Dancing letters and ticks that buzz around aimlessly:  
On the origin of crowding†

Hans Strasburger
Department of Medical Psychology and Medical Sociology, University of Göttingen; and
Institute of Medical Psychology and Generation Research Program (GRP), University of Munich,
Elsässer Strasse 13, 81667 Munich, Germany; e-mail: strasburger@uni-muenchen.de
Received 17 February 2014, in revised form 18 June 2014, published online 4 September 2014

Abstract. When we see an object, we know where it is. Or do we? Perhaps not in indirect vision, as
was observed by the gestalt psychologist Korte in 1923. Objects and object parts appear to ‘dance
around’, and these phenomena may underlie a part of what is called the crowding effect today. From
Korte’s account of pattern recognition in indirect vision, I select two phenomena: a loss of the positional
code for letter parts and a loss of the same for whole letters. Using these examples, I present a novel,
speculative explanation for a contradiction in the literature: patterns located more peripherally
than a target show more interference than do more centrally located patterns, yet for whole-letter
confusions the asymmetry is the other way round. The inward, not the outward, flanker is increasingly
confused with a target at increasing target eccentricities. I propose that feature-binding decreases with
eccentricity such that free-floating letter parts more often intrude from the periphery and whole letters
from the centre. I conclude with a few remarks on computational modelling as, hopefully, a challenge
to neural computationalists.

Keywords: illusions, indirect vision, peripheral vision, pattern recognition, gestalt, eccentricity, visual
field, binding

1 Introduction
Visual crowding is the phenomenon that a pattern is less well recognized in the presence
of neighbouring patterns than when seen in isolation. (1) It is a rather strong effect in more
than 99.9% of the visual field, (2) and it happens every moment in everyday life (when the
eyes are open). But we hardly ever notice it. Visual spatial attention is mostly focused
on the fovea, so much so that eye movements are called overt attention, and we seem to
confuse the phenomenal characteristics of foveal visual perception with visual perception
in general. As another example, contours in peripheral vision appear sharp (3) (cf Galvin,
O’Shea, Squire, & Govan, 1997; Galvin, O’Shea, Squire, & Hailstone, 1999), even though
acuity, which would be thought to be at the basis of perceived sharpness, is drastically
reduced there.

† This paper is based on a talk at the Symposium ‘Illusions and delusions’ in The Barn, 2013. Data were
published in the Journal of Vision, 13(1):24, 1–20, http://www.journalofvision.org/content/13/1/24
(1) Reviews for the crowding effect are found in Strasburger, Harvey, and Rentschler (1991), Pelli,
Palomares, and Majaj (2004), Levi (2008), and Strasburger, Rentschler, and Jüttner (2011).
(2) In the central fovea the crowding effect is weak or absent, but a marked effect occurs already
at 2 deg eccentricity—that is, slightly off the very centre of the fovea (Strasburger et al, 1991).
The visual field subtends roughly 90 deg on the temporal, and 60 deg on the nasal, side. A central
circle of 2 deg radius corresponds to 0.06% of the total area—that is, 99.94% area is outside the central
2 deg radius.
(3) The perceptual appearance of sharpness, of something appearing in focus as separate of being in
focus, was first addressed by Jurin (1738) in the distinction of distinct versus perfect vision. See the
appendix for the terms.
Why speak of crowding in the context of visual illusions? A common-sense definition of illusion is that there is a mismatch between a particular description of the stimulus presented to the observer and his or her perception (but see Rogers, 2010, Koenderink, 2014), and thus by studying illusions, we can learn about how the system works. Indeed, this is how Korte (1923) arrived at a framework for a gestalt theory of pattern perception, as we will see below.

For a definition of what is an illusion, Wade (2005) distinguishes an early approach already used in ancient Greece from the more recent view in history. The earlier, atheoretical approach was to compare perceptual experience of the same objects at different times: if there are differences (as in the apparent size of the moon during its transit), then an illusion has occurred under the assumption that the object has not changed physically over that time (the assumption of object permanence). Later in history, with the better physical measurement of objects, a difference between physical and perceptual observations became the basis for identifying an illusion. For peripheral vision, both these criteria apply. The differences in visibility that are experienced in peripheral vision and that are discussed below do not occur with normal central vision. Also, the physical measurements remain the same but the experience differs in peripheral and central vision. So by these criteria, certain major aspects of peripheral vision—even very close to the point of fixation—are illusory.

Illusions violate our trust in the veridicality of perception (think of Plato); they surprise us, and the experience is somehow odd. Similarly, crowding—which is probably the significant feature of peripheral vision—seems curiously different when brought to attention and to be in conflict with our sense of veridicality. Its main characteristics are:

- crowding overrides all other effects in peripheral pattern recognition, yet goes unnoticed;
- crowding is unrelated to spatial resolution (which is widely considered as being the main shortcoming of peripheral vision);
- one of the ingredients of crowding (as measured in standard paradigms) is the phenomenon that a pattern is recognized but is mistaken for the pattern that was intended to be recognized (which seems to conflict with the very concept of recognition; I refer to that as loss of positional coding; Strasburger, 2005).

I am not the first to say that peripheral vision is curiously different: Aubert and Foerster (1857, page 30), in their classic treatise on indirect vision, said:

“When the two points [in indirect view] cease to be distinguished as two … they are not seen as a single point but quite peculiarly indeterminate (‘strangely undefined’) as something black, the form of which cannot be further stated.”

