Journal of Expertise
2018. Vol. 1(2)

ISSN 2573-2773

The Limits of Expertise in Face Recognition: Response to Sunday and Gauthier's "Face Expertise for Unfamiliar Faces: A Commentary on Young and Burton's 'Are We Face Experts?'"

Andrew W. Young and A. Mike Burton
University of York

Correspondence: Andy Young, andy.young@york.ac.uk

Abstract
Sunday and Gauthier (2018) raise a number of interesting points about Young and Burton's (2018) revisionist opinion of expertise for recognizing face identity. We are happy to clarify and debate these matters by explaining further why we proposed independent criteria for expertise, how we see the roles of perceptual experience and the everyday demands of face recognition, why we agree that individual differences are important, why we don't accept the idea of a linear hierarchy of perceptual recognition difficulty from unfamiliar to familiar faces, and the relevance of things we can indeed see in unfamiliar faces.

Keywords
face recognition, expertise, familiarity, familiar faces, unfamiliar faces

Response

As heretics from the prevailing orthodoxy, we realized that our rethink of expertise for recognizing face identity (Young & Burton, 2018a - henceforth Y&B) would ruffle a few feathers. Sunday and Gauthier's (2018) commentary (S&G) defends and brings up to date a perspective that has been developed across more than 30 years, since the study by Diamond and Carey (1986). Although Y&B pointed out that this perspective has never been accepted by everyone (see, for example Kanwisher, 2000), it has nonetheless been widely used.

To put the debate in context, Y&B pointed out that the view defended by S&G is one that considers humans to have acquired expertise at perceiving and recognizing the unique identities of *both* unfamiliar and familiar faces, whereas Y&B maintained that expertise for face identity is largely restricted to recognition of *familiar* faces. Importantly, Y&B didn't claim that humans are completely unable to recognize the identities of unfamiliar faces, only that identity tasks with unfamiliar faces are difficult, and our comparatively poor performance makes the "expert" label unwarranted. Neither did Y&B deny that we can see lots of things other than identity in unfamiliar faces; for example, we might be considered relatively expert for judging the sex, age, or expression of all faces. The core of the debate, then, concerns expertise for face identity (see also Rossion, 2018; Young & Burton, 2018b).

There are some points of substantial agreement; S&G accept the importance of understanding how familiar and unfamiliar face recognition may differ, and they accept that "familiarity is acquired one face at a time, as we

learn idiosyncratic information about each person." These were cornerstones of Y&B's discussion (see also Young & Burton, 2017; Young, 2018).

Having said that, we do still disagree with S&G over the interpretation of a number of other points they raise. Over the years, we have reached the conclusion that the idea of generalized visual face identity expertise that can encompass recognition of unfamiliar face identities is in some ways misleading. Some of the issues raised by S&G are ones we were not able to discuss at length in the original (necessarily short) Y&B opinion piece. We welcome the opportunity to clarify our views on these.

One focus of disagreement involves how to define and measure expertise. In drafting Y&B, we were surprised to discover that we mainly found operational definitions in the literature on face expertise. For example, claims that expertise can be measured through holistic processing or inversion effects are in effect operational definitions that assume (rightly or wrongly) that these will be consequences of expert performance. We don't claim that these operational definitions are implausible, but they do run a risk of circularity; for example, through claiming that face recognition is expert because there are substantial inversion effects and that inversion effects reflect expertise because they are found for faces.

To avoid this circularity, Y&B decided to seek independent criteria for expertise. We drew on widely-used theories concerning perceptual learning and skill acquisition (e.g. Logan, 1988; Shiffrin & Schneider, 1977) to suggest that expert performance should be based on substantial experience, should result in accurate performance, and should become relatively automatic. We then evaluated studies of face recognition against these criteria. We accept S&G's remark that this was Y&B's "own definition of expertise," and we agree with their view that expertise is not simply an all or none phenomenon. We think, however, that offering a set of independent criteria is a step forward.

In previous work, Gauthier et al. (2000, p.191) had suggested two key characteristics for face expertise: "First, faces are recognized at a more specific level of categorization (for example, 'Adam') than most objects (for example, 'chair' or 'car'). Second, although we are experts with faces, we have much less experience discriminating among members of other categories." These suggestions are worth considering carefully. The first characteristic (level of categorization) is in some respects consistent with Y&B's emphasis on the importance of familiarity. However, we doubt that the entry level category for an unfamiliar face is its unique individual identity ("Adam"), though we do agree that it is certainly likely to be at a higher level than the basic-level category of "a face"; perhaps something more like "young man," or even "friendly-looking young man." Nonetheless, it is mainly in terms of the second of Gauthier et al.'s (2000) characteristics of expertise (the role of experience of discriminating between face identities) that we differ from S&G. We will explain why.

