
Mathematizing Human Perception



Brett Jefferson

In this exposition, I aim to encourage mathematicians to learn about and conduct research in the area of cognitive science by walking the reader through the process of utilizing tools and formalism from mathematics to address challenges in perception. First, I introduce the audience to a historically considered phenomena from psychology, called blob processing. Then, I present one approach to mathematically model the phenomena. Throughout, I introduce psychological findings to show how we can incorporate observed experimental results into our mathematical model.

A Brief Foreword on the Importance of Mathematical Psychology

I'd like to pick on the bully in the room a bit. In the last decade, the area of data science has grown massively due to advances in deep learning. It touches almost every

scientific discipline and algorithms are being developed that are undeniably accurate in their predictive power and robustness. However, despite the popularity of machine learning approaches to applied problems, there are many observable phenomena that require more finesse and rigor for an appropriate explanation than deep learning can provide. Psychology is full of such examples. It is amazing how humans are able to pick out a loved one's voice in a noisy crowd, or quickly identify a word before the actual letters are discernible. Likewise, mathematics is much broader than machine learning and provides a way of thinking and working through real-world problems. To describe and model human behaviors, psychologists and mathematicians (not always distinct people) have a long history of working together to *axiomatize* underlying principles, uncover the right decomposition of perceptual parts, and prove theorems about how and why the human experience doesn't simply cease to function! While this exposition is not a review of this history, a wonderful review of how mathematics and psychology have played in the same sandbox can be found in another AMS column article by Joseph Malkevitch [Mal15].

Brett Jefferson is a senior data scientist at Pacific Northwest National Laboratory. His email address is brett.jefferson@pnnl.gov.

Communicated by Notices Associate Editor Emilie Purvine.

*For permission to reprint this article, please contact:
reprint-permission@ams.org.*

DOI: <https://doi.org/10.1090/noti2734>

I encourage this generation of mathematicians to continue this history of marrying psychological principles and mathematical language, as it poses rich opportunities for mathematicians to find exciting new problems and research avenues. Here, I showcase one example of how a psychological observation can benefit from mathematical concepts and the rigorous formalism of mathematics. That methodology can, in turn, formalize the space of open questions around that observation. This article serves to provide just one example of what I call “mathematizing” a psychological observation.

The Blob Phenomena

We first think about a psychology problem (usually presented as an experiment with a puzzling result). Consider three rectangles of varying height and width, $\{(h_1, w_1), (h_2, w_2), (h_3, w_3)\}$. Over multiple rounds, you present a single rectangle very quickly (just a couple of tenths of a second) and task an observer (a priori) with identifying either height only or width only by responding with the appropriate value. First, you use the three rectangles from Figure 1a and, later, you rerun the experiment with rectangles from Figure 1b. The data collected are accuracy and response time. Once the experiments are complete, you compare performance across the two stimulus sets.

One will find that on average, judgements from stimulus set 1 were less accurate and slower than judgements on stimulus set 2. This is a bit of a mystery. Both sets of stimuli have two *dimensions*, height and width, that can be used to discriminate the rectangles. One might even posit that since the dimensions vary in a redundant and monotonic way in stimulus set 1, that we should expect the opposite behavioral result. Understanding how parts come together to form a perceptual experience is a major theme in perception research [TG11, Tre86, PP11]. [Loc72] proposes a descriptive model for why this occurs for the rectangles. The *blob* processing model posits that visual stimuli are initially perceived in a holistic way with the sum being greater than the parts. Higher-order features emerge that are perceived faster and are also easier to discriminate than the part-wise (single-dimension) components. As such, Lockhead suggests that the rectangles in the second stimulus set have more shape features than the squares in the first set. The notion of blob processing has been a persistent model throughout the decades following Lockhead’s note. Namely, in the area of letter and word recognition, there’s evidence that (1) words are identified before individual letters, (2) word shape and orthography are important for word recognition, and (3) neighboring words also serve as cues for word recognition. The type of holistic processing evidenced in this research area points to words (respectively sequences of words) being perceived

as a single object rather than a collection of letters (respectively words) with general shape features playing an important role. In fact, the word “blob” appears frequently in the literature when investigating early visual perception and speeded-judgement tasks and the idea of a “fuzzy” to “sharp” perception is an intuitive notion.