Interestingly, this comment of theirs—which tries to get at the character (the quality) of nonfoveal vision—was never taken up even though Aubert and Foerster’s quantitative results are well known today (cf the ‘Aubert–Foerster Law’). Half a decade later, Korte (1923) undertook the first thorough attempt to analyze the qualitative, phenomenological characteristics of peripheral vision, including what we now call crowding. He suggested a three-stage process of pattern recognition, which will be sketched below. Such a process theory can be seen as a first step towards a neurocomputational theory, as requested by Tyler Rogers (2010, and this special issue) has argued that there is no adequate or satisfactory way of distinguishing between those things that are labelled illusions and those that are considered to be veridical. The problem arises because there is no single description of what the ‘reality’ is with which our perceptions are compared—there are only descriptions of how our perceptual systems work.

(4) See Appendix for use of the term.

(5) “Wenn die zwei Punkte aufhören, als zwei unterschieden zu werden, also jenseits des Gränzpunktes liegen, so sieht man sie nicht als einen Punkt, sondern ganz eigentümlich unbestimmt als etwas Schwarzes, dessen Form weiter nicht anzugeben ist” [page 30, emphasis by Sperrszatl (increased letter spacing) changed to italics].
and Likova (2007); at least, a neurocomputational theory of peripheral pattern recognition should be capable of describing the phenomena that were described very systematically by Korte (1923). Note that we are still far from that point.

The common belief that reduced acuity is the primary characteristic of peripheral vision still pervades; its peculiarity has been recognized only by a few in the past. Lettvin (1976), in “On seeing sidelong”, says

“When I look at something it is as if a pointer extends from my eye to an object. … Things are less distinct as they lie farther from my gaze. It is not as if these things go out of focus—but rather it’s as if somehow they lose the quality of form” (page 10). It was only quite recently, around 1995, that the catch term crowding became popular, although the phenomenon was studied extensively before but using different terminology. Besides crowding, the phenomenon (or an attempted explanation) was variably called lateral inhibition, lateral interference, or a number of other terms. There seem to be two schools, one in vision research/ophthamology/optometry and one in experimental psychology/cognitive psychology/computer science, that tended (and still tend) not to cite each other (Strasburger et al., 1991, 2011, section 5). The term crowding and phenomenon as we understand it today were introduced in writing in 1953 by the Danish ophthalmologist Ehlers (he had mentioned the phenomenon earlier, in 1936, but only in passing). Stuart and Burian (1962, page 471) later called it the “crowding phenomenon”. Ehlers (1953) noted that, in the acuity charts then in use, some children and adults had particular difficulties recognizing the optotype when other optotypes were close by.

2 Korte’s (1923) gestaltist analysis of indirect vision

Let us return to Korte (1923) and his description of the perceptual phenomena of reading in indirect vision. In today’s view we could equally well see this as an early extensive (and quantitative) treatise on pattern recognition. His introduction reads rather contemporarily: “For a long time the prejudice was prevailing that indirect, compared to direct vision, is imperfect and irrelevant, and only very slowly the insight of the fundamental importance of seeing sidelong has prevailed” (page 18). The text continues by saying: “In 1889, Kirschmann has shown that, in reading, the individual letters are not fixated one after the other but that the fixation point jumps, which means that most letters are seen extrafoveally only” (page 18). While the last phrase is a little misleading (in normal reading, letters outside the fovea, and often even outside the foveola are seen but not recognized), the sentence points to the fact that pattern recognition capabilities are different just a little away from the point of fixation. This is also why Korte speaks of indirect rather than peripheral vision. From the present perspective, it should be added, it is not eccentricity per se that limits reading but the size of the uncrowded window (Pelli & Tillman, 2008).

(7) Keywords or labels that were used include crowding effect (instead of just crowding), contour interaction, interaction effects, lateral inhibition, lateral interference, lateral masking, masking, and surround suppression. There is disagreement as to whether or not these terms refer to the same phenomenon, and it is not always clear whether they refer to an explanation or to the phenomenon itself.

(8) “When one is testing amblyopic children with isolated letters or E’s, the visual acuity recorded is often much better than with the ordinary test chart. If the visual field is crowded with letters, the area of the visual field in which the letters can be recognized narrows. This is very easy to demonstrate, as I showed at the Congress of Scandinavian Ophthalmologists in 1936.”

(9) For a summary of Korte’s paper see the appendix in Strasburger et al. (2011).

(10) Kirschmann in 1889 was not the first to describe discontinuous eye movements during reading; Hering (1879) did that earlier (Wade, Tatler, & Heller, 2003).

(11) The fovea’s diameter is ~5.2 deg, and that of the foveola is ~1 deg (Wandell, 1995).
Later in his paper, Korte describes an extensive set of experiments (performed in Leipzig in 1920–22) with isolated letters (pages 20–29) and words of three to eight letters length, both meaningful and meaningless (pages 29–37). From these experiments Korte extracted letter and word characteristics that are “favourable for their recognition” (pages 44–51), and systematized sources of perceptual errors (pages 63–78).

These sections are preceded by a phenomenological description of the recognition process that Korte derived from the subjects’ reports (page 37–44). He distinguishes three phases: the “1st phase of the perceptual process [thus] brings about a notion of the most general characteristics of the sensation as a whole, i.e. for example roundness, angularity, obscurity, length etc” (page 38). Note that according to this account processing starts not from detail but from the whole; the emergence of detail occurs in only the second phase: “The 2nd phase sets in when, out of the change of sensations, something characteristic singles itself out, be it right or wrong. Now, winged by phantasy, the urge of figuration, of creating a Gestalt (Gestaltungsdrang) sets in [in the third phase], and creates, from the clearly perceived and the diffusely remaining, the image of a character” (page 39). The choice of the term Gestaltungsdrang implies that the process is driven towards a goal (Drang means urge), that the goal is a gestalt—that is, fulfils gestalt characteristics like closure—and that the outcome is not fixed: that there is not a uniquely defined perceot for that what is to be perceived (representandum; cf, for example, Held, 2006). In modern-day language we would conceptualize that phase as the time period in which the neural network for recognizing the pattern goes into an equilibrium state, which in turn corresponds to the resulting percept.