Using the criteria for expertise we proposed, Y&B (p.102) observed that "Substantial, lifelong perceptual experience is therefore undoubtedly important to face recognition, and in this sense the first of our criteria for face expertise does seem to be applicable." Evidence cited by Y&B (p.101) included perceptual narrowing in infants and other-race effects in adulthood; to that extent, we agreed with S&G. However, Y&B then went on to note that their other criteria for expertise (accurate and automatic performance) are not met for unfamiliar face recognition and concluded (Y&B, p.208) that "While there is undoubtedly a sense in which our huge experience of looking at the faces around us has created a type of expertise, this does not take the form of the generic ability to recognize any face identity that has been so widely assumed" and that "Although clearly shaped by experience in some ways, as shown for example by other-race effects, unfamiliar face recognition remains generally vulnerable to the impact of the enormous variability in everyday images of faces."

Some of the points raised by S&G reiterate Gauthier et al.'s (2000) emphasis on the impact

of experience, which (as noted above) Y&B never sought to deny; indeed, our own studies have also shown the critical importance to learning new faces of presenting stimuli in typically experienced formats (Kramer, Jenkins, Young & Burton, 2017). Our claim is rather that while experience undoubtedly plays an important role in shaping our face perception abilities, it does not shape them specifically toward recognizing unfamiliar face identity.

There is also an important difference between experience in the sense of mere exposure to a class of visual stimuli and the type of experience that creates expertise. S&G's take on this is that the essential component underlying expertise is the functional need to discriminate between and individuate members of a visual category, which they assume happens often for unfamiliar faces in everyday life. We are not so convinced that precise individuation of unfamiliar faces is actually such an over-riding everyday demand; there are many circumstances in which full individuation of an unfamiliar face is not needed. We suspect that many readers will, for example, have had the experience of assuming that any young man who comes to their table in a restaurant is the waiter and then failing to notice if a different waiter subsequently arrives.

Such anecdotes are strikingly borne out by demonstrations of change blindness with unfamiliar faces (Simons & Levin, 1998). In contrast, if everyday individuation of all seen faces was in itself sufficient to create generic expertise for face identity, Y&B pointed out that performance at tasks such as unfamiliar face matching would be less variable than has proved to be the case, eye witness reports would be less fallible, and passport officers would make fewer mistakes. Indeed, if individual identity is the dominant level of categorization of all faces, phenomena such as change blindness for unfamiliar faces (Simons & Levin, 1998) could not occur. Our opinion is instead that it is individuation of familiar faces that is the critical everyday ability, since it allows us to go beyond immediate contextual categories such as "the waiter" and retrieve the relevant identity-specific information that allows us to interact appropriately with people we know in light of our past experience with them (Bruce & Young, 1986; Kramer, Young et al., 2017; Young, 2018).

This brings us to the issue of variability across individuals. Until the late 1990s, most researchers (including us) tended to assume that nearly all neurologically normal perceivers would perform at a fairly uniform overall level on face perception tasks, and in consequence little attention was paid to individual differences. However, it has turned out that the performance of tests of unfamiliar face matching and recognition by neurologically normal individuals can be remarkably variable; ranging from chance-level to near-perfect (Burton, White & McNeill, 2010). These individual differences form an important phenomenon. Y&B used them to emphasise that unfamiliar face recognition isn't consistently expert-like because the criterion of accurate performance isn't met by most people, but we fully agree with S&G that the reasons for these individual differences in performance are interesting and that studying them is a useful enterprise. We certainly weren't seeking to "discourage interesting research avenues."

However, we suspect that when we have a better understanding of individual differences, their causes will turn out to be different from those emphasised by S&G. Part of the reason can be seen in S&G's Figure 2, which shows a purported hierarchical continuum of difficulty of object recognition judgements that is claimed to run from easiest (S&G level 1 - basic category level recognition) to hardest (S&G level 4 - subordinate level recognition of familiar exemplars). Figures like this occur in many textbooks and exemplify one version of a traditional view of recognition. Although they can exert a powerful intuitive appeal, we think they are incorrect. Instead the relatively high error rates in unfamiliar as compared to familiar face recognition show that, if anything, S&G's level 4 recognition is not the most difficult; often it is easier to achieve than S&G level 3 (subordinate level recognition of unfamiliar exemplars).

These observations follow naturally from Y&B's point (see also Burton, 2013; Burton, Kramer, Ritchie & Jenkins, 2016; Young & Burton, 2017) that the variability of different everyday images of the same person's face is to some extent identity-specific; for example, the underlying dimensions across which images of the face of (say) Tom Cruise vary will be different from the dimensions of variability for the face of Hugh Jackman (Burton et al., 2016). Recent computational studies of face familiarity (Kramer, Young, Day & Burton, 2017; Kramer, Young & Burton, 2018) have shown how it is possible to create a model that can cope with everyday image variability to accurately identify images of a trained set of familiar faces (level 4 recognition in S&G's terms). Yet the same model doesn't perform nearly as well with images of unfamiliar (i.e. untrained) faces (S&G level 3); it shows only limited ability to generalize beyond the trained set of identities. It is the existence of identity-specific variability that makes S&G level 3 intrinsically difficult, because the nature of the variability between images of an unfamiliar face is unknown when it is encountered for the first time.