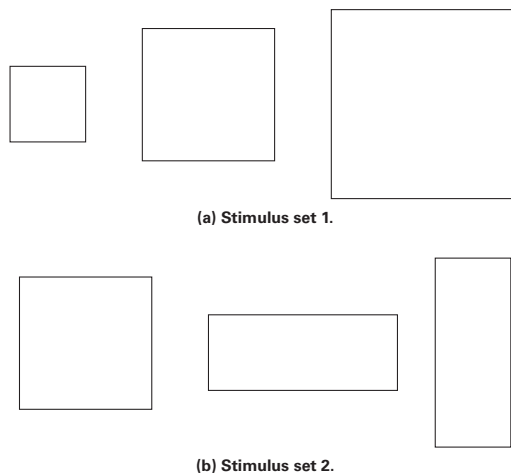


Figure 1. Two sets of stimuli for identification task.

This phenomena is not restricted to rectangular stimuli. In the 1980s Townsend published two studies in which he created a stimulus set comprised of the factorial combinations of three line parts and a curve part (Figure 2). Participants were asked to identify which components were shown. Townsend showed that the components/ parts were not perceived independent of one another. Findings included that the probability of a correct identification increased when there were more components present, the angle between line parts increased, and the stimuli was presented for longer durations [THE8408, THK8807]. These studies also highlighted the importance of time by modeling this discrimination of parts of a single stimulus as a process whereby over the time course of stimulus onset, exposure, and stimulus offset, more and more information is accumulated about the stimulus.

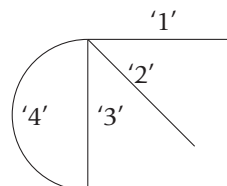


Figure 2. Townsend 1984 stimulus components.

At this point, the reader may have started developing a few theories on why and *how* such processing occurs. The reader may be thinking through the technical process of how our eyes transform light into neurological signals that

our brain then sends down various pathways that lead to perceived visual experiences. However, that level of end-to-end explanation does not quite prove satisfying from the point of view of providing an understanding of mapping direct stimulus inputs to perception ... not to mention the neuroscience community still doesn't have the tools for a complete classification in this way. Rather, we desire a mathematical expression or theory to properly understand how the measurable visual components come together to form this blob and, for that matter, a rigorous definition of what a blob is.

In the next sections, we'll explore a straightforward definition for a blob that is intuitive and testable. After, we'll see how different stimuli can be used to test the subsequent hypotheses. Finally, an analysis method is presented that relates to blob processing theory.

Mathematical Formulations

Mathematizing a repeatable human behavioral phenomena is both beautiful and terrifying at once. While it is satisfying to develop a rigorous framework for study, it can be difficult to know if the mathematics adequately (or appropriately) captures the observation. In mathematics, one tends to constrain the objects and rules for objects to well-studied and agreed upon structures like topological spaces, continuous maps, or algebraic structures. By contrast, in mathematical psychology, the phenomena often are so strange that it can be difficult to relate them to each other via existing structures. It's not uncommon for new areas of study to be derived from observed human behaviors. Fortunately, for the aforementioned phenomena, there are some ready structures that are proposed to define blob processing.

Any definition of blob processing should be compatible with observable processing properties. Those properties are:

1. **Refinement.** When a stimulus is presented, as time increases, the blob for that stimulus should decompose into a fixed set of constituent **parts** for sufficient viewing conditions.
2. **Complexity.** Assuming that for a given stimulus there is a fixed number of n constituent parts, as stimuli become more complex (containing more parts) they become more discriminable¹ from **stimuli** or stimuli comprised of a subset of constituent parts.
3. **Orthogonality.** Given a set of constituent parts, we assume that when two parts are not similar then the parts are discernible and discriminable earlier in time.

¹*Discrimination of two stimuli is measured via accuracy in an identification task (i.e., naming the distinct stimuli when they appear). We note that accuracy is experimentally measured and is usually the mean performance of several trials where performance is modulated by variance in motor movements, environment settings, attention, display properties, cognitive processing stochasticity, stimulus names, and a host of other factors that are well studied in psychology.*

Remark. This list of properties is not intended as an exhaustive list. In fact, property 2 is arguably not necessary as there are many examples where adding more features to a stimulus may serve to slow down discrimination rather than facilitating it. For this exposition, the properties primarily support the findings of [THE8408, THK8807].