Later in the paper seven phenomena are listed that happen during that third phase of gestalt formation in the case of word recognition (which Korte called “Ursachen der Verlesungen”(12)—that is, “causes of/reasons for misreadings”) (see Strasburger et al., 2011, page 61, for an online translation of the list).(13) Six of these are perceptual, and one would be called cognitive today. Out of the six, I shall consider four in the present context.

- **Absorption** (‘Aufsaugung’) and **false amendment**: “A feature of a letter, or a whole letter, is added to another letter, or a detail becomes so dominant that it absorbs everything else” (page 64). Today, this would be referred to as the wrong allocation of features. It was formalized by Wolford (1975) in his feature perturbation theory and corresponds to what Pelli et al. (2004) call excessive feature integration. In his model Wolford (1975) uses the feature space provided in Lindsay and Norman’s (1972, page 120) famous textbook which we will come back to later. The paper was written while Wolford visited Rumelhart (and others) in San Diego. Rumelhart later became a spiritus rector of neural network theory together with McClelland (cf Rumelhart, McClelland, & the PDP research group, 1986).
- **Floating of ‘features’**: “It has already been mentioned that the perceptions are extra-ordinarily wavering. They do not keep still for their regard but are permanently moving. This goes as far as that subjects frequently speak of a ‘dance’. Particularly erratic are the horizontal strokes, the ticks, the arches etc. They aimlessly buzz around, so to say. One minute up, next minute down, then right, then left” (Korte, 1923, page 40).

---

(12) The distinction between phenomena, mechanisms, underlying causes, etc is a thorny issue when speaking about crowding. Korte speaks of ‘Ursachen’, a term which is rather basic to thinking and reasoning in the German language; the best-fitting translations for that would be ‘causes’, ‘reasons’, ‘sources’, or ‘origins’. The syllable ‘Ur-’ translates to Irish but not to English (https://en.wiktionary.org/wiki/ur). Another translation of Ursache is ‘originating source’ (http://www.marxists.org/archive/herder/1772/origins-language.htm). ‘Cause’ comes perhaps closest, but Ursache has fewer connotations to ‘causal’ than does ‘cause’.

(13) That phase is erroneously called second phase there.
• Floating of whole characters: “In the (perception of) words, to the fleetingness of the constituent elements the bouncing of whole characters is added. Firm localization of detail is extremely difficult; it is possible, at most, for the first and, less so, for the last letter. … With ‘kä’, subject R had the impression ‘These are two manikins dancing with each other’. For ‘vais’, the same subject described it as ‘two ‘o’ that jig about in the word’. Subject WD for ‘wauß’ said: ‘the whole word jumps to and fro’. Several times there was the report ‘I am in doubt whether there is a dot on the i. It’s there and then it isn’t.’ With ‘tot’, subject B who was particularly trained in indirect vision saw a ‘t’ and an ‘o’ but was unable to say whether the ‘o’ was on the right or on the left, or whether even half an ‘o’ was on either side of the ‘t’, i.e. ‘db’” (page 41). Elsewhere, we have called this process letter-source confusion (Strasburger & Malania, 2013), in parallel to the term source confusion in cognitive psychology (eg Huber, Shiffrin, Lyle, & Quach, 2002) which refers to the loss of information about where an individual feature came from. Nandy and Tjan (2007) provide a computational model for the latter. Empirical evidence for letter-source confusion will be presented below.

• Perceptual shortening in a certain zone (of the visual field): this point concerns the perceived word length, estimated by the number of letters. Words appear shorter than they are in a visual field region between the blind spot and the fovea. “This phenomenon is so peculiar that we need to attend to it more closely. As the outermost limits of this ‘zone of shortening’, 30 and 5 mm, respectively, can be marked down. Most pronounced is the shortening between 25 and 15–10 mm” (Korte, 1923, page 66). Korte’s experiments were done at 1 m viewing distance on the horizontal meridian, with stimuli slowly moving into the visual field, so this peculiar region extends from 0.29 deg to 1.7 deg, with a maximum effect at from 0.57–0.86 deg to 1.4 deg. These are rather small eccentricities and well within the fovea. The phenomenon is said to happen with only meaningless syllables. Typical statements are: “The word becomes shorter” and “now it becomes longer again” (page 67). Figure 1 summarizes more subject reports for that phenomenon.

**Figure 1.** Korte’s (1923) examples of what he called perceptual shortening happening in the fovea (page 67). The notation in the figure is shorthand for “the meaningless syllable sif was reported four times as ff, twice as ss, and once as sf, sl, and if, respectively (same for läunn, etc). The maximum effect is at the border of the foveola.

Pelli et al. (2004, page 1139) and, from that source, Levi (2008) cite this as the first description of crowding. They were led by a quote close to the end of that section (disregarding the context): “It is as if there is a pressure on both sides of the word that tends to compress it. Then the stronger, i.e. the more salient or dominant letters, are preserved and they quasi ‘squash’ the weaker, i.e. the less salient letters, between them” (Korte, 1923, page 69). Since these papers are concerned with crowding outside the fovea, this is the wrong phenomenon and the wrong quote.

