Table 1 Informational requirements of a variety of facial processing tasks.

Task	*Informational Requirements*	*Citations*
Recognition	Information (e.g., episodic) specific to a target face, global familiarity of a target face, similarity of a presented face to all stored representations	(Ashby & Perrin, 1988; Hintzman, 1988; Massaro, 1998; Murdock, 1993; Nairne, 1990; Nosofsky, 1988a, 1991; Shiffrin & Steyvers, 1997b; Townsend & Landon, 1983)
Recall	Information (e.g., semantic) specific to a target face, including information supporting the expression (e.g., via retrieval or redintegration) of labels (e.g., names) associated in prior experience	(Hintzman, 1988; Massaro et al., 1991; Schweickert, 1993; Thomas, 1996)
Detection	*Feature detection*: derivation of structures or characteristics that comprise the feature within the face. *Face detection*: Generic facial structure, features shared by all faces	(Ashby & Maddox, 1993; Maddox & Ashby, 1993; Massaro, 1998; Nosofsky, 1986, 1991; Townsend et al., 1984; Townsend & Nozawa, 1995)
Categorization	Information in the target face that is shared with some or all faces in a particular category	(Ashby & Perrin, 1988; Ashby et al., 1994; Ashby & Alfonso-Reese, 1995; Cohen & Massaro, 1992; Nosofsky, 1986, 1988b; Nosofsky & Palmeri, 1997; Palmeri, 1997)
Facial expression	Information pertinent to particular features or sets of features that are regularly associated with affect, emotional status, states of intent (e.g., aggression), etc.	(Ellison & Massaro, 1997), see citations for Categorization
Same-different	Relative proximity of two faces along featural dimensions or equivalence of labels associated with two faces	(Ashby & Schwartz, 1996; Thomas, 1996), see citations for recognition

Vetter, T. & Troje, N. F. (1997). Separation of texture and shape in images of faces. *Journal of the Optical Society of America, A, 14*, 2152–2161.

Weldon, M. S. & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 670–686.

Weldon, M. S., Roediger, H. L. & Challis, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory & Cognition, 17*, 95–105.

Wenger, M. J. & Payne, D. G. (1997). Cue integration across study tasks and direct and indirect retrieval instructions: Implications for the study of retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 102–122.

Wenger, M. J. & Townsend, J. T. (1999, in pressa). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology.*

Wenger, M. J. & Townsend, J. T. (1999, in pressb). Spatial frequencies in short-term memory for faces: a test of three frequency-dependent hypotheses. *Memory & Cognition.*

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Young, G. & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika, 3*, 19–22.

van Zandt, T. & Townsend, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual and memory search: An evaluative review. *Perception & Psychophysics, 53*, 563–580.

Notes and Acknowledgments

AO gratefully acknowledges support from NIMH grant 1R29MH5176501A4. JTT and MJW gratefully acknowledge the support of the Indiana University Training Program in the Modeling of Cognitive Processes and NIMH National Research Service Award 1F32MH11491-04, which provided support for their initial collaborations. Some readers will notice the similarity of the title of this chapter to Shepherd, Ellis and Davies (1981) classic book *Perceiving and Remembering Faces.* Their book represents the first attempt to pull together diverse research efforts together and has been, over the years, a very strong influence and invaluable resource for many of us. We have aspired in the present chapter, and indeed this book, to follow in the same vein for integrating and developing the computational and quantitative sides of the diverse efforts first laid out in that book.

Tanaka, J. W. & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A*, 225–245.

Tanaka, J. W. & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*, 583–592.

Tarr, M. J. & Bülthoff, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple views? *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1494–1505.

Thomas, R. D. (1996). Separability and independence of dimensions within the same-different judgment task. *Journal of Mathematical Psychology, 40*, 318–341.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology, 25*, 168–199.

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–168). Hillsdale, NJ: Erlbaum.

Townsend, J. T. (1990a). A potpourri of ingredients for horse (race) soup. Technical Report 32, Indiana University Cognitive Science Program.

Townsend, J. T. (1990b). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should be) distinguished. *Psychological Sciences, 1*, 46–54.

Townsend, J. T. & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory (vol. 3)* (pp. 200–239). Hillsdale, NJ: Erlbaum.

Townsend, J. T. & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge: Cambridge University press.

Townsend, J. T., Hu, G. G. & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics, 36*, 35–49.

Townsend, J. T. & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *Mathematical Social Sciences, 4*, 25–71.

Townsend, J. T. & Nozawa, G. (1995). On the spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*, 321–359.

Townsend, J. T. & Thomas, R. (1993). On the need for a general quantitative theory of pattern similarity. In S. C. Masin (Ed.), *Foundations of perceptual theory: Advances in psychology: Vol 99* (pp. 297–368). Amsterdam: North-Holland/Elsevier Science Publishers.

Townsend, J. T. & Wenger, M. J. (1998). Correlated dynamic linear accumulators with gaussian noise: Initial predictions for target redundancy. In *Proceedings of the International Psychophysical Society* (pp. xxx–xxx). International Psychophysical Society.

Turk, M. & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3*, 71–86.

Ulrich, R. & Miller, J. (1997). Tests of race models for reaction time in experiments with asynchronous redundant signals. *Journal of Mathematical Psychology, 41*, 367–381.

Uttal, W. R., Baruch, T. & Allen, L. (1995a). Combining image degradations in a recognition task. *Perception & Psychophysics, 57*, 682–691.

Uttal, W. R., Baruch, T. & Allen, L. (1995b). The effect of combinations of image degradations in a discrimination task. *Perception & Psychophysics, 57*, 668–681.

Uttal, W. R., Baruch, T. & Allen, L. (1997). A parametric study of face recognition when image degradations are combined. *Spatial Vision, 11*, 179–204.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology, 43A*, 161–204.

Valentine, T. & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception, 15*, 525–536.

Vetter, T. & Poggio, T. (1996). Image synthesis from a single example image. In B. Buxton & R. Cippola (Eds.), *Computer Vision ECCV96* (pp. 652–659). Cambridge, UK: Springer Verlag.

typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 700–708.

Nosofsky, R. M. (1988b). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 54–65.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 3–27.

Nosofsky, R. M. & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Nozawa, G., Hughes, H. C. & Townsend, J. T. (1997). Complementary relationships between response times, response accuracy, and task requirements in a parallel processing system. Technical Report 204, Indiana University Cognitive Science Program.

Nozawa, G., Reuter-Lorenz, P. A. & Hughes, H. C. (1995). Serial and parallel processes in the saccadic control system. *Biological Cybernetics*, *72*, 19–34.

O'Mara, D. (1997). Automatic location of profile fiducials in 3d images of the human face. Eigth Annual Australasian Conference on Mathematical Psychology.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D. & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory & Cognition*, *22*, 208–224.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D. & Abdi, H. (1998). The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition*, *26*, 146–160.

O'Toole, A. J., Millward, R. B. & Anderson, J. A. (1988). A physical system approach to regognition memory for spatially transformed faces. *Neural Networks*, *1*, 179–199.

O'Toole, A. J., Vetter, T., Troje, N. F. & Bülthoff, H. H. (1997a). Sex classification is better with three-dimensional head structure than with image intensity information. *Perception*, *26*, 75–84.

O'Toole, A. J., Vetter, T., Volz, H. & Salter, E. M. (1997b). Three-dimensional caricatures of human heads : Distinctiveness and the perception of facial age. *Perception*, *26*, 719–732.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 324–354.

Rhodes, G., Brennan, S. E. & Carey, S. (1987). Identification and ratings of caricatures : Implications for mental representations of faces. *Cognitive Psychology*, *19*, 473–497.

Schooler, J. W. & Engstler-Schooler, T. Y. (1990). Verbal over-shadowing of visual memory. *Cognitive Psychology*, *22*, 36–71.

Schooler, L. J. & Anderson, J. R. (1997). The role of process in the rational analysis of memory. *Cognitive Psychology*, *32*, 219–250.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, *21*, 168–175.

Sergent, J. (1984). An investigation of component and configural processes underlying face recognition. *British Journal of Psychology*, *75*, 221–242.

Sergent, J. (1989). Structural processing of faces. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 57–91). Amsterdam: Elsevier.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54–87.

Shiffrin, R. M. & Steyvers, M. (1997a). A model for recognition memory: Rem–retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.

Shiffrin, R. M. & Steyvers, M. (1997b). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.

Sirovich, L. & Kirby, M. (1987). Low dimensional procedure for characterization of human faces. *Journal of the Optical Society of America, A*, *4*, 518–519.

Stevenage, S. V. (1995). Can caricatures really produce distinctiveness effects? *British Journal of Psychology*, *86*, 127–146.

Kadlec, H. & Townsend, J. T. (1992). Signal detection analysis of dimensional interactions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 181–228). Hillsdale, NJ: Erlbaum.

Kersten, D. J., Troje, N. F. & Bülthoff, H. H. (1996). Phenomenal competition for poses of the human head. *Perception*, *25*, 367.

Kohonen, T. (1977). *Associative memory.* Springer-Verlag.

Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling.* Sage Publications.

Lando, M. & Edelman, S. (1995). Receptive field spaces and class-based generalization from a single view of in face recognition. *Network*, *6*, 551–576.

Laughery, K., Rhodes, B. & Batten, G. (1981). Computer-guided recognition and retrieval of facial images. In G. Davies, H. Ellis & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 251–271). London, UK: Academic Press.

Light, L., Kayra-Stuart, F. & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 212–228.

Lockhead, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, *72*, 94–104.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Logan, G. D. (1992). Shapes of reaction time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 883–914.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147–180). Hillsdale, NJ: Erlbaum.

Maddox, W. T. & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*, 49–70.

Marr, D. (1982). *Vision.* New York: Freeman.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a general principle.* Cambridge, MA: Bradford.