At a deeper theoretical level, Y&B's perspective emphasizes recognition of familiar faces as a primary social task that is essential to appropriate interactions based on previously stored identity-specific knowledge of an individual (cf. Bruce & Young, 1986; Kramer, Young et al., 2017; Young, 2018). For unfamiliar faces, there is no stored identity-specific knowledge and context alone will often supply much of the information we need; the person who comes to your table in a restaurant is probably the waiter, the person behind the hotel front desk will be the receptionist, and so on. In Bruce and Young's (1986) terms, everything else we can tell from an unfamiliar face must involve visually-derived semantic information. This is not to deny characteristics that may be more or less closely related to identity can be seen in unfamiliar faces and that visually-derived information can be very rich, encompassing the perception of age, gender, ethnicity, expression, gaze direction and a variety of social attributions (Bruce & Young, 1986, 2012; Todorov, 2017; Young, 2018). It is of course important to understand how we learn to do these things with unfamiliar faces, but we consider that these abilities do not necessarily involve expertise for analyzing unfamiliar face identity *per se*.

This is a significant point of difference from S&G's position, since their claim is that face expertise derives from constantly seeking to individuate unfamiliar faces in everyday life. Instead, Y&B noted that computational studies show that at least some visually-derived semantic information (gender and race) can be a consequence of learning to recognize a relatively small number of familiar face identities without any additional form of training involving unfamiliar faces (Kramer, Young et al., 2017). Moreover, our simulations show that training a model to be able to recognize everyday images of familiar faces can also lead to a modest but measurable benefit on recognizing the identities of unfamiliar faces, again without any training for unfamiliar face identity itself (Kramer et al., 2018).

This small benefit to recognizing unfamiliar face identity accrues as an indirect consequence of the fact that local regions of a purely perceptual space must be reshaped in order to cope with identity-specific variability; in effect bringing different images of Tom Cruise closer together than they would be based on their purely visual attributes (see Kramer et al., 2018). It seems that some of this reshaping can benefit the encoding of unfamiliar face identity, but we reiterate that this happens to only a limited extent. The observation is consistent with one of the main points emphasised by S&G—that most of us are able to recognise unfamiliar face identity to some degree—but in Kramer et al.'s (2018) study it was simply a by-product of creating expertise in recognizing a specific set of familiar faces in a system that received no training whatsoever for recognizing the identities of unfamiliar faces.

An issue that Y&B didn't address concerns whether face perception has domain-specific properties (cf. Kanwisher, 2000; Rossion, 2018) or is simply the same in nature as any other form of visual expertise (Gauthier et al., 2000).

This issue has in the past generated much discussion because it relates to the wider question of the extent to which apparent brain specializations for face perception might instead be consequences of expertise (Gauthier et al., 2000; Kanwisher, 2000). Y&B set aside the issue because we think that it cannot at present be resolved without a better understanding of the sense(s) in which it may be appropriate to say that we are face experts. However, we note that substantial parts of S&G's commentary are predicated on the assumption that face expertise is no different from any other form of visual expertise. We do not have a firm view on this matter but, given our emphasis on the variety of social signals derived from faces, it seems to us likely that at least some aspects of face perception will prove to be domain-specific, whether or not they qualify as expert.

Understanding human face recognition ability is of substantial theoretical and practical importance, so it is essential to discuss and debate these issues. In our opinion, S&G's commentary offers a fair, clearly-expressed, and courteous restatement of what has been a mainstream view of face expertise that sets a useful tone for this debate. We have tried to respond in kind, aiming to set out points of disagreement clearly and dispassionately, so that readers can form their own balanced opinion.

Authors' Declarations

The authors declare that there are no personal or financial conflicts of interest regarding the research in this article.

References

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Bruce, V., & Young, A. (2012). *Face perception.* Hove, East Sussex: Psychology Press.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology, 66*, 1467-1485.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40*, 202–223.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods, 42*, 286-291.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General, 115*, 107-117.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*, 191-197.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*, 759-763.

Kramer, R. S. S., Jenkins, R., Young, A. W., & Burton, A. M. (2017). Natural variability is essential to learning new faces. *Visual Cognition, 25*, 470-476.

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review, 124*, 115-129.

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46-58.

Logan, G. D. (1988). Automaticity, resources, and memory: theoretical controversies and practical implications. *Human Factors, 30*, 583−598.

Rossion, B. (2018). Humans are *visual* experts at unfamiliar face recognition. *Trends in Cognitive Sciences, 22*, 471-472.

Shiffrin, R. M., & Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127−189.

Simons, D. J. & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5*, 644-649.

Sunday, M. A., & Gauthier, I. (2018). Face expertise for unfamiliar faces: A commentary on Young and Burton's "Are We Face Experts? *Journal of Expertise, 1*, 35-41.

Todorov, A. (2017). *Face value: the irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Young, A. W. (2018). Faces, people and the brain: the 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71*, 569-594.

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science*, *26*, 212-217.
Young, A. W., & Burton, A. M. (2018a). Are we face experts? *Trends in Cognitive Sciences*, *22*, 100-110.
Young, A. W., & Burton, A. M. (2018b). What we see in unfamiliar faces: a response to Rossion. *Trends in Cognitive Sciences*, *22*, 472-473.

.

Received: 1 June 2018
Revision received: 3 August 2018
Accepted: 3 August 2018