---

(14) “Diese Erscheinung ist so eigentümlich, daß wir uns genauer mit ihr befassen müssen. Als alleräußerste Grenzen dieser ‘Verkürzungszone’ können wohl 30 und 5 mm bezeichnet werden. Am stärksten tritt die Verkürzung zwischen 25 und 15–10 mm auf.”
3 Feature confusions or letter confusions?
What does Korte’s (1923) paper tell us today? On the one hand, it is probably the most complete phenomenological description of letter and word pattern recognition in indirect vision ever written. The challenge for present-day theories of pattern recognition is to explain just a few of them, and keep in mind just how many there are. On the other hand, even if the classification into seven phenomena (or 'causes’) is just one way to organize the reported percepts, which may be couched in different terminology today, it suggests that there is not the one correct mechanism underlying crowding, but there are likely several, and they are likely to contribute to different extent depending on what is there to be perceived. One hint in that direction (discussed above) was that, close to the very centre of the fovea, there is a peculiar squashing of ‘weak’ letters (page 69) which does not happen farther out. In the remainder of the paper I would like to give an account of how two of Korte’s phenomena listed above—the floating of features and the floating of whole letters—appear to contribute differentially. This will be presented in the context of two sets of experiments that are published and online (Strasburger, 2005; Strasburger & Malania, 2013). The data and description of the two asymmetries are exposed there, but in the present paper I propose a novel explanation of how they could come about and which, at the same time, could resolve a seeming contradiction in the literature.

Subjects were presented with letters of various sizes on the right horizontal meridian, either singly or surrounded by a flanker (of same contrast and size) left and right (see inset in figure 2). Crowding was measured in the letter-contrast threshold paradigm (Strasburger et al., 1991): contrast thresholds for the recognition of the target were determined (in an adaptive procedure), and the increase in contrast threshold caused by adding the flankers is the measure of crowding. In a number of the tested conditions a ring cue around the target was added and used to attract transient attention (Nakayama & MacKeben, 1989), but this can be disregarded for the present discussion.

![Figure 2. False localizations: influence of flanker distance and target eccentricity (modified from Strasburger & Malania, 2013). The inset shows the basic stimulus.](image-url)
By the definition of threshold, a constant percentage of recognition errors occur in the paradigm, about 33% in our case. Now, in the crowding condition these errors can be classified into two groups: those in which the subject reports a flanker instead of the target (correspondences), and those where some other character is reported. The former I interpret as mislocalizations or letter-source confusion—that is, as caused from the loss of the character’s positional information. Of the second group of errors, a few contribute to crowding (Strasburger, 2005) but the majority just mean that the target is not recognized for some other reason (like contrast being too low or from mislocalized individual features). Mislocalizations of whole characters would be interpreted as arising from mislocalized features by Wolford’s (1975) perturbation model or Pelli’s feature-integration concept (Pelli et al., 2004; see also Freeman, Chakravarthi, & Pelli, 2012), but it remains the case that the circumstance where these are the features that make up a flanker and are bound in some way sets whole-letter mislocalizations apart from mere feature mislocalization. So, for simplicity I refer to the first group of errors as letter mislocalizations and to the second group as failure of feature detection or feature mislocalization.

Figure 2 shows these mislocalizations as a function of flanker distance, for three eccentricities (2 deg, 4 deg, and 6 deg on the right horizontal meridian; mean over twenty subjects). There are four things to observe:

- At a given eccentricity, mislocalizations decrease with flanker distance (which is to be expected since crowding happens only when flankers are within what is called critical distance).
- There is a maximum of mislocalizations (correspondences) when the flankers are close by, but that seems to occur not at the very lowest flanker distance but with flankers at a slight distance away.
- With increasing eccentricity, that maximum occurs at larger flanker distances.
- At the maximum, more than a third of all errors are in fact mislocalizations. In the present context we might call them illusions.

Note that, even though these errors are widely reported (see Strasburger & Malania, 2013, for review), current theories of crowding do not speak to them. As a separate source, they are either ignored or explained away as arising from feature mislocalization.

4 Asymmetry

Figure 2 shows the sum of confusions with both flankers, so it is natural to ask whether the two flankers contribute equally. Indeed, there have been numerous reports that this is not the case (reviewed in Strasburger & Malania, 2013). The general consensus is that a peripheral flanker exerts more influence than one centrally located (Mackworth, 1965). So, out of curiosity and following a suggestion by Solomon, we reanalyzed our data in the two studies with respect to that question. To our surprise, there was indeed a reliable asymmetry—but to our greater surprise, the asymmetry was the wrong way round. Figure 3 shows the data from the second study (those in the first are similar). Confusions increase with eccentricity for the inward flanker and remain at a constant, above-chance level for the outward flanker. As a result, the ratio of inward-to-outward flanker confusions increases systematically with eccentricity.
The seeming contradiction can be resolved by looking more carefully at what kind of flanker interactions there are. Among the papers on interaction asymmetry, there is indeed one that reports an opposite asymmetry, which has been largely neglected: Chastain (1982) found that, with respect to the effects of similarity and confusion, the inward flanker plays the more important role. He also pointed out that the confusability increases with eccentricity. Doubting his own results, Chastain also carried out a reanalysis of data in the well-known paper by Krumhansl (1977), and found that, counter to what was said in its text, it also supported the reverse asymmetry. Chastain (1982) further pointed out that his result of letter similarity being more influential for a centrally than for a peripherally located flanker was incompatible with the prevailing explanation which suggested that it results from feature movement from the peripheral to the central letter (since then similar flankers should exert less, not more, influence).