These processing properties directly describe the observed phenomena from Townsend's study. Lockhead's review specifies that when there are multiple dimensions to a stimulus (like length and width) then the additional dimensions contribute to the perception and discrimination of the stimuli. This is akin to modeling stimuli in a multi-dimensional metric space rather than projecting to a uni-dimensional space. Note that there are two ends of the discussion here. The first end is at the early stage processing where stimuli can be thought of as barely detectable and coarse. Here the concept of dimensions isn't relevant because not enough information has been processed to even guess at the dimensionality of the stimulus. On the other end of the discussion, at a later stage of processing, stimuli are modeled together in a higher-dimensional space (often the experimenter defined manipulations each being a dimension) and as being comprised of perceivable parts. Each part can be modeled in the higher-dimensional space. For example, the rectangles in Figure 1 can be embedded as points in the 2-dimensional space with width corresponding to x-coordinates and length to y-coordinates. Simultaneously, the lower line segment of each rectangle can also be embedded in this space where the y-coordinate is 0. More commonly, a set of *parts* is defined first, and stimuli are created from combinations of parts. For this discussion, we take this later view and treat the transition from blob to not blob as a question of when the parts become discernible.

Definition 1 (Part). Let \mathcal{T} be a task for participants to complete, and \mathcal{S} be a fixed set of visually presented stimuli for that task. Let $\{p_i\}_N$ be a set of subsets of \mathbb{R}^2 along with an inner product such that for each $s \in \mathcal{S}$, $s = \{p_{j_i}\}_{1..N_s}$. We will call the p_j the **parts** of a stimulus, s . It is also useful to call the inner product the **similarity** between parts.

[Jef18] followed on [THE8408] and proposed that blob processing for a stimulus is a time sensitive function on the Fourier decomposition of visual stimuli. Look ahead to Figure 3 for example stimuli and their Fourier representations in polar coordinates. Let $F \times O$ be the Cartesian product of spatial frequencies (measured in the number of cycles subtended at the eye per degree) and orientations (measured in degrees of rotation from horizontal). For each frequency-orientation combination, (f, o) , of the decomposition, there is an amplitude determined by the stimulus. When the amplitude at (f, o) is greater than a threshold specific to that (f, o) a human can perceive

that component. We refer to the threshold as the **sensitivity** of the human to a frequency-orientation component (higher sensitivity implying lower threshold). Sensitivity (as a broadly used term in psychology) can refer to neuron responses, just-noticeable-differences, minimum settings before accurate detection occurs, minimum settings before accurate discrimination can occur, or, yet still, minimum setting before accurate identification can occur. The basic setup in the detection, discrimination, or identification tasks is that an experimenter will have a participant view stimuli through an apparatus (such as a digital monitor) for variable duration and respond (commonly through a button press) as quickly as possible. In a detection task, the participant will be given a cue to respond if a stimulus appeared or not. In the discrimination task, the participant may be asked to distinguish two or more stimuli ('same'/'different' judgements for example). In the identification task, the participant will be asked to respond with the unique identifier (maybe a name or a unique button) for each distinct stimulus presented. Conditions for presenting stimuli are manipulated to induce errors so that theorized deficits in stimulus processing can be studied. Manipulations can include duration of presentation, contrast of the stimulus with background, and size. Furthermore, sensitivity is not static, but is a function of time, contrast, and previously perceived points (f, o) in the decomposition.

Conjecture 1. For a task, \mathcal{T} , and set of stimuli, \mathcal{S} , let time from stimulus onset be denoted, $t \in T$. **Blob processing** is a real-valued function of a Fourier decomposition for stimulus, s ,

$$\mathcal{B}_{t,s} : F \times O \rightarrow \mathbb{R}$$

that predicts sensitivity thresholds for each frequency-orientation pair. We denote the amplitudes of a stimulus decomposition as $\mathcal{A}(s) : F \times O \rightarrow \mathbb{R}$.

Under this framework, a stimulus is not discernible at time t if $\mathcal{A}(s) < \mathcal{B}_{t,s}$ for all $(f, o) \in F \times O$. In other words, the human is not sensitive enough to the stimulus to finely perceive it but it can be detected. Time is a critical consideration here. In this model, the time scale for becoming more sensitive is typically on the order of milliseconds and is directly related to stimulus duration. For extremely short periods of duration, sensitivity will never reach a level sufficient enough for discrimination. For longer durations, there is a chance the sensitivity thresholds will be reached, but there are other considerations experimentally that are typically controlled for (dark adaptation, practice, stimulus size, etc.). Assuming that the participant is familiar with the stimuli, that the size is sufficient to avoid eye movements while being large enough to recognize parts, and that lighting is sufficient, sensitivity should reach threshold for **enough** frequency-orientation points that the stimulus becomes perceptible in a matter

of a couple hundred milliseconds. At this point a *blob* can now be defined.

Definition 2 (Blob). For a task, \mathcal{T} , and set of stimuli, \mathcal{S} , a stimulus s is perceived as a **blob** at time t if for some part of the stimulus, p_i , $\mathcal{A}(s)$ restricted to the frequency and orientation components of p_i is less than $\mathcal{B}_{t,s}(p_i)$ for a sufficient amount of frequencies and orientations.

