Journal of Experimental Psychology:
Human Perception and Performance
2001, Vol. 27, No. 3, 538–541

0096-1523/01/$5.00 DOI: 10.1037//0096-1523.27.3.538

Illusory Conjunctions Are Alive and Well: A Reply to Donk (1999)

William Prinzmetal, Jörn Diedrichsen, and Richard B. Ivry
University of California, Berkeley

When presented with a red *T* and a green *O*, observers occasionally make conjunction responses and indicate that they saw a green *T*. These errors have been interpreted as reflecting separable processing stages of feature detection and integration with the illusory conjunctions arising from a failure at the integration stage. Recently, M. Donk (1999) asserted that the phenomenon of illusory conjunctions is an artifact. Conjunction reports are actually the result of confusing a nontarget item (*O* in the example above) for a target item (the letter *T*) and (correctly) reporting the color associated with the (incorrectly) selected target. The authors demonstrate that although target–nontarget confusion errors are a potential source of conjunction reports, there is a plethora of findings that cannot be accounted for by this confusion model. A review of the literature indicates that in many studies, illusory conjunctions do result from a failure to properly integrate features.

Many theories of object recognition emphasize that perception involves an analytic process in which image components are initially extracted and then integrated to form coherent objects. Critical components have been variously described in terms of spatial frequencies (e.g., De Valois & De Valois, 1980), geons (Biederman, 1987), or features (e.g., Selfridge, 1959). Regardless of the components of object recognition, all recognition by component models face a similar problem. Given a complex scene with many different objects, the visual system must appropriately combine the components. This problem is referred to as *the binding problem.*

The visual system is apparently adept at solving the binding problem: In everyday life, we are not aware of combining image features. Indeed, until the pioneering work of Treisman (e.g., Treisman & Gelade, 1980) and Wolford (1975), the problem of feature binding was generally neglected. However, the seminal work of Treisman and Schmidt (1982) described several experimental paradigms that made the problem of feature binding manifest and also provided a method for studying how the visual system combines visual features (also see Wolford & Shum, 1980). Treisman and Schmidt called this phenomenon *illusory conjunctions* or errors in feature binding. For example, when briefly presented with a multicharacter display that contained a red *T* and green *O*, observers sometimes reported that the display contained a green *T*.

In the past two decades, the phenomenon of illusory conjunctions has proven useful for exploring feature binding. Experiments involving illusory conjunctions have been productive in determining the stimulus information that the visual system uses to bind features (see Prinzmetal, 1995, for a review), the role of attention in feature binding (e.g., Cohen & Ivry, 1989; Kleiss & Lane, 1986; Prinzmetal, Henderson, & Ivry, 1995; Prinzmetal, Presti, & Posner, 1986), the effect of perceptual organization on feature binding (Baylis, Driver, & McLeod, 1992; Khurana, 1998; Prinzmetal, 1981), and cognitive factors in feature binding (e.g., Prinzmetal, Hoffman, & Vest, 1991; Prinzmetal & Keysar, 1989; Prinzmetal & Millis-Wright, 1984; Prinzmetal, Treiman, & Rho, 1986; Treisman, 1986). The neuropsychology of feature binding has been explored (e.g., Arguin, Cavanagh, & Joanette, 1994; Cohen & Rafal, 1991; Friedman-Hill, Robertson, & Treisman, 1995). Finally, several formal models have been proposed to account for illusory conjunctions (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Logan, 1996; Wolford, 1975). It is important to note that the studies mentioned here are only a sample of the work using illusory conjunctions to understand feature binding.

One problem encountered in all of these experiments is how to measure binding errors separately from errors that result from misperceiving features. An experiment by Prinzmetal et al. (1995, Experiment 1) illustrates this problem. Observers were presented with an array of letters in the periphery for 5 s while they simultaneously monitored a string of digits at fixation. The array of letters consisted of a colored target letter that could be either the letter *T*, *X*, or *L* and a colored nontarget letter (the letter *O*). These two colored letters were flanked by two white *O*s. Following the presentation of the letters, the observers were required to report the identity of the target (i.e., *T*, *X*, or *L*) and its color (i.e., red, green, or blue). Consider the situation where the display contained a red *T* and green *O*. There are six possible response categories (shown in Table 1). For example, observers responded with the correct response (C) on 77% of the trials. Conjunction responses (CRs) occur when the observer reports the correct target letter and the color of the nontarget letter. In this experiment, conjunction responses occurred on 17% of the trials.