Massaro, D. W., Weldon, M. S. & Kitzis, S. M. (1991). Integration of orthographic and semantic information in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 277–287.

Mauro, R. & Kubovy, M. (1992). Caricature and face recognition. *Memory & Cognition*, *Vol 20*, 433–440.

Miller, J. (1986). Timecourse and coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*, 331–343.

Miller, J. (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 160–169.

Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.

Mordkoff, J. T. & Egeth, H. E. (1993). Response time and accuracy revisited: Converging support for the interactive race model. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 981–991.

Mordkoff, J. T. & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 520–538.

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, order, and associative information. *Psychological Review*, *100*, 183–203.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251–269.

Neath, I. (1998). *Human memory: An introduction to research, data, and theory.* Pacific Grove, CA: Brooks/Cole.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1988a). Exemplar-based accounts of relations between classification, recognition, and

Cohen, M. M. & Massaro, D. W. (1992). On the similarity of categorization models. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 395–448). Hillsdale, NJ: Erlbaum.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology, 34*, 253–275.

Colonius, H. (1995). The instance theory of automaticity: Why the Weibull? *Psychological Review, 102*, 744–750.

Colonius, H. & Townsend, J. T. (1997). Activation-state representation of models for the redundant-signals-effect. In A. A. J. Marley (Ed.), *Choice, decision, and measurement: Essays in honor of R. Duncan Luce* (pp. 245–254). Hillsdale, NJ: Erlbaum.

Cottrell, G. W. & Metcalfe, J. (1991). Empath: Face, gender and emotion recognition using holons. *Advances in neural information processing systems, 3*, 564–571.

Cottrell, G. W., Munro, P. & Zipser, D. (1987). Learning internal representations of gray scale images: An example of extensional programming. In *Proceedings of the 9th Annual Cognitive Science Society Conference* (pp. 462–473). Erlbaum, Hillsdale.

Craw, I. & Cameron, P. (1991). Parameterising images for recognition and reconstruction. In P. Mowforth (Ed.), *Proceedings of the British Machine Vision Conference* (pp. 367–370). London: Springer-Verlag.

Deffenbacher, K. A. & Horney, J. (1981). Psycholegal aspects of face identification. In G. Davies, H. Ellis & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 201–227). London, UK: Academic Press.

Deffenbacher, K. A., Vetter, T., Johanson, J. & O'Toole, A. J. (1998). Facial aging, attractiveness, and distinctiveness. *Perception, 27*, 1233–1244.

Diederich, A. (1991). *Intersensory facilitation: Race, superposition, and diffusion models for reaction time to multiple stimuli.* Frankfurt: Peter Lang.

Ellison, J. W. & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception and Performance, 23*, in press.

Farah, M. J., Wilson, K. D., Drain, M. & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.

Fleming, M. & Cottrell, G. W. (1990). Categorization of faces using unsupervised feature extraction. *Proceedings of International Joint Conference on Neural Networks-90, Vol 2*, 65–70.

Galper, R. E. & Hochberg, J. (1971). Recognition memory for photographs of faces. *American Journal of Psychology, 84*, 351–354.

Garner, W. R. (1974). *The processing of information and structure.* New York: Wiley.

Golden, R. M. (1996). *Mathematical methods for neural network analysis and design.* Cambridge MA: Bradford.

Golomb, B. A., Lawrence, D. T. & Sejnowski, T. J. (1991). Sexnet: A neural network identifies sex from human faces. In J. M. R. P. Lippman & D. S. Touretsky (Eds.), *Advances in Neural Information Processing Systems Vol 3* (pp. 572–577). San Mateo, CA: Morgan Kaufmann.

Gray, M., Lawrence, D. T., Golomb, B. A. & Sejnowski, T. J. (1995). A perception reveals the face of sex. *Neural Computation, Vol 7*, 1160–1164.

Hancock, P. J. B., Burton, A. M. & Bruce, V. (1996). Face proccessing:human perception and principal components analysis. *Memory & Cognition, Vol 24*, 26–40.

Hill, H. & Bruce, V. (1993). Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion. *Perception, 22*, 887–897.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528–551.

Johnston, R. A., Milne, A. B., Williams, C. & Hosie, J. (1997). Do distinctive faces come from outer space? an investigation of the status of a multidimensional face-space. *Visual Cognition, 4*, 59–67.

Kadlec, H. & Hicks, C. L. (1998). Invariance of perceptual spaces and perceptual separability of stimulus dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 80–104.

References

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396–408.

Ashby, F. G. (1992). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1–34). Hillsdale, NJ: Erlbaum.

Ashby, F. G. & Alfonso-Reese, L. A. (1995). Categorization as probability density estimation. *Journal of Mathematical Psychology*, *39*, 216–233.

Ashby, F. G., Boynton, G. & Lee, W. W. (1994). Categorization response time with multidimensional stimuli. *Perception & Psychophysics*, *55*, 11–27.

Ashby, F. G. & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372–400.

Ashby, F. G. & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, *95*, 124–150.

Ashby, F. G. & Schwartz, W. (1996). A stochastic version of general recognition theory. Paper presented at the 1996 Meeting of the Society for Mathematical Psychology, Chapel Hill, NC.

Ashby, F. G. & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.

Atick, J. J., Griffin, P. A. & Redlich, A. N. (1996). Statistical approach to shape from shading: Reconstruction of 3d face surfaces from single 2d images. *Neural Computation*, *8*, 1321–1340.

Atkinson, R. C., Holmgren, J. R. & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, *6*, 321–326.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation, vol. 2* (pp. 89–195). New York: Academic.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Clarendon.

Beale, J. & Keil, F. (1995). Categorical effects in the perception of faces. *Cognition*, *57*, 217–239.

Benson, P. J. & Perrett, D. I. (1991). Perception and recognition of photographic quality caricatures : implications for the recognition of natural images. *European Journal of Cognitive Psychology* (pp. 105–135).

Benson, P. J. & Perrett, D. I. (1994). Visual processing of facial distinctiveness. *Perception* (pp. 75–93).

Bergen, J. R. & Hingorani, R. (1990). Hierarchical motion-based frame rate conversion. Technical report, David Sarnoff Research Center, Princeton NJ.

Beymer, D. & Poggio, T. (1996). Image representations for visual learning. *Science*, *272*, 1905–1909.

Bishop, C. M. (1995). *Neural networks for pattern recognition.* Oxford UK: Clarendon.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 657–668.

Bower, G. H. & Karlin, M. B. (1974). Depth of processing pictures of faces amd recognition memory. *Journal of Experimental Psychology*, *103*, 751–757.

Brennan, S. E. (1985). The caricature generator. *Leonardo*, *18*, 170–178.

Buhmann, J., Lange, J. & v.d. Malsburg, C. (1989). Distortion invariant object recognition by matching hierarchically labeled graphs. In *Proceedings of the International Conference on Neural Networks* (pp. 155–159). Washington.

Bülthoff, H. H. & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Science* (pp. 60–64).

Busey, T. A. (1998). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science*, *9*, 476–483.

Carello, C., Grosofsky, A., Shaw, R. E. & Pittenger, J. B. (1989). Attractiveness of facial profiles is a function of distance from archetype. *Ecological Psychology*, *1*, 227–251.

Christie, L. S. & Luce, R. D. (1956). Decision structure and time relations in simple choice behavior. *Bulletin of Mathematical Biophysics*, *18*, 89–111.

and work, and can be examined in both relative and absolute terms, at a variety of levels of analysis (e.g., the item or "feature" level, or the total task level; see Townsend & Ashby, 1983, Chapter 4). In relative terms, it is possible to consider a system's capacity for completing a fixed level of workload across a variety of stimulus manipulations; this is illustrated for a detection task across levels of facial "gestalt" in the chapter by Wenger and Townsend.

In absolute terms, three "flavors" of capacity have been suggested and can be roughly understood by considering what might happen as a result of increasing the processing workload. A system has *unlimited capacity* if performance (i.e., accuracy and latency) does not change as workload increases. A system has *limited capacity* if performance declines as workload increases.[12] Finally, if performance actually improves as a function of increasing workload then the system's performance is positively affected, and is said to have super-capacity. As before, if one were to allow the units of workload to be defined with respect to emergent properties of an input pattern space, as might result from the specification of an evidence space, then performance of the system in terms of the processing of this evidence space could be examined in terms of capacity.

In terms of faces being processed as configural objects, increasing the number of features to be processed in (for example) a parallel system with positively or negatively correlated channel rates could result in distinct effects in system capacity (e.g., Townsend & Wenger, 1998). For example, if positive correlations were used to represent the hypothesis of configural processing in a parallel self-terminating system, the presence of three rather than two features in the input might actually eventuate in distinctly higher, rather than lower, levels of processing capacity. In the particular case of parallel self-terminating processing, it is possible to assess capacity in an absolute sense in a particular experimental context using a derived measure known as the *capacity coefficient* (see Townsend & Nozawa, 1995). In other types of processing systems, it is possible to take advantage of another characterization of processing times (the integrated hazard function $H(t)$) to assess relative capacity across different types of stimuli (see Townsend & Ashby, 1978; Wenger & Townsend, 1999, in pressa, and the chapter by Wenger & Townsend, this volume). The ability to do this allows for tests of hypotheses such as the notion that the configural nature of faces (vs. other visual objects) may allow for increases in process capacity. As we hope is apparent in these brief descriptions, each of the informational and process characteristics has importance and relevance for questions in facial processing. In addition, there appear to be a number of ways in which these constructs can be used to represent important hypotheses about face processing.

Concluding Remarks

In summary, there seem to be rather natural connections between information processing and computational models—connections that admittedly remain largely to be developed—that could allow the theoretical distinctions and empirical tools of the information processing approaches to take advantage of the rich, formal descriptions of the pattern and evidence spaces provided by the computational models.