Remark. The failure of a few components of $\mathcal{A}(s)$ to reach threshold doesn't prevent accurate perception of stimulus parts, much like using fewer terms in the Taylor expansion of a function doesn't prevent understanding of the overall function shape. In fact, there is a large body of research surrounding critical Fourier components for foveal (central vision) and peripheral stimulus identification.

Each of the desired processing properties can now be stated in terms of \mathcal{B} and \mathcal{A} .

1. **Refinement.** The refinement property requires that for $s \in \mathcal{S}$, there exists $\hat{t}_s \in T$ so that whenever $t > \hat{t}_s$, $\mathcal{B}_{t,s} \leq \mathcal{A}(s)$ for sufficiently many $(f, o) \in F \times O$.
2. **Complexity.** There exists a minimum $\delta_d \in \mathbb{R}$ so that for $s_i, s_j \in \mathcal{S}$ with $s_i \subset s_j$, there exists $t_{d,ij} \in T$ so that whenever $t \geq t_{d,ij}$ we have $\|\mathcal{B}_{t,s_i} - \mathcal{B}_{t,s_j}\| \geq \delta_d$ and humans can accurately distinguish s_i from s_j . We refer to δ_d as the discrimination threshold and $t_{d,ij}$ as the discrimination time. Note that δ_d is not stimulus dependent and is minimum over all stimulus comparisons. Also, $t_{d,ij}$ is ideally small enough that parts are not discernible. Complexity is how blobs are distinguished from one another.
3. **Orthogonality.** Given four parts, p_1, p_2, p_3 , and p_4 , where
 - $t_{d,12}$ is the discrimination time for $\{p_1\}$ and $\{p_2\}$,
 - $t_{d,34}$ is the discrimination time for $\{p_3\}$ and $\{p_4\}$,
 - and
 - $\langle p_1, p_2 \rangle \leq \langle p_3, p_4 \rangle$,
 we have $t_{d,12} \leq t_{d,34}$.

With a working definition of a blob, blob processing, and mathematical expressions for the three processing properties, we make note of a few directions in which one could go. First, one could start to study the nature of the sensitivity function, \mathcal{B} . This includes understanding if frequency and orientation have independent effects on sensitivity, asking if sensitivity is monotonic with time, or how sensitivity changes with different and related stimulus sets. In the next sections we explore some of these questions.

Another direction that one could pursue is characterizing the similarity inner product. This is one of the more difficult directions to validate in psychology because similarity judgements and responses can be modulated by experimenter instructions, context, and experiment environment settings . . . all things that **ideally** would have no

bearing on a true similarity representation. We often desire a static similarity space to make understanding other phenomena more feasible. There are models in psychology dedicated to unveiling such a representation, but this is outside of the scope of this exposition. For our purposes, we assume there is such an inner product and move on.

Testing Framework

Testing the blob model assumption that the Fourier decomposition is the *right* way to describe the stimulus space faces many challenges. This assertion is one based in human performance and humans, after all, vary in behavior, learn, adapt, and evolve. In truth many psychology models are, at best, contextually true; that is, they are specific to a stimulus set, population demographic, experiment settings, and measure of behavior. Strong theories show consistent results when these conditions change. A theory of visual perception based on spatial frequency channels (or dedicated processors) has been explored for an array of simple and complex stimuli. While alternative theories exist, as in [Lup79], there still remains strong physiological and psychophysical evidence that humans are wired to detect spatial frequency and orientation information [BC69, DeV82, VSG89]. The blob processing model brings to the forefront the question of time-course of processing. Namely, for early stage processing, we can ask three questions:

1. Do distinct, yet co-occurring frequencies facilitate, inhibit, or act independently on relative sensitivity? Using our definitions, for fixed $o \in O$, sufficiently low $t \in T$, and $f_1 \neq f_2 \in F$, what is the relationship between $\mathcal{B}_{t,s}(\{(f_1, o), (f_2, o)\})$, $\mathcal{B}_{t,s}(\{(f_1, o)\})$, and $\mathcal{B}_{t,s}(\{(f_2, o)\})$?
2. Likewise, do distinct, yet co-occurring orientations facilitate, inhibit, or act independently on relative sensitivity?
3. Lastly, is sensitivity to frequency independent of sensitivity to orientation? We desire the relationship between two frequencies (resp. orientations) to be independent of orientation (resp. frequency). Equivalently, are the following true:

- (a) Given frequencies f_1 and f_2 , there exists a constant, c_o so that for any orientation, \hat{o} ,

$$\mathcal{B}_{t,s_{f_1}}(\{(f_1, \hat{o})\}) = c_o \cdot \mathcal{B}_{t,s_{f_2}}(\{(f_2, \hat{o})\})$$

for stimuli that only contain the respective frequencies and

- (b) Given orientations o_1 and o_2 , there exists a constant c_f so that for any frequency, \hat{f} ,

$$\mathcal{B}_{t,s_{o_1}}(\{(\hat{f}, o_1)\}) = c_f \cdot \mathcal{B}_{t,s_{o_2}}(\{(\hat{f}, o_2)\})$$

for stimuli that only contain the respective orientations.