Not all of the CRs represent incorrect feature binding. Some proportion of them undoubtedly result from misperceiving either the target color or the target letter and guessing a color or letter. To correct for guessing, Prinzmetal et al. (1995) used an adaptation of the multinomial processing procedures (Batchelder & Riefer, 1999; Batchelder, Riefer, & Hu, 1994; Riefer & Batchelder, 1988) that had been developed to study illusory conjunctions (Ashby et

William Prinzmetal, Jörn Diedrichsen, and Richard B. Ivry, Department of Psychology, University of California, Berkeley.

Correspondence concerning this article should be addressed to William Prinzmetal, Department of Psychology, University of California, Berkeley, California 94720. Electronic mail may be sent to wprinz@socrates.berkeley.edu.

Table 1
Response Categories From Prinzmetal et al. (1995)

Response	Response name	Results
Red T	Correct (C)	77%
Green T	Conjunction response (CR)	17%
Blue T	Color feature error (CF)	1%
Red F or X	Letter feature error (LF)	2%
Green F or X	Letter feature conjunction response (LFCR)	4%
Blue F or X	Color–letter feature error (CLF)	0%

Note. There were three possible target letters (X, T, and L), and there were three possible colors (red, green, and blue). On each trial, one target letter and two colors were present in the display. In this example, the target was a red T and the nontarget was a green O.

al., 1996). From this method, Prinzmetal et al. estimated the probability of correctly perceiving the target color and target letter and correctly binding features. The parameter for feature binding, called α, can range from .5 (random feature binding) to 1.0 (no errors in feature binding). The parameter α in Prinzmetal et al. (Experiment 1) averaged .83, which indicates good but not perfect binding. This result was typical of the results from the experiments in Prinzmetal et al. and other studies reported in the literature.

Donk (1999) recently suggested that illusory conjunctions are an artifact. Rather than supporting that illusory conjunctions result from the incorrect combinations of correctly perceived features, Donk proposed that they occur because observers incorrectly perceive the nontarget letter as a target letter. In the example from Prinzmetal et al. (1995), consider a trial in which the target was a red T and the nontarget item a green O. If observers misperceived the O as a T, a CR would result. Previous studies of illusory conjunctions have not entertained this possibility. A thought experiment readily demonstrates the viability of this hypothesis. Suppose the target and nontarget were both Ts but differed by one pixel. Obviously, in this situation target–nontarget confusions would readily occur, and the occurrence of conjunction reports would be quite high. Donk questioned an assumption that is made by many researchers, namely that such confusions do not occur because the target and nontarget letter sets are highly distinct (e.g., X and T for the target and O and C for the nontarget).

Donk (1999) thus concluded that the phenomenon of illusory conjunctions is a chimera. This issue is important because CRs have been used as a valuable experimental tool to understand how the visual system binds features. Fortunately, a relatively straightforward demonstration can be made that the problem suggested by Donk does not apply to most of the existing data on feature integration.

Consider, as an example, the experiment discussed above. What Donk (1999) correctly pointed out is that there is a nonzero probability that the nontarget letter will be erroneously identified as the target. This possibility is depicted in Figure 1. The parameter, s, corresponds to the probability of correctly distinguishing the target from the nontarget item. Correspondingly, $1 - s$ represents the probability of making a target–nontarget confusion. In this figure, there is a probability, $1 - s$, that the nontarget O will be confused with one of the target items (T, L, or X). The consequence of misperceiving the O as a T is a CR. The consequence of misperceiving the O as one of the other target letters is what we label a *letter feature conjunction response* (LFCR; see Table 1). We do not know the confusion matrix of O with T, L, or X, but because the target letter was randomly selected, we can assume that over trials and targets, the three outcomes are equally likely.[1] Thus, misperceiving the nontarget as a target would result in twice as many responses in the LFCR category as in the CR category. However, in that experiment, there were many more CRs (17%) than LFCRs (4%). This pattern is opposite to what one would expect if the conjunction reports were actually the result of target–nontarget confusion errors given that there are two paths resulting in LFCRs and only one resulting in a CR. Indeed, we know of no previous data that conform to the predictions of Donk's confusion model.