The application of the traditional information processing constructs to perceptually rich representations of the information in faces poses some new and daunting challenges to both approaches. Most importantly, and perhaps most pragmatically, there needs to be some set of formalisms for allowing the computationally derived models of the pattern and evidence spaces to serve as "front ends" to the information processing models. Although we believe that interfacing these two approaches may be difficult, and perhaps even impossible in totality, we believe that much will be learned in the attempt.

We set out to compose a chapter that would put some type of general analysis on a rather complex set of problems, and in doing so have proposed some rather general terminology and distinctions. From our perspective, the exercise has revealed a set of challenges and possibilities, particularly with respect to the nascent connections between computational and information processing approaches, connections that we feel may inject some vitality into both domains.

[12] A special type of capacity limitation is one where the available processing capacity must be allocated across all the elements or features to be processed, and then remains the same without reallocation. This is a situation referred to as fixed capacity processing (Townsend & Ashby, 1983, pp. 85 ff.).

features are pooled in a single output channel. This architectural possibility has been labeled *coactivation* and represents an interactive variant of parallel processing (c.f. Mordkoff & Yantis, 1991; Townsend & Nozawa, 1995).

In terms of the processing of perceptual pattern spaces such as those provided by the computational models described in the first part of the chapter, one could envisage a *parallel system* in which the feature dimensions are processed simultaneously, a *serial system* in which feature dimensions are processed one at a time, or a *coactivation* system in which the feature dimensions are combined to produce "derived features" by pooling the activations of a set of feature dimensions initially processed in parallel (Colonius & Townsend, 1997; Miller, 1982). Returning to the question of face configurations, the latter of these three possibilities is probably closest to the intuitive notion of a face being processed as a "configuration." That is, the information from the different features is pooled in the potentially unitary percept of the face. Thus, if evidence for serial processing (for example) were to be obtained, one might question the degree to which the stimulus is being processed as a configural whole. Strong tests for hypotheses regarding architecture exist in the work of Townsend and Nozawa (1995), and the chapter by Wenger and Townsend (this volume) illustrates the application of these tests to a facial processing (feature detection) task.

Stopping Rule

The stopping rule for a system refers to the criterion for deciding when to cease processing and emit a response. Historically, two stopping rules have been considered: *self-terminating* and *exhaustive processing* (e.g., Townsend & Ashby, 1983; van Zandt & Townsend, 1993). In a self-terminating process, a minimum of the evidence is evaluated (e.g., at least one feature or dimension meets or surpasses a criterion for the target response). In an exhaustive process, all of the evidence is evaluated (e.g., all of the features or dimensions must meet or surpass criterion). Consideration of a system's stopping rule can be done independent of consideration of processing architecture; but, as the coactive architecture posits only one output channel, the stopping rule needs only be considered relative to serial/parallel architecture.

One of the most frequently invoked combinations of architecture and stopping rule is the combination of parallel processing with self-termination. In these types of models, there is some minimum amount of information (specified, for example, in terms of basic elements of a stimulus) that is both necessary and sufficient for the cessation of processing. When completion of any of the presented items permits a correct response, self-terminating models are referred to as horse race or minimum time, or first completion time models (Colonius, 1995; Diederich, 1991; Logan, 1988, 1992; Mordkoff & Yantis, 1991; Townsend, 1990b; Townsend & Nozawa, 1995; Ulrich & Miller, 1997), since the task the race) is over as soon as the first element (dimension, feature, element, etc.) has been completed (i.e., as soon as the first horse crosses the finish line). If one were to use a computational model to derive a set of "features" from a pattern space, one could generate competing hypotheses for the processing of these features in terms of both architecture and stopping rule. In terms of faces being processed as configural wholes, one might intuitively posit parallel exhaustive processing (e.g., the need to consider both face shape and nose size). Or one might imagine that a task like gender classification could be completed on the basis of the first feature dimension (e.g., facial hair) to reach criterion (parallel and self-terminating).

Independence

As a processing characteristic, independence should be kept logically distinct from the types of informational relations discussed earlier. Independence in processing refers to the rate at which any one aspect of the stimulus affects the rate at which any other element is processed (c.f. Colonius, 1990). This is relevant for all three types of system architectures and for both types of stopping rules. Again, as noted in the preceding paragraphs, if a pattern space were to be established by means of a computational model, the processing of the resulting evidence space could be examined rather naturally with respect to the preservation or violation of processing independence of the emergent "features." In terms of faces being processed as configural wholes, one might hypothesize parallel processing of the features in which the rates of processing of those features do exhibit some temporal correlations. For example, such a system could show an increase in the rate of processing one feature (e.g., the eyes) in the presence of another feature (e.g., the mouth) in a task such as expression evaluation. Positive correlations such as these might be the most intuitive but not necessarily the only realistic possibility (see Townsend & Wenger, 1998).

Capacity

Processing capacity refers to the way in which a system responds to variations in workload. Capacity is thus linked to notions of physical energy, power, non-integral (e.g. Lockhead, 1966), and integral vs. separable (e.g. Garner, 1974). The associated empirical methods of determining whether particular psychological dimensions are or are not independent have instigated a tremendous body of basic and applied research.These methodologies, based primarily on operation definitions (that is, defining the concept through the outcomes of the experimental manipulations meant to test or reveal them), emanating from different laboratories, lacked a common theoretical underpinning which made merging findings from the various methods problematic. The absence of a rigorous framework led to some definitions suggesting results that contradicted concepts or findings associated with a different method, supposedly measuring the same thing. The absence of a common underpinning made it difficult to know exactly what to make of empirical results from the different methods, even when they seemed to agree. A universal theoretical foundation is not yet with us, but progress has been made in providing a rigorous meta-model and common set of concepts and language that underlie a number of earlier and new approaches.

One missing component from a number of approaches has been the failure to preserve the distinction (emphasized in signal detection theory) between the accumulation or acquisition of the psychological evidence and the decisions that operate on that evidence (see also Maddox, 1992). Preserving that distinction in a multi-dimensional context requires that separability, independence, and other related notions be specified with respect to their preservation or violation at either the evidentiary or decisional levels. A general set of definitions for these types of informational characteristics was set out by Ashby and Townsend (1986), and empirical tests and applications across a range of tasks have been documented in the body of work contributed by Ashby, Townsend, and colleagues (e.g., Ashby & Townsend, 1986; Ashby & Alfonso-Reese, 1995; Kadlec & Townsend, 1992; Thomas, 1996).

In the simplest terms, mapping these distinctions onto a computationally derived pattern space for faces translates into a question about the "localizability" the information necessary to carry out a task. For example, perceptually separable (and possibly independent) stimulus information would be "located" on separate (possibly orthogonal) axes/dimensions in the pattern or evidence spaces. The notion of the separability of stimulus dimensions, in combination with some processing constructs, may be quite relevant for understanding the long-debated arguments about the configural nature of a face, particularly as it relates to the degree to which different stimulus features may or may not remain distinct in the encoded representation (e.g., Farah, Wilson, Drain & Tanaka, 1998; Tanaka & Farah, 1993; Tanaka & Sengco, 1997).

The data for testing hypotheses regarding separability and independence at perceptual and decisional levels often takes the form of identification/confusion matrices. And the methods for testing for the maintenance or violation of separability and independence at these two levels involves application of multi-dimensional signal detection measures (e.g., Kadlec & Townsend, 1992; Kadlec & Hicks, 1998) and fitting of sets of multidimensional Gaussian models (e.g., Ashby, 1992). A variety of tasks have been used in the development of this approach (see Thomas, 1996, for a review), each of which has the potential for use in pursuing questions in facial cognition.

Processing Characteristics

There are four general characteristics of any information processing system (Townsend & Ashby, 1983) that need to be considered relative to operations on the pattern and evidence spaces. These four characteristics are the system's (a) architecture, (b) stopping rule, (c) independence in time and space, and (d) capacity. Although each of these dimensions has been the focus of decades of research effort, only recently has it been possible to empirically assess performance in a way that supports the simultaneous characterization of processing on several of these dimensions simultaneously (see Townsend & Nozawa, 1995). The chapter by Wenger and Townsend (this volume) represents one of the first applications of that theoretical and empirical technology outside its original domain (see also Nozawa, Reuter-Lorenz & Hughes, 1995; Nozawa, Hughes & Townsend, 1997) and illustrates how the theoretical and empirical technology can be applied to a facial processing task (feature detection in this case).

Architecture

The architecture of a system is the manner in which that system accomplishes its processing goals in space and time. Historically, questions of architecture have focused on the distinction between parallel and serial processing (e.g., Atkinson, Holmgren & Juola, 1969; Christie & Luce, 1956; Townsend, 1972, 1974, 1990a; Townsend & Ashby, 1983). More recently, a third architecture has been proposed based on research investigating redundancy gain (e.g., Miller, 1982, 1986, 1991; Mordkoff & Yantis, 1991; Mordkoff & Egeth, 1993). This is a form of parallel processing in which the outputs of individual ceptual tasks is presented in Table 1. In this table, we lay out what might be the minimal characteristics of the psychological evidence spaces required to support performance in each of the tasks. We also point to a small (and admittedly incomplete) set of citations illustrating information processing approaches that could be coupled to the models of the pattern space discussed in the early part of this chapter. Our intent is not to highlight models that necessarily have been applied to facial processing tasks. Rather, we think the models represented by these citations are excellent candidates for providing connections between the computational and information processing models, in the context of questions regarding facial cognition.