Note that the stimulus corresponding to a single frequency and single orientation is a 2-dimensional linear sinusoidal pattern that changes in brightness. The stimulus corresponding to a single frequency and all orientations is a radial sinusoidal pattern; and the stimulus corresponding to a single orientation and all frequencies is a line. See Figure 3.

Answering the posed questions requires determining values for $\mathcal{B}_{t,s}$ (human sensitivities) for different stimuli. This is done experimentally. The blob phenomena presented here were observed in identification experiments. The output of such an experiment is a confusion matrix, $n \times n$ where n is the number of stimuli and possible responses. The matrix of integers describe response frequencies with rows corresponding to individual stimuli and columns to responses. We must tie responses to perception at this point so we make the following assumption.

Assumption 1. *If $\mathcal{A}(s) \geq \mathcal{B}_{t_{s_i}, s_i}$ at minimum time t_{s_i} for sufficiently many points for stimulus s_i , then s_i may be a response at time $t \geq t_{s_i}$.*

We don't assume that s_i will be a response because several stimuli may meet this criterion. At this point, we note there are many candidate response models corresponding to choice experiments. One class of response models considers behaviorally derived weights for each potential response and chooses the maximum weight. Another class represents alternative responses in a metric space and the response closest to a behaviorally derived representation is chosen. Yet a third class of response models are race models in which the first response to reach a threshold is the chosen response. For this application we use the race model to make a second assumption regarding choice. This is made out of convenience rather than evidence.

Assumption 2. *For $\{s_{i_1}, \dots, s_{i_M}\}$ in the response candidate set, the response corresponding to the minimum of $\{t_{s_{i_1}}, \dots, t_{s_{i_M}}\}$ is chosen.*

Lastly, since responses are not deterministic, we must introduce stochasticity. Again, there are many models for how stochasticity enters responses. One class of models considers information accumulation to be a noisy process, where information is gained (learned) and lost (forgotten or, more conservatively, inaccessible). In this model, this might be introducing Gaussian noise to $\mathcal{B}_{t,s}$ at each point in the decomposition. Alternatively, we may consider sampling the stimulus to be a noisy process. This would mean that $\mathcal{A}(s)$ is the the correct amplitude for a random sample of frequencies and orientations. Out of convenience, we model the former.