One might propose that the higher rate of CRs is due to the fact that such errors can occur even when the target letter is correctly identified, whereas LFCRs are restricted to situations in which the target letter is not perceived. Observers on some proportion of trials might misperceive the target color (red) as the nontarget color (green). Or they may correctly identify the target letter but fail to perceive its color and guess the nontarget color. An important aspect of multinomial models is that they allow one to estimate the contribution of such errors. In most experiments, color feature errors are rare and contribute minimally to the CR category. For example, in Experiment 1 of Prinzmetal et al. (1995), color feature errors occurred on about 0.5% of the trials. These results are typical of those in the field.

Donk (1999) correctly pointed out that previous correction for guessing procedures (e.g., Ashby et al., 1996; Prinzmetal et al., 1995) does not estimate the probability of directly confusing the nontarget with the target.[2] It is important to note that for this analysis, however, it is not necessary to estimate the value of the parameter, s (see Figure 1). The relation between CR and LFCR responses is the same regardless of the value of s. Unfortunately, the above analysis cannot be applied to the results in Experiments 1–4 of Donk's article. Only two possible values were used for each dimension, and both of these values on the dimension that were used to select the target were present on every trial. Thus, the necessary response categories do not exist in Donk's Experiments 1–4.

[1] Indeed, this assumption is at the heart of Donk's (1999) model. If the nontarget letter is more likely to be confused with the target letter than a letter not in the display, CRs would be more likely than LFCRs. The influence of the target letter on the perception of the nontarget letter is the basis of illusory conjunctions.

[2] It is possible, in principle, to estimate Donk's (1999) parameter, s, which is the probability of misperceiving the nontarget for the target, from the data in Ashby et al. (1996). In some of their models, Ashby et al. not only estimated the probability of correctly perceiving the target letter (TL) but also the probability of perceiving the nontarget letter (NL). From these, an estimate of s can be obtained. The value of s would equal $1 - (1 - TL)(1 - NL)$. However, the estimates of the NL parameter proved to be unstable and had little impact on the goodness of fit. We expect this is true for two reasons. First, the $(1 - TL)(1 - NL)$ branch only influences the response outcome on a small percentage of the trials. Second, Ashby et al. did not require the observers to report the nontarget letter. The confusion parameters estimated by Donk, sc and so, appear to behave in a similarly unstable manner.

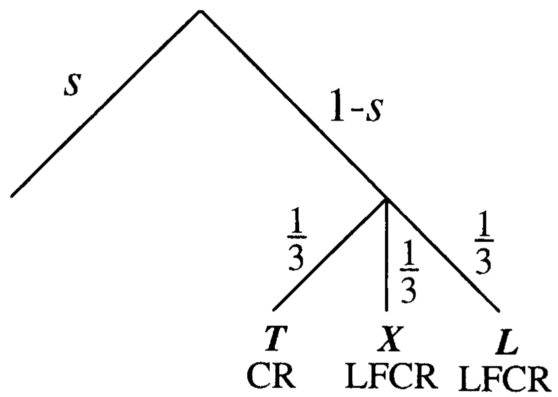

Figure 1. The consequences of confusing the nontarget with the target (*T*, *L*, or *X*) for Prinzmetal et al. (1995). The probability of confusing the nontarget with the target is $1 - s$. s = the probability of correctly distinguishing the target from the nontarget; CR = conjunction response; LFCR = letter feature conjunction response.