Now all the ingredients are coming together to draft a coherent model: flankers interfere with the recognition of the target, and the consensus is that the outward character is more disturbing than the inward, in that respect. At the same time, there are trials where the target is not reported simply because something else is mistaken for it, and in that respect the inward character is the more important. Both cases can be interpreted on the basis that the positional code (or the source of information) is lost in the processing. For the outward flanker, this shows up as individual features floating towards the centre and getting assimilated with the target (Wolford & Hollingsworth, 1974). For the inward flanker, it more often happens that the whole character moves outward—that is, we recognize a character, but do not know which one it is. Hence, the difference is that in the first case features move to some extent freely, and in the second case they are not free but are more closely bound together. An overarching explanatory bracket would be that, apart from a general loss of positional coding, feature binding decreases with eccentricity.

The suggestion is that there is:

- separate neural coding of pattern content and pattern position—that is, of what and where;
- an underlying mechanism to crowding where feature-binding decreases with eccentricity such that …;
- free-floating letter parts more often intrude from the periphery and whole letters from the centre.

---

**Figure 3.** Asymmetry of confusions (modified from Strasburger & Malania, 2013; correspondences in %, collapsed over all errors and all cue sizes). (a) Flankers separately; (b) inward–outward ratio.

---

**5 Solution: a two-mechanism model**

The seeming contradiction can be resolved by looking more carefully at what kind of flanker interactions there are. Among the papers on interaction asymmetry, there is indeed one that reports an opposite asymmetry, which has been largely neglected: Chastain (1982) found that, with respect to the effects of similarity and confusion, the inward flanker plays the more important role. He also pointed out that the confusability increases with eccentricity. Doubting his own results, Chastain also carried out a reanalysis of data in the well-known paper by Krumhansl (1977), and found that, counter to what was said in its text, it also supported the reverse asymmetry. Chastain (1982) further pointed out that his result of letter similarity being more influential for a centrally than for a peripherally located flanker was incompatible with the prevailing explanation which suggested that it results from feature movement from the peripheral to the central letter (since then similar flankers should exert less, not more, influence).

Now all the ingredients are coming together to draft a coherent model: flankers interfere with the recognition of the target, and the consensus is that the outward character is more disturbing than the inward, in that respect. At the same time, there are trials where the target is not reported simply because something else is mistaken for it, and in that respect the inward character is the more important. Both cases can be interpreted on the basis that the positional code (or the source of information) is lost in the processing. For the outward flanker, this shows up as individual features floating towards the centre and getting assimilated with the target (Wolford & Hollingsworth, 1974). For the inward flanker, it more often happens that the whole character moves outward—that is, we recognize a character, but do not know which one it is. Hence, the difference is that in the first case features move to some extent freely, and in the second case they are not free but are more closely bound together. An overarching explanatory bracket would be that, apart from a general loss of positional coding, feature binding decreases with eccentricity.

The suggestion is that there is:

- separate neural coding of pattern content and pattern position—that is, of what and where;
- an underlying mechanism to crowding where feature-binding decreases with eccentricity such that …;
- free-floating letter parts more often intrude from the periphery and whole letters from the centre.
Modelling crowding: Selfridge’s Pandemonium Model

For a satisfactory model of letter crowding, or its asymmetry, we would eventually expect the model to be neuroanalytic—that is, “one must be able to specify the processing circuitry in terms of known functions of cortical neurons, generating an output from a neural array that performs the task in a manner that can be implemented in a computational simulation” (Tyler & Likova, 2007, page 3). The concepts discussed above—floating of features, feature binding, and floating of characters—could then be ingredients that are made explicit in such a model or could emerge as its properties. There is as yet no such model of crowding (shortcomings of several model accounts have been discussed by Tyler & Likova, 2007; a recent review of crowding modelling can be found in Strasburger et al., 2011, section 8.4), but we might be close. In the remainder I would thus like to draw attention to a few points in this respect.

For one, the concept of feature integration is still a major contender for letter recognition and (thus) letter crowding (cf Freeman et al., 2012; Pelli et al., 2004; Song, Levi, & Pelli, 2014; Strasburger et al., 2011; Tyler & Likova, 2007). It is perhaps instructive to go back to the roots of that and remind ourselves that Selfridge’s (1959) Pandemonium Model (figure 4) was already computational and close to a neural implementation (since the demons

![Figure 4. Selfridge’s (1959) computational Pandemonium Model, as depicted by Lindsay and Norman (1972) (illustration: Leanne Hinton). Collectively, the demons recognize the letter ‘R’. Note that the feature space in Wolford’s (1975) Feature Perturbation Model was taken from Lindsay and Norman; there were seven types of features including vertical lines, acute angles, and continuous curves. Also note that feature locations are not coded in the model—that is, a floating of features is intrinsic to the model.](image-url)
might correspond to cell assemblies located, from today’s knowledge, in the ventral pathway up to the inferotemporal cortex or the fusiform gyrus. Which features are integrated (ie how, for example, Pelli et al.’s 2004 feature integration field is implemented) would need to be made explicit. But another of the requirements for our proposed explanation is already in place—namely, the floating of features: since feature locations are not coded in the model in the first place, floating of features is intrinsic to the model.