In addressing the link between the pattern and evidence spaces we must characterize the relationships both within and between these spaces. These relationships have been considered separately up to this point in the chapter. We believe that one of the possible ways to link the pattern and evidence spaces involves the inherent capacity to address both informational and process characteristics simultaneously (see the chapters by Campbell et al., and Wenger & Townsend for related discussions). We begin with the assumption that the pattern space contains all of the information that we are capable of extracting from faces. Psychological evidence spaces are more difficult to define, but entail, generally speaking, all of the information required for completing a particular task with faces. Construction of the space generally involves some transformation, computation, or analysis on the pattern space. What they will not usually explicitly provide is a specification of the processes (or the characteristics of the processes) that operate on those spaces.[11]

Thus, a complete account of the performance of any task involves a definition of the germane perceptual pattern space, the psychological evidence space(s), and the task logic and processes that operate on those spaces. For simplicity and clarity, we assume that each task must make contact with some subset of the information available in the pattern space. Also, although as pointed out earlier, certain milieus might require more than one pattern or evidence space, we will speak as if only one of each were required. The psychological space combines the relevant subspaces, and is constrained both by the nature of the information contained in the subspace and by the manner in which a system can operate on this subspace to achieve the task goal.

Data consistent with this general idea have come from a variety of sources, including applications of spatial frequencies analyses to object and face processing. In particular, a number of studies have documented task dependencies in the use of spatial frequencies (e.g. Sergent, 1984, 1989; Uttal, Baruch & Allen, 1995a,b, 1997), illustrating how the same pattern space (i.e., the set of input spatial frequencies) may make contact with different aspects of stored information, resulting in different psychological evidence spaces specific to different tasks (e.g., discrimination vs. recognition). For example, Uttal and colleagues (e.g. Uttal, Baruch & Allen, 1995a,b, 1997, also, this volume) have documented the differential ability of high spatial frequency components of the input pattern to support performance in discrimination tasks and low spatial frequency components to support performance in recognition tasks. Wenger and Townsend (1999, in pressb) have further shown how such task dependencies, relative to specific components of the input pattern space, can be revealed as a function of manipulations of both task instruction (e.g., discrimination vs. recognition) and retention interval. In related work, O'Toole and colleagues (O'Toole, Deffenbacher, Valentin, McKee, Huff & Abdi, 1998) have demonstrated the contributions of different aspects of a single underlying evidence space to performance on gender classification, recognition, and attractiveness ratings.

The information processing literature contains a variety of possibilities for predicting performance on the basis of psychological evidence spaces and process characteristics, but these structures have yet to be connected in any concrete way to the pattern space. And since only a limited amount of theoretical and empirical work has been done on the psychological evidence spaces and processes specific to facial cognition, the paragraphs that follow emphasize the *potentials* for applying, extending, and connecting existing work in the information processing literature to questions in the facial processing literature and models of the facial pattern space.

Characteristics of the Evidence Spaces

In terms of informational characteristics, a longstanding and important issue has been the degree to which the various dimensions of a stimulus may be *separable* or *independent*. Traditional distinctions pertinent to this issue have included unitary vs. analyzable (e.g. Shepard, 1964), integral vs.

[11] A simple example, to preview the paragraphs to follow, might be found in a task requiring a comparison between two faces that have been coded in terms of vectors of pixel values. The comparison of these codes could either be done sequentially across all of the pixels (i.e., a serial process) or done on all of the pixels concurrently (i.e., a parallel process). The point to be made is that the coding in this case does not constrain the characteristics of the processes that might operate on the codes.

the category of faces with all of its composite subcategories, how do we access the information we need from this face space? The information processing literature has considered this question in a general sense for many years and has much to offer here, in the context of the more perceptually based computation models we have discussed.

As one aim of this chapter is to highlight potential connections among the structural, temporal and logical/computational pieces of the larger problem of facial cognition, we note the first of these potential bridges here. The codes described to this point demonstrate the rich sources of information potentially available to perceptual and cognitive processes. These codes can conceivably be used to provide a precise set of inputs to traditional information processing models. Additionally, as noted several contributors to this volume, (see the chapters by Uttal, Campbell et al., Wenger & Townsend, Busey, Steyvers & Busey, and Valentine), there are numerous perceptual and cognitive models of information processing that could be applied to the more precise analysis of face spaces constructed by using the codings described in the preceding sections of this chapter. These information processing models have as a particular strength a focus on decisional goals and logic and rules for operating on informational sources. What they currently lack, however, are mechanisms for connecting with any type of precise model of the pattern space.

Specification of the mechanisms for connecting precise models of the input patterns with information processing models may be one of the most promising avenues for developing connections among the computational and information processing approaches. For example, approaches such as the fuzzy logical model of perception (described by Campbell et al., this volume), cognitive stochastic process theory and systems factorial technology (developed by Townsend, Schweickert, and numerous others, and briefly summarized by Wenger & Townsend, this volume), and the dynamic extension to time stochastic process theory (described by Wenger & Townsend, this volume) all begin with intentionally general specification of input signals, sources of evidence, and features. It is possible that the codes described so far, either as a body of input data, or summarized or aggregated in some form (as described later in this chapter) could provide explicit form for these intentionally general constructs.

It is worth noting also that the generality of the information processing models was intended originally (at least in a loose sense) to allow for the application of the various approaches across many psychological milieus. Providing a computational specification for the input constructs allows for direct application to the specific issues of facial cognition and does not compromise the generality of the approaches. This claim holds in spite of the fact that such a step inevitably involves computational instantiations of analytic models. Thus a path for developing these possibilities might include the invention of general, analytic, and conceptual frameworks designed to hold for any implementation (e.g., Townsend & Nozawa, 1995). These frameworks could then be implemented computationally using a model of the input information space. The result should be a coupling of the precise specification of the information available in the stimulus with models that explicitly require specification of alternative information processing architectures and rules for operating on that information. These latter types of models have a demonstrated track record of lending themselves to strong experimental tests.

A similar case can be made for the ability of the codes described so far to act as the foundation for specifying the general stimulus dimensions required by psychological models that use "features" as the dimensions or axes of representational spaces. We note a few of these here but will take up this discussion more thoroughly after we return to the concept of a face space. Approaches such as multi-dimensional scaling (e.g., Nosofsky, 1986, see also chapters by Busey, Steyvers & Busey, and Valentine, this volume) or multi-dimensional extensions to signal detection theory in either its static (e.g., Ashby & Townsend, 1986; Ashby & Alfonso-Reese, 1995; Thomas, 1996) or dynamic (Ashby & Schwartz, 1996) forms rely implicitly on feature axes that are intentionally specified at a very general level. As with the process models just described, the coupling of these models of the psychological evidence spaces with the computational models of the pattern space could lead to explicit links between information available in the stimulus and the logical relations among the psychological dimensions that derive from that information. By logical relations we mean factors central to the accessibility of the stimuli or stimulus components at perceptual and decisional levels, including the classic psychological constructs of integrality/separability and the independence of dimensions. Such coupling could again allow for the use of the strong experimental tests that have been the hallmark of the information processing approaches.

Information and Information-Processing Requirements

A generic analysis of the informational requirements in a representative sample of memory and perrelative proportion of the source individuals in the resulting morphed face. Beale and Keil (1995) found that observers perceived the identity of the morphed faces "categorically," in the sense that there was a very abrupt discontinuity in the proportion of identifications of the individual across the morph continuum.

For present purposes, morphing technology may allow another point of contact between computational models of the pattern and psychological evidence spaces and models of the psychological operations that are applied to those spaces. For example, one might imagine a scenario in which the paths described by the image morph could be systematically varied with respect to a set of feature dimensions (determined, potentially, by a model of the individual face space), resulting in tests of competing psychological models for the use of this feature information.

The use of morphing to create stimuli is unquestionably an exciting and potentially highly informative way of exploring a face space. Nevertheless, the field would likely profit from careful thought about the nature of the information underlying the face space, the "meaning" of the direction and shape of trajectories within the space, and the need to exert control over those trajectories in the service of hypothesis testing. A quite captivating look at the perceptual-computational oddities of treks through face space using morphing as a navigation tool can be seen in Busey (this volume, also Busey, 1998). We will not detail that work here but recommend it highly to readers interested in using morphed stimuli for psychological experiments. We will however, make a few general points. First, at the risk of being repetitive, the kind of face encoding that underlies most morph procedures is but one kind of face encoding. Due to the fact that most psychological researchers make use of commercial software, the precise nature of the codes corresponding to the morph trajectories is often not actually known,[9] nor is it subject to experimenter control.

Further, to our knowledge there is presently no compelling evidence either for or against the psychological relevance of these codes. For example, Hancock et al. (1996) found morph codes only very slightly better than a raw, image-based code for fitting data on recognition and typicality.[10] Second, where careful examinations of the similarity structure among morphed faces have been made, some surprising deviations from the anticipated perceptual geography of a morphed-based face space have been found (Busey, this volume). Third, where qualitatively different kinds of face encodings have been used (e.g., full corresponded three-dimensional codes), the perceptual geography of the space has turned up even more surprises, including facial age variations as radiant trajectories in the face space (see also Carello, Grosofsky, Shaw & Pittenger, 1989; O'Toole, Vetter, Troje & Bülthoff, 1997a).

Should these obstacles put a damper on the hope that there might be a tight link between the psychological and computationally-derived notions of a face space? As our discussions to this point have suggested, we think the answer is "no." The constraints that result from the fact that quantitative models must commit to a particular operational definition of the input representation of a face suggest distinct opportunities for strong inference and falsifiability at the level of the pattern space, with additional opportunities to connect to models of the psychological decisions and processes. Certainly, and as one might expect, there is good evidence that the perceptual and cognitive implications of the choice of input representations cannot be ignored. Clearly, a "straight line" in a face space resulting from one encoding may be a curve in a space based on another encoding (see Townsend et al., this volume for more on this issue). Still, it is equally clear that computational models can be used as very powerful tools for exploring the nature of human representations of faces. This can be done by varying the nature of the input coding and comparing the human and model performance both qualitatively and at the level of individual faces. In addition, as we have taken pains to emphasize, the harnessing of computational models of the pattern and evidence spaces to formal descriptions of the rules and processes that operate on these spaces suggests a wide set of possibilities for "grounding" psychological inquiry.