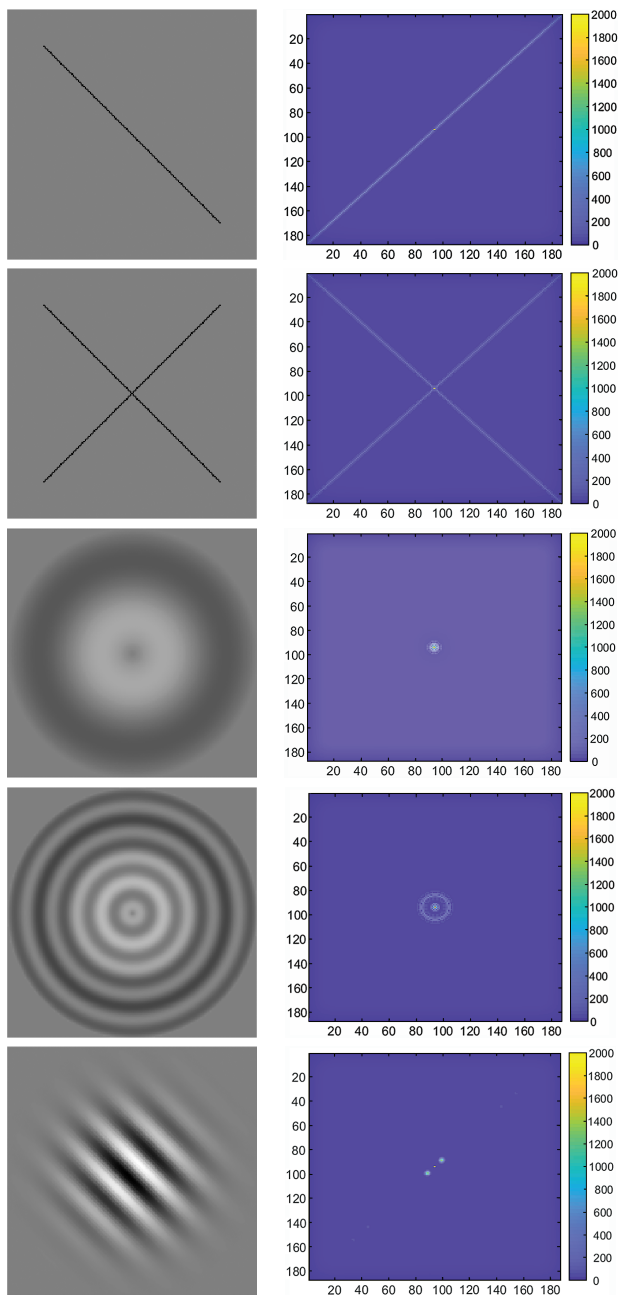


Figure 3. Example stimuli: (Top-to-Bottom) single orientation, two orientations, single frequency, two frequencies and orientation. (Left) Stimuli as presented to participants. (Right) Fourier space representation of stimuli. These are polar coordinates with angle denoting orientation and distance denoting spatial frequency. Brighter color denotes higher amplitude values.

Assumption 3. *There exists an experimentally determined parameter ϵ so that for all stimuli and all $(f, o) \in F \times O$,*

$$B_{t,s}(f, o) = \mathcal{N}(0, \epsilon) + B_{t,s}$$

for deterministic function $B_{t,s}$.

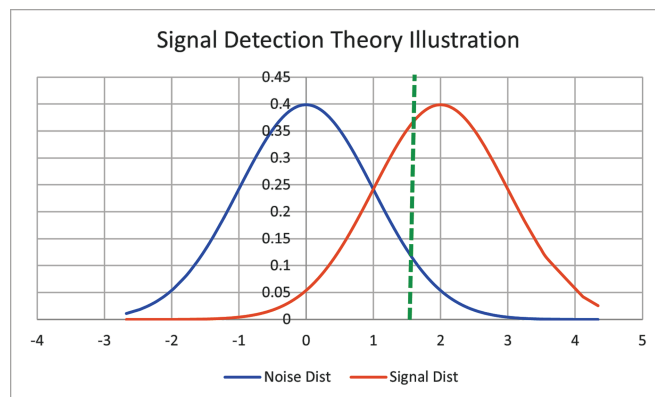


Figure 4. Signal detection theory: (orange) signal distribution, (blue) noise distribution, (green) decision boundary, x^* .

Signal detection theory. After attaining judgements in a confusion matrix, the experimenter will need an analysis technique to separate signals from different sources (or stimuli). A widely used approach is signal detection theory. We will call the no-stimulus condition the *noise* and the stimulus-presented condition the *signal* due, in part, to the history of signal detection theory in psychology and to focus the reader on a particular stimulus of interest. In psychology, signal detection theory is an application of statistical decision theory to stimulus detection. It describes a perceptual process where there is no fixed threshold for detecting a stimulus (or part of a stimulus) but rather an observer-specific threshold. Briefly, we assume that while the participant views a stimulus (including the null stimulus) samples are drawn that are classified as either coming from a signal distribution or a noise distribution. We model these distributions as Gaussian $\mathcal{N}(0, \sigma_N)$ and $\mathcal{N}(\mu_S, \sigma_S)$ on a common support where a decision boundary (x^* on the support) determines if participants respond that a stimulus (or part) was present or not. We fix the mean of the noise distribution to $\mu_N = 0$ and variance of the noise distribution to $\sigma_N = 1$ without loss of generality. Modellers often assume the variance of the signal distribution is equal to that of the noise distribution ($\sigma_N = \sigma_S = 1$) for computational simplicity and this assumption provides an easier interpretation later. From data, the mean of the signal distribution, μ_S , and x^* are fit so that

$$\int_{x^*}^{\infty} f_S(x) dx = P(\text{Respond Signal} \mid \text{Signal})$$

and

$$\int_{x^*}^{\infty} f_N(x) dx = P(\text{Respond Signal} \mid \text{No Signal}),$$

where f_S and f_N are density functions for the respective distributions. A picture communicates the concepts more efficiently (see Figure 4). Note, that the larger μ_S is, the

more correct identifications occur and the better sensitivity is. The standardized distance between the means is referred to as *sensitivity* and is denoted d' . Let H (*hit*) denote $P(\text{Respond Signal} \mid \text{Signal Stimulus})$ and FA (*false alarm*) denote $P(\text{Respond Signal} \mid \text{No Stimulus})$. Then

$$d' = z(H) - z(FA),$$

where z denotes the z -score. Since we are already using the word *sensitivity* to be inversely proportional to thresholds from \mathcal{B} we can call this computation **SDT sensitivity**.

