of course, would place serious constraints on the type of features described in Point (1) above. For instance, gabor jets certainly provide a feature system for actual, visual faces, but the question of how gabor-jet based features could be transformed to a common format for images and verbal descriptions is at best unclear and at worst impossible.
3. Accepted rules for constructing or selecting fillers (e.g., Wells, 1993).
4. And last, but decidedly not least, a theory of what one is trying to optimize in this kind of identification procedure. Such a theory which, in principle, should constitute the formal foundation for the rules indicated in (3) above is, at present, sorely lacking in either the scientific or the forensic literature. An optimization theory would necessarily go beyond facial cognition in scope, beginning with the question: What exactly should be optimized when a witness is trying to identify some previously viewed perpetrator? Overall proportion correct? Hit rate? Correct rejection rate? This question, while easily couched within the familiar domains of recognition memory and signal detection must be an-

swered within the context of issues that are fundamentally legal and philosophical. Once the optimization basis is decided upon, the question becomes: How does one achieve such optimization by selecting or generating appropriate fillers in a photo lineup?

This and other related problems are crying out for clever solutions. So go to it, facial-cognition theorists! Channel some of the impressive creativity and energy that you have so convincingly demonstrated in this book into solutions of some of the many genuine, very pressing, and intellectually challenging problems of today's society. Solving them will validate facial-cognition theory, it will offer a public image of Psychology that's an alternative to the standard little bearded Freudian, and it will give you something interesting to talk about with your airplane seatmates when infinite-dimensional Riemannian spaces cause their eyes to begin glazing over. Moreover, if you get the solution right, you can patent it and make a ton of money! Why not? It's the American way!

References

- Atkinson, R.C. (1972). Ingredients for a theory of instruction. *American Psychologist*, 27, 921-931.
- Buckout, R. (1974). Eyewitness testimony. *Scientific American*, 231, 23-31.
- Green, P. (1989) Instrument panel displays and ease of use. *UMTRI Research Review*, 19, 1-14.
- Loftus, G.R. (1983). The continuing persistence of the icon. *The Behavioral and Brain Sciences*, 28.
- Malpass, R.S. & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, 13, 330-334.
- Rakover, S.S. & Cahlon, B. Race Recognition (2001). *Cognitive and computational processes*. Amsterdam: John Benjamins.
- Rakover, S.S. & Cahlon, B. (1989). Face Recognition: To catch a thief with a recognition test: The model and some empirical results. *Cognitive Psychology*,
- Scheck, B. Neufeld, P., & Dwyer, J. (2000). *Actual Innocence*. New York: Doubleday.
- Wells, G. L. & Seelau, E. P. (1995). Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law*, 1, 765-791.
- Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48, 553-571.