Perceptual Codings and Information Processing Approaches

In the final part of this chapter, we address issues concerned with accessing and operating on the information that is encoded in a face space. In other words, once we have quantified the information in faces, set up a face space framework for representing

[9] In fact it is proprietary. One of us has actually telephoned Gryphon software, the commercial software company that markets "Morph," the program used by most psychologists to make face morphs. After speaking to several engineers at the company to request information about the nature of the representation and algorithm used in the program, it became clear that many relevant details were not, for good commercial reason, available to the public.

[10] Additionally, their implementation is but one of many possible ways of defining such a code.

logical model of perception, models based on multidimensional scaling, or multidimensional generalizations of signal detection theory) or abstract (as in the Valentine, 1991, theory), are necessarily concrete in the computational implementations of the physical face space. The necessity of specifying the source of the feature axes on which the computational models operate suggests that the inputs to the traditional information processing models need not remain arbitrary or abstract. For example, it is easy to imagine a situation in which the pattern space generated by a computational model would serve as a constrained set of inputs to particular models such as the FLMP (e.g., Massaro, 1998, see also Campbell et al., this volume), or generalized psychological evidence spaces such as those used as the basis for model testing within general recognition theory (e.g., Ashby & Townsend, 1986). It is also easy to envisage process models such as those described by Wenger and Townsend (this volume, see also Townsend & Nozawa, 1995) that use the computational principles described in this and previous sections to determine the form of the input signal. Each of these possibilities suggests that the formal descriptions of the stimulus pattern spaces or more general evidence spaces (in the case of a particular task application) can be connected to models that allow for strong inferences regarding the psychological relations among and utilities of the input feature dimensions and the characteristics of the processes that operate on those dimensions.

Navigating Through Face Spaces

Once we have defined a physical face space using a particular input or encoding, we can then "navigate" through the space in any direction we like, producing a continuously varying facial stimulus as we go. There is, indeed, a very long tradition of navigating through face spaces for fun and profit, particularly with the goal of face synthesis. Perhaps the earliest use of a computationally defined face space was made by Brennan (1985) in her "automated caricature generator." Her algorithm operated on line drawings of faces using the locations of fiducial points as a face encoding. Recall that this is a two-dimensional configural coding. The x, y locations of these points were recorded for a large number of faces and an average of the fiducial points for these faces was computed. Next, to create a caricature of an individual face, a measure of the deviation of the face from the average two-dimensional configuration was computed. Finally, the line drawing of the face was re-sketched with the "distinctive" or unusual features of the face exaggerated to produce the caricature.

A number of psychologists have made good use of this kind of caricature generator as a method for directly manipulating the distinctiveness of faces using line drawings (Rhodes, Brennan & Carey, 1987; Benson & Perrett, 1994) and photographic quality images (Benson & Perrett, 1991). Both the line drawing and photographic representations yield compelling caricatures, and there is now evidence indicating that these caricatures can be recognized more quickly and more accurately than the veridical images (e.g., Benson & Perrett, 1994; Mauro & Kubovy, 1992; Stevenage, 1995). Further, these caricatures are rated as better likenesses of individuals than the veridical images (Benson & Perrett, 1994). Such a result has both an intuitive appeal (as even a casual consumer of political cartoons might realize) and a counter-intuitive quality, since it is not altogether clear how most psychological models of either individual or aggregate face spaces could account for the effect.

Recent work with this caricature algorithm on fully-corresponded three-dimensional surface codes has yielded another surprising and theoretically challenging result: Three-dimensional caricatures of faces appear *older* than the veridical faces (O'Toole, Vetter, Volz & Salter, 1997b). This is reasonable in that creases and facial wrinkling are three-dimensional features of faces, and making them more distinct is clearly likely to age the faces. This result poses at least one theoretical challenge to face space models not specific about the details of the perceptual representation. Specifically, distance from the average in a face space derived from the three-dimensional features of faces differs qualitatively from distance from the average in a face space based on the two-dimensional configural features of faces. Consequently, it is possible that the latter encoding would either not be able to make predictions regarding, or would mis-predict, the effects of three-dimensional caricature.

Another interesting technique for navigating through face space "between faces" is morphing, a popular technique for blending images of objects and faces in an apparently continuous fashion. The software needed to morph images is now widely available and runs on just about any type of computer. This accessibility has allowed researchers to begin making use of its potential as a tool for addressing questions about the psychology of the perceptual and evidentiary spaces.

To our knowledge, the first researchers to make use of morphing in the context of face perception were Beale and Keil (1995), who morphed together the faces of famous individuals to produce a smooth "linear" transition between source faces, varying the components, PCs, or eigenvectors). The PCs represent orthogonal *patterns* of the individual "feature" elements. The analysis itself is applied to a matrix of the cross-products of the face measures (e.g., pixels, etc.), and the resultant PCs can be ordered according to the eigenvalue associated with each. This eigenvalue is related to proportion of variance each PC explains in this covariation matrix.

At the level of sets of faces, the eigenvalue is a quite tangible measure of the importance of each PC for describing the *entire set* of faces. A psychological interpretation of this eigenvalue might be, simply, "How useful is its associated feature (PC) for describing faces, in general?" This property of the approach has further psychological appeal as a way of simulating some aspects of the sensitivity of human memory to the statistical structure of experience (e.g., Anderson & Schooler, 1991; Schooler & Anderson, 1997). For faces, one of us has argued that the classic "other-race effect," the recognition advantage for same- as opposed to other-race faces, may be an example of human sensitivity to the statistical structure of our own experience with faces (O'Toole, Deffenbacher, Valentin & Abdi, 1994).

At the level of individual faces, each face can be described in this space by its coordinates on the set of axes or PCs. These coordinates comprise an abbreviated code for the face that consists of a list of numbers (i.e., coordinates). Each coordinate measures the extent to which the PC is needed to construct the face. In other words, it represents the face's value with respect to each "feature" or axis. This gives a very tangible measure of the importance of each axis for describing the information in an *individual* face.

It is worth noting that when PCA is applied to a physical measure of faces, such as pixels, surface values, or pre-morph codes, the resultant PCs are of the same form. What this means is that they can be viewed (if they are image-based), constructed (if they are surface based), or synthesized (if they are derived from pre-morph codes). In fact, the feature value or coordinate of the face with respect to each axis is a direct measure of the similarity of the face vector to the eigenvector or PC. With the possibility of viewing PCs or eigenvectors, one can attempt to visually interpret the information they capture (e.g., Hancock, Burton & Bruce, 1996; Turk & Pentland, 1991; Vetter & Troje, 1997; O'Toole, Vetter, Troje & Bülthoff, 1997a; O'Toole, Vetter, Volz & Salter, 1997b).

The PCA approach is also related to neural network approaches to simulating face processing tasks. Though not always noted explicitly, these links become rather obvious when one views things in terms of physical or computationally-based face spaces. In fact, the basics of linear systems analysis were first applied to faces by Kohonen (1977) over 20 years ago. He used a linear auto-associative neural network as a content addressable memory for faces as an example, because it was easy to "see" the performance of the network on individual images; he also noted that the auto-associative network was implementing PCA. Although contemporary neural network and PCA approaches differ in terms of implementation (e.g., Bishop, 1995; Golden, 1996), they remain closely related. Even three-layer (and more) back-propagation networks, which have been used by neural network researchers for face classification (Fleming & Cottrell, 1990; Cottrell & Metcalfe, 1991; Golomb, Lawrence & Sejnowski, 1991), can be shown to implement a (rotated) PCA at the level of the hidden unit activations (cf., Cottrell, Munro & Zipser, 1987).

Addressing Psychological Questions with Physical Face Space Hypotheses

Connecting the computational conception of a physical face space to the psychological face space is relatively straightforward and goes a long way toward resolving what has been called "the paradox of most information process models: the form of the information, and what happens to it, are usually absent from the models" (Townsend & Thomas, 1993, p. 340). In both cases, the generic principles of an informational space apply; however, the computational implementations add two things. First, a computational model will create a face space contingent on the kind of input encoding used: The same set of faces can give rise to many different experience-specific face spaces. Thus, the face space that derives from two-dimensional information may well be different than that which derives from a three-dimensional coding, and each of these may vary as a function of the nature of the processing goals and operations at encoding. This will obviously lead to different predictions about which faces are typical versus distinct and will also lead to different predictions about which faces are similar. It also has the potential for making long-standing cognitive constructs, such as encoding specificity and transfer appropriate processing, much more concrete. Finally, it holds out the possibility of addressing questions about human representations of faces in that spaces based on different kinds of input encodings/representations can be compared vis-a-vis the quality of predictions they make about human performance on face processing tasks.

Second, the feature axes, which are either arbitrary (e.g., operationalized with respect to the experimental manipulations, as in work with the fuzzy (i.e., the points that correspond to the projection of a point x, y, z in the scene onto the left and right retinae). In the latter case, one needs to match the points in an image taken at time t to the points in an image taken at time $t + \Delta t$. For faces, objects that share many common features, the principle is the same: One tries to match the "corresponding" points in two faces. The techniques have been borrowed from the literature on structure-from-motion and rely on elaborated optic flow algorithms (e.g., Bergen & Hingorani, 1990).

This concept of completely corresponding points has its analogue in practical and abstract notions of producing faces as mappings or functions, for instance, from a rectangle on the plane or perhaps as a function of points on a cylinder or sphere (e.g., O'Toole, Vetter, Troje & Bülthoff, 1997a; O'Toole, Vetter, Volz & Salter, 1997b, also, Townsend et al., this volume). If one knows exactly which point, of say, a rectangle is matched to each point of Face A, it is in principle straightforward to find the corresponding point of Face B as an image of the same point of the original rectangle.