The other requirement is for some (explicit or implicit) mechanism of binding. Even if some of the features buzz around like little rascals for Korte’s subjects, features cannot just end up anywhere. They need to get somehow associated with a letter, and the degree to which that happens needs to be variable for the present purpose. In the new, high-dimensional pooling model from Rosenholtz’s group (Balas, Nakano, & Rosenholtz, 2009), the visual field is tiled with overlapping pooling regions (corresponding loosely to integration fields) whose sizes roughly correspond to a critical distance— that is, they are typically larger than a single letter and increase with eccentricity. Within these, the visual system computes a rich set of summary statistics. At a later stage, some ‘decision demon’ responds with what letter was the target. Feature-floating here translates to locational ambiguity. Binding and misbinding in this model emerge from the way the summary statistics within the pooling region are set up (Keshvari & Rosenholtz, 2013). The increased ambiguity from the, on average larger, pooling regions lying peripherally might bring about decreased binding, or perhaps one of the summary statistics could have a parameter in it that varies with eccentricity.

Note that, when we speak of feature binding here, it is meant in the narrow context of pattern recognition, not in the meaning of Treisman’s feature integration theory (Treisman & Gelade, 1980). The features in question are pattern elements and pattern characteristics (eg ‘roundedness’, and, moreover, are of same colour since crowding characteristics change when flankers have different colour or contrast polarity; Pelli et al., 2004). In this sense, I wish to keep “the binding baby” (Wolfe, 2012, page 307). Note further that, in pattern recognition by hierarchical re-entrant systems (eg Mumford, 1991, 1992)—which Di Lollo (2012) proposes as a replacement for solving the (allegedly ‘ill-posed) binding problem, binding is perhaps not explicitly modelled but still takes place—implicitly—namely, in the way the correlational process prefers certain combinations over others. Features are still an ingredient of such models, and a loss of binding towards the periphery, as proposed here, might then be implemented in some other way, perhaps by part-relational features (Strasburger et al., 2011, section 8.6).

In a recent paper, Pelli suggests that perhaps letter substitution—though prominent in the data—might after all not be required as an explanatory mechanism; the paper argues against “rampant unpooled substitution during crowding” (Freeman et al., 2012, page 388), allegedly suggested by our previous work (Strasburger, 2005) and that of others. The argument is based on the premise that a simple substitution process is blind towards the similarity of the target and the flanker. Note, however, that their stochastic model is not computational—that is, it does not describe the pattern recognition process, but rather places constraints on that process. Binding, or some equivalent process, is not addressed, so the question remains as to why certain combinations of features that constitute a flanker occur so often.

A final point to be made concerns the question of what are the features that are actually used by an observer. A computational approach of how these could be determined was presented by Nandy and Tjan (2007). Their approach was based on reverse correlation (ie averaging patterns depending on the observer’s answer), where the patterns for averaging were (within constraints) left free with respect to their location. As a result, feature extraction became

(17) That is, the distance within which crowding takes place, which is roughly half the eccentricity value; cf Bouma’s Law in Strasburger et al. (2011, section 5.3).
independent of the feature’s location—thus permitting the quantitatively separate processing of pattern content from pattern location. Nandy and Tjan’s approach therefore covers source confusion, for both features and whole letters. In sum, whether and how Korte’s phenomena and concepts and our extensions to the latter are implemented in a neuroanalytic model is an open question—it should be a challenge to neural computationalists.

Acknowledgments. Thanks go to Nick Wade, Brian Rogers, Ruth Rosenholtz, and Ingo Rentschler for thoughtful comments on the manuscript; to Brian Rogers for meticulous language corrections; and to Nick Wade for pointing out the historical dimension and suggesting the writings of Purkinje and Jurin. Thanks go further to the Deutsche Forschungsgemeinschaft (grant STR 354/7-1 to HS) which provided the funds for open access publication.

References
Aubert, H. R., & Foerster, C. F. R. (1857). Beiträge zur Kenntniss des indirecten Sehens. (I). Untersuchungen über den Raumsinn der Retina. Archiv für Ophthalmologie, 3, 1–37.
Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. Journal of Vision, 9(12):13, 1–18.
Bex, P. J., Dakin, S. C., & Simmers, A. J. (2003). The shape and size of crowding for moving targets. Vision Research, 43, 2895–2904.
Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 13, 767–782.
Carrasco, M. (2011). Visual attention: The past 25 years. Vision Research, 51, 1484–1525.
Chastain, G. (1982). Confusability and interference between members of parafoveal letter pairs. Perception & Psychophysics, 32, 576–580.
Chastain, G. (1983). Task and contrast effects on performance with parafoveal stimulus pairs. Psychological Research, 45, 147–156.
Clemmesen, V. (1944). Central and indirect vision of the light-adapted eye. Acta Ophthalmologica, 22, 317–318.
Collier, R. M. (1931). An experimental study of form perception in indirect vision. Journal of Comparative Psychology, 11, 281–290.
Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. Trends in Cognitive Sciences, 16, 317–321.
Ehlers, H. E. (1953). Clinical testing of visual acuity. Archives of Ophthalmology, 49, 431–434.
Estes, W. K., Allmeyer, D. H., & Reder, S. M. (1976). Serial position functions for letter identification at brief and extended exposure durations. Perception & Psychophysics, 19, 1–15.
Estes, W. K., & Wolford, G. L. (1971). Effects of spaces on report from tachistoscopically presented letter strings. Psychonomic Science, 25, 77–80.
Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. Attention, Perception, & Psychophysics, 74, 379–396.
Galvin, S. J., O’Shea, R. P., Squire, A. M., & Govan, D. G. (1997). Sharpness overconstancy in peripheral vision. Vision Research, 37, 2035–2039.
Galvin, S. J., O’Shea, R. P., Squire, A. M., & Hailstone, D. S. (1999). Sharpness overconstancy: The roles of visibility and current context. Vision Research, 39, 2649–2657.
Geissler, L. R. (1926). Form perception in indirect vision. Psychological Bulletin, 23, 135–136 (Abstract).
Held, C. (2006). Mental models as objectual representations. In C. Held, G. Vosgerau, & M. Knauff (Eds.), Mental models & the mind (pp. 237–253). Amsterdam: Elsevier.
Hering, E. (1879). Über Muskelgeräusche des Auges. Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-naturwissenschaftliche Klasse. Abt. III, 79, 137–154.
Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Quach, R. (2002). Mechanisms of source confusion and discounting in short-term priming: 2. Effects of prime similarity and target duration. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28, 1120–1136.
Hueck, A. (1840). Von den Gränzen des Sehvermögens [On the limits of vision]. Archiv für Anatomie, Physiologie und wissenschaftliche Medicin, 82–97.
Jonides, J. (1981). Voluntary versus automatic control of over the mind’s eye’s movement. In J. B. Long, & A. D. Baddeley (Eds.), Attention and performance (Vol. 9, pp. 187–204). Hillsdale, NJ: Erlbaum.
Jurin, J. (1738). An essay on distinct and indistinct vision. In R. Smith (Ed.), *A compleat system of opticks in four books, viz. a popular, a mathematical, a mechanical, and a philosophical treatise* (pp. 115-171). Cambridge: Published by the author.