The SDT sensitivity can provide insight on the relationships between functions \mathcal{B}_{t,s_i} and \mathcal{B}_{t,s_j} . [Jef18] conducted an experiment with diagonal line stimuli (see Figure 3 for example) presented for very short durations (50 ms) with contrast level adjusted for each participant so that accuracy in the identification task was 60%. There were two line parts (line 1 at 45 degrees, referred to as $l1$ and line 2 at 135 degrees, referred to as $l2$) that provided for four stimuli ($l1$, $l2$, $l1 + l2$, and blank). That study showed

- the probability of responding accurately in the identification task when both diagonals were present is lower than the product of probabilities in single line presentations (inhibiting; non-independence of orientation),
- SDT sensitivity decreased for each line part with an increase in the number of parts, and
- higher correct “blank” responses than correct responses when a stimulus was presented.

The SDT decrease implies the probability of $\mathcal{B}_{t,l1+l2}$ or $\mathcal{B}_{t,l1} \leq \mathcal{A}(l1)$ first is higher than the probability of $\mathcal{B}_{t,l1+l2}$ or $\mathcal{B}_{t,l1} \leq \mathcal{A}(l1 + l2)$ first. However, we’ve defined \mathcal{A} to be the amplitudes corresponding to the Fourier decomposition and since the study only considered simple stimuli, $\mathcal{A}(l1 + l2)$ agrees with $\mathcal{A}(l1)$ on the corresponding components of $l1$. This may lead us to believe that either the sensitivity functions are modulated by the exact stimulus being presented, or the other sensitivity functions (one for $l2$ and one for “blank”) are winning the race model more frequently for the $l1 + l2$ stimulus. In fact, the third finding shows that the “blank” sensitivity is not negligible.

These results may appear to be at odds with the complexity blob processing requirement, but they are not. In [Jef18] the $l1 + l2$ stimulus was not correctly identified above chance (25%), in many cases being confused with the “blank” stimulus, where nothing was shown. So the discrimination threshold for that stimulus was not reached to satisfy the complexity property. Concurrently, for higher-contrast stimuli, many authors have shown evidence that for longer display times, two oblique-angle stimuli should act independently of one another (single frequency [CK66] and line stimuli [Gil68]) and provide higher accuracy and faster response times in discrimination tasks. We conjecture the following:

Conjecture 2. *For small enough t , Blob processing is highly sensitive to low-frequency stimulus content. That is, for all stimuli, $\|\frac{d}{dt}\mathcal{B}_{t,s}(\text{small } f, -)\|$ decreases as t increases.*

In another experiment from the same study, [Jef18] ran the same identification task with a different set of stimuli. Two spatial frequencies were provided. For each frequency, a sinusoidal radial stimulus was produced with that frequency, one with 1 cycle (1c), and a second with 5 cycles (5c). A combination stimulus (1c+5c) and a “blank” stimulus were also used (see Figure 3). These frequencies have been largely agreed upon to be independent through adaptation studies [DeV77], masking studies [KKS73], and detection studies [GN71]. [Jef18] showed that SDT sensitivity increased when going from the 5c stimulus to the 1c+5c stimulus, but not from the 1c stimulus. The study concluded that there was a dominance effect of the 1c frequency over the 5c stimulus in that participants were more likely to respond that a stimulus was the 1c stimulus when the 1c part was added to the 5c part. This study provides some support for the second conjecture. Additionally, the 5c stimulus was often confused for the “blank” stimulus.

In a final experiment, participants were asked to identify which of four stimuli were presented. Stimuli were Gabor patches. The lower image in Figure 3 shows a Gabor patch, a stimulus with alternating black and white linear patterns at different orientations and frequencies. This task tests the ability to simultaneously discriminate between orientation and frequency. [Jef18] found that stimulus frequency discrimination was largely unimpaired by changes in orientation, but low-frequency perception deteriorated separation of same-oriented stimuli. A model that could explain this observations is that the sensitivity functions may have a uniform continuity constraint for a given t . The observation that responses for a given stimulus have systematic confusions across dimensions can be explained by sensitivity functions being more or less consistent across bands of orientations and frequencies. A final conjecture on sensitivity functions follows.