Although these codes are a very much improved way of representing the information in faces, there are some practical questions that limit their utility. Foremost, these algorithms have a great deal of difficulty dealing with naturalistic images of faces (e.g., faces with hair), and require *very* good pre-alignment of the faces before proceeding—certainly as good or better than the alignment required for making sensible use of raw image codes. In any case, when the correspondence can be achieved, the flexibility of completely corresponded codes for face synthesis is impressive.

To summarize, in choosing a method for initially quantifying the information in faces (i.e., for constructing a representation of the initial pattern space), a number of factors must be considered. First, the difficulty of discriminating a large number of individual faces has led most computational modelers to the use of codes that retain at least some of the basic perceptual information in faces. Second, in doing this, the question of retaining access to the location of the discrete facial features has been addressed also by supplementing the purely image-based codes with information about the two-dimensional configural structure of the face. Combined, these kinds of codes can be manipulated or weighted as necessary to provide a reasonably complete description of a face, at least from a single viewpoint.

Representing Groups of Faces in a Physical Face Space

Once an encoding system for the faces has been chosen, the next problem concerns representing groups of faces in a more concrete embodiment of a face space that can serve as a base for accomplishing the task at hand. At the level of a common denominator between psychological and physical face space models, the representation/quantification of faces employed by *any* computational model can be used to create a generic physical face space as follows. Because all computational models require an encoding or quantification of the information in faces in terms of a set of measures or "features," we can imagine that each face is encoded by its values on the set of features and thus can be expressed simply as a list or vector of values. To represent a set of similarly encoded faces in a face space, we might proceed by assuming that each of the input features items defines a "feature axis." Considering all of the features, a face then can be thought of as a point in high dimensional space. The dimensionality of the space is simply the number of feature axes—e.g., for a raw image based code, the maximum dimensionality of the space is determined by the number of pixels. Concomitantly, the average face can be computed in a straightforward manner, and the typicality of a face can be assumed to be the distance of the face from the average. To stress the strong potential for connecting computational and information processing models, we should note that the steps just outlined are steps that are involved in constructing general representational spaces using a variety of traditional information processing models.

Before proceeding, a number of links to the statistical underpinnings of a computationally-derived face space and to neural network approaches are in order.

Statistics of the Face Space and Neural Network Approaches

Over the past decade, it has been popular in the computational literature to create a physical face space in a more sophisticated fashion than in the generic way described above. Specifically, PCA or eigen decomposition has been applied to this problem by many researchers (e.g., Sirovich & Kirby, 1987).[8] PCA is a technique used to describe a set of correlated variables (i.e., what we are calling features, e.g., pixels) using a smaller number of uncorrelated or orthogonal variables (i.e., what we will call principal

[8] PCA is related to many tools commonly used in psychology. For example, it is a linear version of multidimensional scaling, has numerous connections to models of categorization and memory, and can be considered as a particular relative of multidimensional generalizations of signal detection theory (Ashby & Townsend, 1986, see also discussions in Steyvers & Busey, this volume).

Pre-morph Codes

A more strictly aligned coding than the purely image-based code described previously was proposed recently by Craw and Cameron (1991) and has been applied to both computational and psychophysical work by Hancock, Burton, and Bruce (1996). In this code, the locations of feature landmarks/fiducial points are combined with information about the face image. To encode a particular face in this way, one begins by defining and locating a set of fiducial points in the face. Hancock et al. refer to the set of these points as the "face shape." Using these points, the face is then "aligned" or warped to the average face shape. Hancock et al. refer to this aligned image as a "shape-free" code for the face. More simply, this combined code can be understood easily in terms of "morphing," a popular technique for blending images of objects and/or faces (see discussions in Busey, 1998). To morph images of two faces together one must first locate a set of corresponding points on the faces. These points include the fiducial points of the face but are often supplemented with additional points to obtain a high quality morph. For example, one might represent the bottom of the mouth with the endpoints of the lips and 12 equally spaced points in between that trace the bottom edge of the lower lip. Once the corresponding points are located, the image data can be warped or interpolated to make a smooth transition between the two faces, using the corresponded points as guides. Morphing operations are considered as a type of trajectory in a high dimensional face space in the chapter by Townsend et al. (this volume).

For brevity, and for lack of a better term, we will refer to the combined shape and shape-free components of a face as a "pre-morph" code. More concretely, the code consists of the x, y locations of some number of "landmark" and supplemental points in a face image, and a vector of pixel intensities taken from the shape-free face image. The nature of the information captured in a pre-morph code has not often been considered explicitly (although see Busey, 1998). However, in light of the popularity of using morphed faces in psychological experiments (e.g., Beale & Keil, 1995), this is well worth doing.

First, the shape code is simply a measure of the two-dimensional configural properties of a face, because it implicitly codes the relative locations of the facial features in the two-dimensional image. Second, the shape-free code contains a shape-normalized representation of the image data—i.e., the pigmentation minus the basic two-dimensional configural properties of the face. Finally, it is worth noting that these codes leave out one kind of information that is available in most of the commercial software for producing morphs—specifically, most programs allow for the placement of curved contours that fit around the outlines of the features. This enables a better quality synthesis of the faces that lie "between" the two source faces by defining a contour that can be used to guide pixels in the warping procedure. We will discuss this synthesis problem shortly.

Although pre-morph codes provide a fairly complete coding of faces as input patterns, one that combines the best properties of image-based (i.e., access to subtle, internal shape and texture information) and geometrically-based codes (i.e., maintaining registration of the discrete features), they have some practical shortcomings. The first is that in most applications, these points must be defined and hand-located by a human operator, a tedious and time-consuming operation, especially given that high quality facial morphs require as many as 200 or 300 points. Although some automatic algorithms are available for locating the major fiducial points on faces(cf., O'Mara, 1997), they are not generally adequate for locating a sufficient number of supplemental points to produce good quality morphs. A second practical problem for computational models concerns the relative importance of the shape- and shape-free codes. By sheer quantity, the number of image-based measures far outweighs the number of fiducial and supplemental points in the shape code. Though intuitively unsatisfying, it is not yet obvious how else one ought to combine the parts of these codes. The decision will nonetheless have a very potent effect on the performance of any model that uses these codes as the basis for processing.

Completely Corresponded Pre-Morph Codes

Recent work in computer vision has demonstrated that under some conditions (see below), it is actually possible to set two faces into a completely registered coordinate system such that all of the points in one face image can be made to correspond to all of the points in another face image (Beymer & Poggio, 1996; Vetter & Poggio, 1996). The generation of such completely corresponded morph codes makes it possible to completely automate the process of morphing. In fact, the problem of locating the fiducial and supplementary points on a face, as is required for morphing, can be seen as a special case of the well-known "correspondence problem" in computer vision that is commonly associated with the computation of structure-from-stereopsis and structure-from-motion. In the former case, one tries to locate "corresponding" points on the left and right retinae most common consists of "reasonably well-aligned" image data—often coded simply as of vectors of pixel intensities (e.g., Fleming & Cottrell, 1990; Gray, Lawrence, Golomb & Sejnowski, 1995; Kohonen, 1977; Lando & Edelman, 1995; O'Toole, Millward & Anderson, 1988; Sirovich & Kirby, 1987; Turk & Pentland, 1991). By reasonably well-aligned, we mean face images of about the same size, viewed from a frontal pose, and centered in an image so that the eyes (or some other reference feature) are located at a roughly equivalent level.[6] These can be considered "low-level" visual codes in the sense that they are arguably close to the kind of data with which a human observer begins the process of face perception and cognition, i.e., two-dimensional images on the retina. Von der Malsburg and colleagues (e.g., Buhmann, Lange & v.d. Malsburg, 1989) and Lando and Edelman (1995) have taken these codes one step further to incorporate some basic aspects of early visual processing in the cortex. Specifically, they have used image-based codings that mimic the operations of oriented line detectors across a range of resolutions.

With the recent availability of laser scan technology, similarly unprocessed three-dimensional data have also been used as input to computational models (Atick, Griffin & Redlich, 1996; O'Toole, Vetter, Troje & Bülthoff, 1997a). Laser scanners provide a "ground-truth measure" of three-dimensional shape of a face. In other words, they provide the solution to the elusive ill-posed, "inverse optics" problem of computer vision. The inverse optics problem refers to the problem of computing the three-dimensional structure of a scene that "caused" a particular two-dimensional image.[7] For faces, the goal of the inverse optics problem is to derive a representation of the three-dimensional structure of the face from a two-dimensional image. Although classic theories of vision (e.g., Marr, 1982) assume the primary task of the human visual system is to solve this problem, in recent years, this assumption has been questioned both in the context of face (Valentin et al., this volume) and object recognition (Bülthoff & Edelman, 1992; Tarr & Bülthoff, 1995). At present, the extent to which humans represent faces in terms of their two- rather than their three-dimensional structure is highly controversial (cf., Valentin et al., this volume, for a complete discussion of the issue).

[6] Turk and Pentland (1991) have proposed algorithms for achieving at a certain degree of size scaling, and face location in the image. As we will describe shortly, Von der Malsburg and colleagues (e.g., Buhmann, Lange & v.d. Malsburg, 1989) have added a degree of view invariance to their algorithm.

[7] This problem is known to be under-constrained and unsolvable without imposing additional constraints, which are not easy to define in a general fashion.

In the sense that laser scan data provide direct and accurate information about the three-dimensional structure of a face, one might argue that using these data as input to a model is "cheating." However, very useful comparisons can be made by pitting the predictions of two- versus three-dimensional face representations against each other in a computational model and comparing the model and human performance. This should give some insight into the extent to which human observers represent faces or objects in terms of their two- versus three-dimensional features. Note that this raises the questions of the necessity and sufficiency of both of these types of information, relative to the particular task. To answer such questions at a reasonably general level, it is necessary to have some definition of the decisional goals and rules of the candidate systems along with demonstrations of the relative abilities of the candidate computational models to solve the problem and account for the human data.