Keshvari, S., & Rosenholtz, R. (2013). A high-dimensional pooling model accounts for seemingly conflicting substitution effects in crowding. *Journal of Vision, 13*, 576.

Kirschmann, A. (1889). Über die Helligkeitsempfindung im indirecten Sehen. *Philosophische Studien (Wundt), 5*, 447–497.

Kirschmann, A. (1908). Über die Erkennbarkeit geometrischer Figuren und Schriftzeichen im indirecten Sehen. *Archiv für die gesamte Psychologie, 13*, 352–388.

Koenderink, J. (2014). The All Seeing Eye? *Perception, 43*, 1–6.

Köllner, H. (1920). XII. Die Untersuchung des indirekten Sehens. In E. Engelking, H. Erggelet, H. Köllner, F. Langenhan, J. W. Nordenson, & A. Vogt (Eds.), *Die Untersuchungsmethoden. Handbuch der gesamten Augenheilkunde* (pp. 394–552). Berlin: Julius Springer Verlag.

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie, 93*, 17–82.

Krumhansl, C. L. (1977). Naming and locating simultaneously and sequentially presented letters. *Perception & Psychophysics, 22*, 293–302.

Lettvin J. Y. (1976). On seeing sidelong. *The Sciences, 16*, 10–20.

Levi, D. M. (2008). Crowding—an essential bottleneck for object recognition: A minireview. *Vision Research, 48*, 635–654.

Lindsay, P. H., & Norman, D. A. (1972). *Human information processing: An introduction to psychology*. New York: Academic Press.

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science, 3*, 67–68.

Mumford, D. (1991). On the computational architecture of the neocortex I. The role of the thalamo-cortical loop. *Biological Cybernetics, 65*, 135–145.

Mumford, D. (1992). On the computational architecture of the neocortex II. The role of cortico-cortical loops. *Biological Cybernetics, 66*, 241–251.

Nakayama, K., & MacKeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*, 1631–1647.

Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision, 7*(2):5, 1–26.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*, 1129–1135 (plus online supplement).

Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: Not to be confused. *Journal of Vision, 7*(2):12, 1–9.

Purkinje, J. E. (1823). *De Examine Physiologico Organi visus et Systematis Cutanei*. Breslau.

Purkinje, J. E. (1825). *Neue Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Berlin: G. Reimer.

Rogers, B. (2010). Stimuli, information, and the concept of illusion. *Perception, 39*, 285–288.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1986). *Parallel distributed processing. Vol. 1: Foundations*. Cambridge, MA: MIT Press.

Selfridge, O. (1959). Pandemonium: A paradigm for learning. Paper presented at Proceedings of the Symposium on Mechanisation of Thought Processes, National Physical Laboratory, Teddington, November 1958 (Vol. 1, pp 513–526). London: HMSO.

Song, S., Levi, D. M., & Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision, 14*(5):3, 1–37.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11), 1024–1037.

Strasburger, H., Harvey, L. O. J., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49*, 495–508.

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision, 13*(1):24, 1–20.
On the origin of crowding

Strasburger, H., & Pöppel, E. (2002). Visual field. In G. S. Adelman, B.H. Smith (Eds.), Encyclopedia of neuroscience (3rd ed., pp 2127–2129). Amsterdam: Elsevier.

“Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. Journal of Vision, 11(5):13, 1–82.

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. American Journal of Ophthalmology, 53, 471–477.

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

Tyler, C. W., & Likova, L. T. (2007). Crowding: A neuroanalytic approach. Journal of Vision, 7(2):16, 1–9.

von Helmholtz, H. (1871). Über die Zeit, welche nöthig ist, damit ein Gesichtseindruck zum Bewusstsein kommt [On the time required for a visual impression to reach consciousness]. Monatsberichte der Königlichen Preußischen Akademie der Wissenschaften zu Berlin, Juni, 333–337.

Wade, N. J. (2005). Perception and illusion: Historical perspectives. New York: Springer.

Wade, N. J. (2007). Image, eye, and retina (invited review). Journal of the Optical Society of America A, 24, 1229–1249.