It is possible to use nontarget letters that are confusable with the target letter. If the nontarget is confused with the target letter, the confusion model in Figure 1 should predict the relation between CR and LFCR responses. In Experiment 5, Donk (1999) used three possible colors, two target shapes, and a nontarget shape. If conjunction reports result from a confusion between the nontarget shape and the target shape in this situation, then the proportion of CRs should equal the proportion of LFCRs. However, if at least some of the conjunction reports are actual errors in binding, then the proportion of CRs should be greater than the proportion of LFCRs. (It is important to note that in Prinzmetal et al. [1995], there were three possible targets, and thus the proportion should be twice as many LFCRs as CRs; Donk had two possible targets.)

The results of Experiment 5 are revealing. Donk (1999) used two stimulus sets. One stimulus set consisted of colored letters. Similar to many of the previous studies of illusory conjunctions, the target letters, *X* and *T*, were not particularly confusable with the nontarget letters, *C* and *S*. The proportion of CRs was more than twice the proportion of LFCRs (.13 to .06). The other stimulus set consisted of rectangles, and the target and nontarget shapes were very similar. The rectangles could be oriented with the long axis either vertical or horizontal. The target rectangle was always 5 × 8 pixels, whereas the nontarget rectangle was 4 × 10 pixels. With these stimuli, the proportion of CRs (.17) was much closer to the proportion of LFCRs (.14). Thus, it would appear that Donk's confusion hypothesis can account for the occurrence of CRs when the target and nontarget are confusable. However, in most experiments by other researchers, the nontargets are chosen to be distinct from the targets. Donk's confusion model cannot account for the results of these previous studies.

The method adopted by Donk (1999) in Experiments 1–4 turned out to be problematic. In these studies, there were only two values on each dimension, a paradigm that is unfortunate for three reasons. First, as we have mentioned, it precludes the simple analysis described in Figure 1. Second, the probability of misperceiving the target shape or color (feature errors) was estimated from trials on which illusory conjunctions could not occur. These were trials on which the critical feature was repeated (e.g., two red objects), referred to as *identical trials.* The assumption is that the feature errors estimated on the identical trials would provide an estimate of the feature error rate on trials in which the target and nontarget color were different. However, evidence exists to suggest that detecting stimuli with identical features might be different than detecting stimuli with heterogeneous displays (Bamber, 1969; Bamber, Herder, & Tidd, 1975). If Donk had included more than two features on each dimension, which was the approach taken in previous studies of illusory conjunctions, she would not have had to make the assumption that feature errors on identical and heterogeneous displays were the same. Third, by adding another color and shape to the stimulus set, the number of response categories would be increased from three to six, which makes it easier to model the data.

The problem of estimating parameters with the procedure in Experiments 1–4 is apparent when considering the estimated values for the parameters that relate to confusing the nontarget with the target, *si* and *sc*, respectively. These values vary widely across individuals. For example, in Experiment 1, they varied from 1.0 to .01. This range suggests that in some conditions, some of the observers never confused the nontarget with the target, whereas other observers essentially failed to distinguish between the target and nontarget on all trials, which seems unlikely. We suspect that the estimates of these parameters are unstable and reflect the fact that the models need to account for different types of errors. Such instability is not observed for parameters that represent the perception of the target features or the binding of these features, which are both in previous studies that have used multinomial models (Ashby et al., 1996; Prinzmetal et al., 1995) as well as in the studies of Donk (1999).

It is possible that in several of Donk's (1999) experiments, observers did not make illusory conjunctions. This observation, by itself, would not be surprising. The visual system is quite adept at integrating features, and our experience has shown us that it can be quite difficult to elicit illusory conjunctions. For example, illusory conjunctions are more likely (a) when the features are aligned horizontally or vertically rather than on an oblique line (Lasaga & Hecht, 1991; Prinzmetal, 1981), (b) when the items are horizontally aligned rather than vertically aligned (Prinzmetal & Keysar, 1989), and (c) when the major axis of symmetry of the display items is aligned (Gallant & Garner, 1988; Lasaga & Hecht, 1991). The conditions favoring the occurrence of illusory conjunctions were only present in Donk's replication of Ashby et al. (1996).