The three-dimensional head surface provided by a laser scanner is sometimes, though not always, accompanied by a "reflectance map" that captures the efficiency with which the sample points reflect light of various wavelengths, i.e., the colors. These sample points are usually coded in a standard RGB format. Although the reflectance map is in some ways comparable to an image, it is unlike an image in that it is inherently view-independent. This is because it bends or "wraps" around a three-dimensional head surface. Indeed, using standard computer graphics, the reflectance map can be wrapped around the head surface and rendered from any viewpoint and under any illumination conditions. The possibility of computing and making predictions on these two divisible components of the faces has been explored computationally (O'Toole, Vetter, Troje & Bülthoff, 1997a), though to our knowledge not psychophysically (though see Kersten, Troje & Bülthoff, 1996, for a discussion of the nature of this representation without its associated surface map).

The roughly aligned, unprocessed pixel and surface codes have been criticized in recent years due to the fact such codes do not maintain a perfect registration of the classically defined discrete features in faces (e.g., the tip of the nose, etc.). Although not a theoretical problem for cognition, the problem for face synthesis (as we will see shortly) is more serious. It is worth noting explicitly that the question for researchers who use image-based codes is not *whether* to align the faces into a common coordinate system, but rather *how much to align them.* It must be emphasized that *all* current computational models do assume some degree of alignment.

whereas more complex spaces where a face is itself a function or manifold may be required or useful in other situations (see the chapter by Townsend et al., this volume).

We would suggest that faces may be an ideal "guinea pig" stimulus for combining the computational and information processing approaches. There are two reasons for this. First, as we will discuss shortly, faces comprise a single important category of objects for humans. As such, computational models can operate exclusively and successfully within the boundaries of this category. This may render the problem somewhat easier than the more general problem of visual scene analysis, for example. Second, perhaps due to the convenient limitations provided by the first reason, computational models of faces currently exist in a variety of alternative forms and can be compared empirically in terms of their accord with human data. This potentially limits the need for the information processing models to rely too heavily on untested assumptions about the nature of the human perceptual evidence space. Although there are not yet definitive data on the nature of the perceptual evidence spaces, we think the methods needed to identify such processes are available. We believe that it is just a matter of time before these methods are applied to a sufficient number of relevant problems to begin to provide informative answers about the nature of the perceptual evidence, as well as the prospective pattern spaces.

Quantifying the Information in Faces

Given an understanding of the kinds of tasks we must accomplish with human faces and a representation framework based on a complex multidimensional space, we now consider the logical components of implementing these notions more concretely. The first step is to "encode" or internalize a face from an external stimulus. In short, this requires a quantification of the information in a human face that is sufficient to accomplish the task(s) at hand (e.g., recognition, categorization, etc.). As noted, though not a primary focus of work in the information processing tradition, the issue of psychologically relevant and computationally expedient encoding systems has received much attention in the computational face processing literature. Indeed, the encoding assumptions of any computational model of facial cognition comprise perhaps the most important factor in determining the operation and characteristic behaviors of the model. The level of specificity of these assumptions ranges widely in the literature and has been determined primarily by the theoretical focus and goals of the particular model. Whatever the level of specificity, however, these initial representational assumptions constrain the types of relations and operations that can be applied to the different pattern and evidence spaces.[5]

The purpose of this section is to review the kinds of encoding systems that have been used in computational models and to examine both the advantages and disadvantages of these systems as models of human representation. As noted previously, we have tried to be very thorough in the review of these codes, and so readers who are somewhat less interested in the issue of encoding may skip through to the next section.

The earliest computational models for facial processing began with abstract geometric codes (e.g., important facial dimensions such as distance between eyes) and/or verbal labels that described the features of a face in much the same way as a human eyewitness might do, e.g., brown hair, brown eyes, light skin, (cf., Laughery, Rhodes & Batten, 1981, for a comprehensive review of these codes and the logic behind them). The primary problem with such codes is that they are often not adequate for quantifying and communicating enough information about an individual face to distinguish it from the multitude of competing similar candidates. In addition, such codes (particularly the verbal) are not rich enough to allow them to be used as sources of evidence at a variety of levels of task analysis. The limits of abstract, descriptive codes are also well-documented in the human eyewitness identification literature (c.f., Deffenbacher & Horney, 1981; Schooler & Engstler-Schooler, 1990). For this reason, most computational models of face processing have used a feature code that retains a more complete representation of the basic perceptual information in faces, at least at the initial stages of computation. Raw image or surface-based codes are an example of these, preserving shading and contour information that would be eliminated from the discrete codes.

Raw Image and Surface Codes

Among the variety of raw, unprocessed codes that have been used for recognition algorithms, the

[5] As one of the recurring themes in the chapters that follow is the intellectual background to the various approaches and models, we think it appropriate to note that this process of model constraint resulting from initial representational assumptions is one with a long history in contemporary cognitive psychology. Indeed, such constraints were explicitly considered in the types of control processes that could operate on an informational architecture in what arguably (e.g., Baddeley, 1986; Neath, 1998) may be the information processing model to have perhaps the most pervasive impact on contemporary cognitive psychology (Atkinson & Shiffrin, 1968).

coding system (e.g., image data, discrete features). These spaces are generally derived using linear systems analysis procedures (e.g., principal components analysis, PCA; metric multidimensional scaling; for additional discussion, see the chapter by Steyvers & Busey, this volume) applied to the face physical-similarity data, coded using the particular encoding system. We will discuss the application of PCA to this problem in detail later in this chapter. Physical face spaces are psychologically relevant only in so far as they can be shown to (a) be systematically related to psychological spaces, or (b) succeed in predicting either qualitative or quantitative aspects of human performance on face processing tasks. An important test of the relevance of individual computational models rests on their accord with the predictions made by psychological face spaces derived from human empirical data.

The Pattern and Evidence Spaces

The multifaceted nature of the face space concept in the information processing literatures can be seen in "pattern" and "psychological evidence spaces," which have been proposed by Townsend and Thomas (1993) in the context of arguing for the embodiment of abstract models of information processing in terms relevant for perceptual pattern analysis. From either the computational or information processing perspective, two important and closely linked problems must be considered in order to encode and represent faces. The first involves quantifying individual faces in terms of a set of "features" or other measurable aspects. This procedure then yields a quantifiable encoding system for faces. The second problem involves the construction of a representation of all faces, from which psychological predictions involving more than one face (e.g., confusability, similarity, etc.) can be made.

We consider a pattern space to be a relatively low-level representation wherein the effects of a set of stimuli can be specified or located. A psychological feature space (as in models specific to facial cognition, or more general models of memory derived from dimensional representations of abstract features, e.g., Hintzman, 1986; Nairne, 1990; Shiffrin & Steyvers, 1997a; Valentine, 1991), is a natural example of a pattern space. The term "relatively low-level" is advisable, since for some tasks, the psychological pattern space might involve a fair degree of processing leading up to a representation suitable for a particular task environment. An evidence space is juxtaposed with but follows a pattern space and is intended to deliver task-specific measurements or evidence pursuant to successful performance on that task. In classic signal detection theory, for example, the pattern space would consist of an ensemble or set of possible signal patterns, for instance, sinusoids embedded in noise. Then, a particular such stimulus would be submitted to one or more filters or templates and the outputs of those filters would serve as evidence for any of a number of candidate stimuli. Thus, an ideal detector might compute the likelihood (evidence) of the presence or absence of a signal, based on the activation caused by the original pattern. It can be seen that there could exist more than one evidence space (and perhaps even more than one pattern space) attendant on a particular situation. Unfortunately, pattern and evidence spaces have been confounded in many areas of cognitive research, for at least two reasons. First, there is a general lack of detailed process models for many cognitive applications of signal detection theory. Second, the log-likelihood statistic of even a multidimensional Gaussian-distributed pattern is itself normally distributed.

It is worth observing that despite the potentially complementary nature of the information processing and computational approaches, they have not yet been applied in concert to the problem of face representation and processing, nor generally speaking, to problems in cognition in general. Even the ubiquitous information processing approach itself has not been immune to this deficit. As Townsend and Thomas (1993) observed, the information processing approach to perception, memory, and elementary cognition has in the past been ironically marked by a total absence of specification of the information being transformed, remembered, etc.. The closest one usually gets is an abstract notion of "features" or other attributes that are not concretely delineated in terms of the stimuli and/or mental events.

One possible reason for this is that perceptually realistic pattern spaces (e.g., those capable of housing representations that retain the richness of a visual object or scene, for example) are enormously data-intensive. Even when attempting to use tractable, analytically-based information processing methods, it is easy to loose one's logical bearings in representational spaces with the kind of topographical complexity and high dimensionality that even a simple structural analysis of the perceptually-based evidence can yield. Furthermore, the attendant "space" may be distinct for, say, search for a certain kind of eyes in stimulus faces, as opposed to more holistic tasks, such as recognition (see e.g., Uttal, Baruch & Allen, 1995a,b, as well as the chapter by Uttal, this volume). Thus, a face as a point in a finite dimensional feature space may be appropriate for certain kinds of tasks ity/femininity. People are also quite willing to rate faces for personality characteristics (e.g., generosity, friendliness, etc.).[2]

Computational models of facial cognition must grapple with the problems presented by the competing nature of the facial information useful for perceptual categorization, recognition, and identification. In particular, to categorize a face, one must extract the information that a face *shares* with an entire category of faces (e.g., male faces). By contrast, to recognize a face, one must extract the information that makes the face unique or *different* from all other faces in the world (related points are discussed further in the chapter by Cottrell et al., this volume). The dichotomous nature of this information is an important factor in understanding the design of computational models aimed at solving perceptual or memory tasks. We will address this problem in more detail in the section on encoding faces.