Wade, N. J., Tatler, B. W., & Heller, D. (2003). Dodge-ing the issue: Dodge, Javal, Hering, and the measurement of saccades in eye-movement research. Perception, 32, 793–804.

Wagner, J. (1918). Experimentelle Beiträge zur Psychologie des Lesens. Zeitschrift für Psychologie, 80, 1–75.

Wandell, B. A. (1995). Foundations of vision. Sunderland, MA: Sinauer.

Wertheim, T. (1894). Über die indirekte Sehschärfe [On indirect visual acuity]. Zeitschrift für Psychologie & Physiologie der Sinnesorgane, 7, 172–187.

Wolf, J. M. (2012). The binding problem lives on: Comment on Di Lollo. Trends in Cognitive Sciences, 16, 307–308.

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

Appendix: Terminology

Indirect vision: We use the term indirect vision rather than peripheral vision here to denote vision off the point of fixation, and with sustained attention directed towards the visual target. The term was popular in this meaning in the 19th and first half of the 20th century. Purkinje (1823) first used it in formal writing (cf Wade, 2005, page 176; 2007). He elaborated on it in chapter 1 of his second book on subjective visual phenomena (“Indirectes Sehen”, Purkinje, 1825) where he writes:

“All other lines drawn from the objects to the eye fall on it more or less aslant, and determine the direction of the seen objects, besides and off the eye’s axial point where clear, direct vision takes place. Vision that is mediated by these lines may be called indirect, or lateral sight (‘Nebensicht’)” (page 4).

Today’s definition in Merriam-Webster Online is close to that phrasing: “Indirect vision: vision resulting from rays of light falling upon peripheral parts of the retina.” Aubert and Foerster (1857) use the term in the title of their paper, as do Kirschmann (1889, 1908), Wertheim (1894), and quite a few others (eg Clemmesen, 1944; Collier, 1931; Geissler, 1926; Köllner, 1920). The role of spatial attention is often implicit in these writings, in which the observers are typically instructed to fixate at a certain point, yet to process a target at some other, previously known location. Hueck (1840), who uses the Latin terms ‘visio obliqua’ versus ‘visio directa’, addresses the role of attention directly:

“it remains rather difficult, however, even when one keeps attention directed at such a sideways located object, to recognize it, in particular since the object’s colour becomes, the farther from the centre, the more ambiguous” (page 94).
The kind of attention that is meant in these historical writings would, in modern terminology, probably be referred to as spatial, sustained (or voluntary) attention. Note that the attentional status is often not clear in studies of noncentral vision. Yet, spatial attention was already described by von Helmholtz (1871); it refers to attention being directed towards a nonf ixated region of the visual field. The distinction of sustained, as opposed to transient, spatial attention was coined by Nakayama and MacKeben (1989) who pointed out differences in time constants between slow, consciously controlled ‘sustained’, and fast, reflex-like ‘transient’ attention. In many ways, that distinction is synonymous with that of voluntary versus automatic attention as distinguished by Jonides (1981) (see Strasburger et al., 2011, for these terms). Visual performance off the point of fixation depends importantly on where and how spatial attention is focused (see Carrasco, 2011, for review).

The role of attention when speaking of indirect vision is also implicit in the way indirect is used in everyday language: one can look at, or view, something directly or indirectly, but in any case one looks at it—that is, turns one’s focus of (sustained spatial) attention towards it. With respect to retinal location, the terms indirect vision and peripheral vision are often used synonymously. The emphasis is a little different, though. The term indirect vision takes the fovea as the reference, whereas peripheral vision reminds of the peripheral visual field in perimetry and the visual peripheral border, both of which are quite somewhere else (Strasburger & Pöppel, 2002; cf also the comments on terminology in Strasburger et al., 2011, page 3). The important characteristics of ‘peripheral’ pattern recognition discussed here begin to show already in the fovea, as mentioned in the main text. The term indirect vision avoids that confusion.

Distinct vision: A blurred image does not necessarily appear blurred, as already noted by Jurin in the 18th century. Jurin (1738) elaborates in his “An essay on distinct and indistinct vision” on the observation that, with respect to clearness versus blurriness, the visual percept is to some extent independent of the retinal image being in focus. He thus distinguishes “Vision perfectly distinct or perfect vision”—where the retinal image is in focus—from distinct vision, where the retinal image is out of focus but large letters still appear nonblurred. He defines the latter, distinct vision, as follows:

“11. Vision imperfectly distinct, or simply Distinct Vision, is that, in which the rays of each pencil are not collected into a sensible point, but occupy some larger space upon the Retina, yet so as that the object is distinctly perceived, as the larger point in Art. 5 and 6” (page 116).

The paragraph that he refers to describes his observation and experimental method by which distinct vision is separated out—namely, by bringing printed text closer to the eye than accommodation (which is discussed in the same essay) would allow:

“5. Afterwards, bring the book by degrees so near, as that the letters of the smallest print may now begin to appear a little confused, and cannot by any endeavour or straining of the eyes be rendered so distinct as they were before. Then, keeping the book at the same distance, look at a print somewhat larger than the former, and that larger print shall seem perfectly distinct without any the least appearance of confusion.”

“Here, it is manifest from the less distinct appearance of the smaller print, that at this distance the rays of each pencil are not accurately united in a sensible point of the Retina, notwithstanding which the larger print appears distinct” (page 116).

Jurin’s essay refers to foveal vision, but today’s research on peripheral vision (eg work by O’Shea cited above) shows that the phenomenon is much more general.