The picture of feature integration that has emerged over the last 20 years is that feature binding is constrained by a rich set of factors ensuring that our perception of objects is generally veridical. However, in normal environments, and even more so under carefully designed conditions, failures of perception can occur. These failures have proven to be an important source of information concerning the processes of feature perception and integration. Donk (1999) has correctly shown that confusions between the target and nontarget items can lead to responses that are erroneously attributed to illusory conjunctions. The much stronger claim, that the confusion hypothesis can account for all previous reports of illusory conjunctions, is not tenable. The phenomenon of illusory conjunctions has taught people much about how the human visual system binds features. Illusory conjunctions remain alive and well.

References

Arguin, M., Cavanagh, P., & Joanette, Y. (1994). Visual feature integration with an attentional deficit. *Brain and Cognition, 24,* 44–56.

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, T. (1996). A formal theory of illusory conjunctions. *Psychological Review, 103,* 165–192.

Bamber, D. (1969). Reaction times and error rates for "same"–"different" judgments of multidimensional stimuli. *Perception & Psychophysics,* 169–174.

Bamber, D., Herder, J., & Tidd, K. (1975). Reaction times in a task analogous to "same"–"different" judgment. *Perception & Psychophysics,* 321–327.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psycholonomic Bulletin & Review, 6,* 57–66.

Batchelder, W. H., Riefer, D. M., & Hu, X. (1994). Measuring memory factors in source monitoring: Reply to Kinchla. *Psychological Review, 101,* 172–176.

Baylis, G. C., Driver, J., & McLeod, P. (1992). Movement and proximity constrain miscombinations of colour and form. *Perception, 21,* 201–218.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Cohen, A., & Ivry, R. B. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 650–663.

Cohen, A., & Rafal, R. D. (1991). Attention and feature integration: Illusory conjunctions in a patient with a parietal lobe lesion. *Psychological Science, 2,* 106–110.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology* (Vol. 31, pp. 309–341). New York: Oxford University Press.

Donk, M. (1999). Illusory conjunctions are an illusion: The effects of target–nontarget similarity on conjunction and feature errors. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1207–1253.

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995, August). Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science, 269,* 853–855.

Gallant, J. L., & Garner, W. R. (1988). Some effects of distance and structure on conjunction errors. *Bulletin of the Psychonomic Society, 26,* 323–326.

Khurana, B. (1998). Visual structure and the integration of form and color. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1766–1785.

Kleiss, J. A., & Lane, D. M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 200–210.

Lasaga, M. I., & Hecht, H. (1991). Integration of local features as a function of global goodness. *Perception & Psychophysics, 49,* 201–211.

Logan, G. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103,* 603–649.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics, 30,* 330–340.

Prinzmetal, W. (1995). Visual feature integration in a world of objects. *Current Directions in Psychological Science, 4,* 90–94.

Prinzmetal, W., Henderson, D., & Ivry, R. (1995). Loosening the constraints on illusory conjunctions: The role of exposure duration and attention. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1362–1375.

Prinzmetal, W., Hoffman, H., & Vest, K. (1991). Automatic processes in word perception: An analysis from illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 902–923.

Prinzmetal, W., & Keysar, B. (1989). A functional theory of illusory conjunctions and neon colors. *Journal of Experimental Psychology: General, 118,* 165–190.

Prinzmetal, W., & Millis-Wright, M. (1984). Cognitive and linguistic factors affect visual feature integration. *Cognitive Psychology, 16,* 305–340.

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance, 12,* 361–369.

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language, 25,* 461–475.

Riefer, D. M., & Batchelder, D. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339.

Selfridge, O. G. (1959). *Pandemonium: A paradigm for learning.* London: Her Majesty's Stationery Office.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255,* 114B–125.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107–141.

Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review, 82,* 184–199.

Wolford, G., & Shum, K. H. (1980). Evidence for feature perturbations. *Perception & Psychophysics, 27,* 409–420.

Received January 18, 2000
Revision received July 6, 2000
Accepted August 15, 2000 ■