Face Space Framework for Representing Faces

Representing sets of faces in a way that enables a recognition judgment or a perceptually-based categorization (e.g., a sex classification) requires a system capable of comparing individual faces with information structures representing other individual faces and with groups of (presumably) known faces in memory.[3] In both the computational and psychological face literatures, the most common theoretical framework for doing this relies on the general, abstract notion of a "face space." Perhaps the most prominent use of this notion can be found in the work of Valentine (1991, see also the chapter by Valentine, this volume), who explicitly invoked this construct as a psychological model of face processing. To quote from one of our own uses of this notion, "A generic face space representation includes only a few basic conceptual elements: (a) faces can be thought of as points in a high dimensional space, (b) the dimensions or axes of this space represent the different 'features' that we use to encode the faces, and (c) the distance between any two faces in this space is a measure of their similarity" (Deffenbacher, Vetter, Johanson & O'Toole, 1998). Generally, this face space is assumed to possess a structure along the lines of a multivariate probability distribution, with the central tendency of this distribution corresponding to the notion of a prototype or average face, and with the density of faces decreasing as a function of the distance from the central tendency.

Like many constructs in contemporary psychology, the notion of a face space has intuitive appeal from a variety of perspectives. This broad appeal has led to some imprecision in the use of the construct. To our knowledge, there are three common referents for the notion of a face space: (a) abstract, (b) psychological, and (c) physical. We will discuss each in turn.

Valentine's model exemplifies the *abstract face space* and can account for the well-known finding that distinctive faces are better recognized than typical faces (e.g., Light, Kayra-Stuart & Hollander, 1979). It can account also for the finding that typical faces are classified as faces more quickly than unusual faces (Valentine & Bruce, 1986). The former occurs presumably due to the fact that typical faces are close to the average face (i.e., in the densest part of the face space) and so are more likely to be confused with other faces. The latter finding occurs presumably due to the fact that classification of the face requires a comparison to the average face, which is assumed to occur more quickly/efficiently for faces close to the average than for faces farther away. These abstract notions recapitulate mechanisms for recognition and categorization that have been explored in more general terms using traditional information processing models (see, e.g., Ashby & Perrin, 1988; Ashby & Alfonso-Reese, 1995; Nosofsky, 1988a, 1991; Palmeri, 1997).

Psychological face spaces are multidimensional descriptions of the perceptual similarity relationships among a set of faces. Psychological face spaces can be derived from human empirical data on the perceived similarity between all possible pairs of faces in some set (e.g., Johnston, Milne, Williams & Hosie, 1997; Kruskal & Wish, 1978; Young & Householder, 1938, see also Busey, and Steyvers & Busey, this volume). These spaces are generally the result of multidimensional scaling procedures applied to human data on face similarity judgments.[4] What results from this kind of analysis is a multidimensional representation of the *perceptual* similarity among a set of faces.

Physical face spaces are derived from the *physical* similarity between all possible pairs of faces in some set, defined relative to a specific facial en-

[2] Before dismissing this willingness as trivial, it is well-worth noting that one of the oldest and most robust findings in the face recognition literature is that faces rated for "deep" characteristics such as personality traits are more accurately recognized than faces rated for "surface" characteristics (e.g., features or gender) (Bower & Karlin, 1974).

[3] Such has to be the case for any conception of memory, be it a pattern of weighted connections or, as is generally the case for models in cognitive psychology, "items" stored in a place (Roediger, 1980).

[4] There is a very close relationship between multidimensional scaling and the statistical models we will consider with respect to the computational models.

groups of faces in a more concrete embodiment of a face space, and (c) accessing the results of the encoding/representation process to solve a particular task (e.g., recognition, categorization by gender, etc.). In the initial (encoding) section, we will review the kinds of codes that have been implemented and point out the advantages and disadvantages of these codes. We will devote a fair amount of space to this endeavor for two reasons. First, we believe the issue of appropriate codes comprises a very important component of the human representation. Second, to our knowledge, no thorough and up-to-date review of this topic exists. In any case, readers less interested in this issue can skip over this section as it is self-contained.

In the second section (representing groups of faces), we make a link between the abstract face space and a more concrete computational embodiment in a complex multidimensional space. We will also consider the potential of navigating through these computationally-defined face spaces via morphing. Finally, in the third section, we will see how the information processing literature provides a very rich structure for understanding how information can be accessed from representational spaces and be put to the service of particular tasks. We will review these structures in the context of problems in computational face cognition. Although the application of these structures to facial cognition is just beginning, we will sketch out the potential of these approaches for complementing the lower-level, perceptually based computational approaches.

Before proceeding, we wish to note that it is not completely clear that the encoding and task components of the problem are divisible in this neat way. Much of the research to date, however, makes the working assumption that these components are logically independent, and so we present the discussion that follows in that light, pointing out as we go the limitations of this assumption.

The Tasks

Like any other visual pattern, a face can constitute an input that can be put to a multiplicity of psychological uses. The same basic pattern can support the recognition of a loved one, the discrimination of sex and race, and the rapid and complex psychological processes that result in attraction, fear, or bigotry.

Perceptual and Memory Tasks with Faces

Face processing tasks can be divided into "perceptual" and "memory"-based tasks. The primary difference between the two is that memory tasks require an observer to retain information about a *particular* face over time, whereas perceptual tasks do not. Memory tasks include recognition,[1] defined here as a decision about whether a face has been encountered previously, and memory-based identification, defined here as the retrieval of the visual information specific to a face, a name, or some other semantic label for a face. "Recall" is not, strictly speaking, a *response* construct in face memory (as it might imply a response such as drawing the face). It is, however, a *task* construct, in the sense that we are able to retrieve information specific to an individual face (e.g., a cued recall situation in which the face is the cue).

An additional, though perhaps not entirely independent task construct is face recognition "generalization" (or, generalization, for short). Generalization is a task construct used commonly in the object processing literature to refer to an observer's ability to identify or recognize an object given a novel view, or under novel lighting conditions. Although in real life, all recognition tasks entail some degree of generalization, radical mismatches in viewpoint and/or lighting between the studied and test stimulus can make face processing very difficult (e.g., Hill & Bruce, 1993), just as mismatches in stimulus characteristics can impact performance in other cognitive tasks (e.g., Blaxton, 1989; Weldon & Coyote, 1996; Weldon, Roediger & Challis, 1989; Wenger & Payne, 1997). In fact, much useful information about human representations of faces has been obtained from examining the kinds of generalizations we *do not* perform efficiently or accurately. For example, recognition of upside-down faces and faces in the photographic negative are notoriously difficult (e.g., Galper & Hochberg, 1971; Yin, 1969). These results may indicate that certain quite literal aspects of the presented image may be retained in internal representations.

Examples of perceptual tasks include feature and face detection, categorization (e.g., by sex, race, age), facial expression analysis, and same/different judgments. These tasks might also include our ability, or perhaps, more appropriate, our "readiness" to rate faces along a variety of dimensions including attractiveness, typicality, familiarity, and masculin-

[1] The term "recognition" has been used in multiple domains to mean multiple things, including identification (e.g., "recognizing" a face as being someone in particular), discrimination (e.g., "recognizing" that a particular face is the same or different from some other face), and the task of determining whether one has seen a particular face before (e.g., "recognizing" a face as being one that was seen in an earlier encounter). We will prefer the third referent in this chapter as it is the one used most commonly in the face literature.

Quantitative Models of Perceiving and Remembering Faces: Precedents and Possibilities

Alice J. O'Toole
The University of Texas at Dallas *

Michael J. Wenger
University of Notre Dame

James T. Townsend
Indiana University

The purpose of this initial chapter is to provide a general sketch of the problems that must be addressed in quantifying the representations and processes involved in converting an image of a face into a meaningful cognitive object. Our approach is to define a structural model of the process components and to point out implicit connections among the logical, computational, and psychological pieces of the problem. We will consider especially potential connections among computational, psychophysical, and traditional information processing approaches that have not been made explicit in the literature. In this way, we hope to help readers organize the chapters that follow into the mutually supportive component parts of a complicated, multifaceted, but (we believe) ultimately tractable problem.

We also hope to convince the reader who may not regularly consume computational and/or mathematical material that the approaches we discuss are natural formalizations of theoretical concepts with which most psychologists are quite familiar, and that the development and application of these formalizations are well-worth pursing. We think that mathematical and computational formalizations of the problem can provide a unified framework for understanding face processing at a level that supersedes its individual components. Although the field is still far from doing research at this "higher" level, we can begin to see where some of these potential lines of inquiry may lie.

The psychological concepts that we believe readers will find familiar in this chapter concern issues of measurement, representation, and task demands. These are issues encountered in nearly all models of psychological phenomena. Specifically, we will ask the following kinds of questions. How do you measure the information in a stimulus when the stimulus is a face? How do we represent sub-categories of stimuli, e.g., for faces, male and female, young and old? Finally, how do the demands of the task and the nature of the processor constrain our access to and use of the information in the representation?

This chapter is organized as follows. We first give a brief overview of the kinds of tasks we must accomplish with human faces. This defines the nature and diversity of the output that computational models must produce to be considered successful. We next present a general abstract framework for face perception and memory in terms of a "face space" representation (e.g., Valentine, 1991, see also chapters by Busey, and Steyvers & Busey, this volume). This model implicitly or explicitly underlies much current psychological and computational work on face processing.

Throughout this chapter we will see a dualism between concepts and theories that are well-known in the mathematical psychology literature, and those that have emerged more recently in the computational face literature. The notion of a complex, multidimensional representational space is an example of this dualism. In the mathematical psychology literature this notion has been presented in terms of constructing "pattern" and "evidence" spaces on which one can operate with traditional human information processing procedures (e.g., Townsend & Thomas, 1993). In the psychological and computational modeling literature, the term "face space" has been likewise employed. We will see that these two approaches are complementary and that much can be gained by combining the tools, techniques, and insights that come from each.

Once we understand the kinds of tasks we must accomplish with human faces, we next divide the problem of facial cognition into its logical components, which consist of (a) "encoding" or internalizing a face from an external stimulus, (b) representing