Conjecture 3. *$\mathcal{B}_{t,s}$ is uniformly continuous in f and o for each t . Further, $\mathcal{B}_{t,s}$ is continuous in t .*

While there may be other models that explain the collective findings across these experiments, we have three conjectures that can be tested by other paradigms. Without using mathematics, these disparate experimental results would lack a unifying concept to organize and tie them together. Recall, that Figure 1 contained rectangular stimuli. We now have a language and framework for describing the “right” features of the stimuli for explaining varied accuracy and reaction times. In Fourier space, the distance between the stimuli in set 2 are larger than the distances between stimuli in set 1. These distances may map directly to perceptual distances and the sensitivity of rectangles is likely

higher than that of squares due to the higher-frequency content.

Open Questions

As a psychological construct, the blob processing model provides a testable model for the concept of the blob. As a mathematical construct, the simple functional approach gives an intuitive and flexible base case.

To the degree possible, I encourage readers to consider some open questions that I find interesting for this particular phenomena:

- The inner product as a similarity measure was left undefined. When judging similarity is there a more appropriate group structure or action that captures behaviors?
- Are there topological properties inherent in the surface of a stimulus sensitivity $\mathcal{B}_{t,s}$ that are indicative of deficits or changes over time? Likewise, are there topological properties applicable to similarity space?
- This discussion avoided making claims about \mathcal{B} outside of a particular range of frequencies and orientations relevant to a stimulus. What would changes in the support of \mathcal{B} do for SDT sensitivity?

There is research in this space that considers more complex stimuli [TS22], and interested readers should consider some of the citations in this notice as a good entry point to the many viewpoints on the relationship between spatial frequency and orientation (especially [VSG89]). More generally, I encourage mathematicians to learn more about the very diverse world of mathematical psychology. There's a strong chance that a phenomena will fascinate you, and there's a guarantee that your expertise can be valuable in helping to describe the phenomena!

References

- [BC69] Colin Blakemore and Fergus W. Campbell, *On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images*, *The Journal of Physiology* **203** (1969), no. 1, 237–260.
- [CK66] F. W. Campbell and J. J. Kulikowski, *Orientalional selectivity of the human visual system*, *The Journal of Physiology* **187** (1966), no. 2, 437–445.
- [DeV77] Karen K. DeValois, *Spatial frequency adaptation can enhance contrast sensitivity*, *Vision Research* **17** (1977), no. 9, 1057–1065.
- [DeV82] Russell L. DeValois, *Early visual processing: feature detection or spatial filtering?*, *Recognition of pattern and form*, Springer, 1982, pp. 152–174.
- [Gil68] Alberta S. Gilinsky, *Orientation-specific effects of patterns of adapting light on visual acuity*, *JOSA* **58** (1968), no. 1, 13–18.

- [GN71] Norma Graham and Jacob Nachmias, *Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models*, *Vision Research* **11** (1971), no. 3, 251–IN4.
- [Jef18] Brett A. Jefferson, *An analysis of spatial frequency and orientation: General recognition theory application and blobloc*, PhD thesis, 2018.
- [KKS73] J. J. Kulikowski and P. E. King-Smith, *Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation*, *Vision Research* **13** (1973), no. 8, 1455–1478.
- [Loc72] G. R. Lockhead, *Processing dimensional stimuli: a note*, *Psychological Review* **79** (1972), no. 5, 410–419.
- [Lup79] Stephen J. Lupker, *On the nature of perceptual information during letter perception*, *Perception & Psychophysics* **25** (1979), no. 4, 303–312.
- [Mal15] Joseph Malkevitch, *Mathematics and psychology*, AMS Feature Column (2015).
- [PP11] James R. Pomerantz and Mary C. Portillo, *Grouping and emergent features in vision: toward a theory of basic gestalts*, *Journal of Experimental Psychology: Human Perception and Performance* **37** (2011), no. 5, 1331.
- [TS22] Hikari Takebayashi and Jun Saiki, *Restriction of orientation variability and spatial frequency on the perception of average orientation*, *Perception* **51** (2022), no. 7, 464–476, PMID: 35578551.
- [TG11] James W. Tanaka and Iris Gordon, *Features, configuration, and holistic face processing*, *The Oxford Handbook of Face Perception* (2011), 177–194.
- [THE8408] James Townsend, G. Hu, and R. Evans, *Modeling feature perception in brief displays with evidence for positive interdependencies*, *Perception & Psychophysics* **36** (1984), 35–49.
- [THK8807] James Townsend, Gary Hu, and Helena Kadlec, *Feature sensitivity, bias, and interdependencies as a function of energy and payoffs*, *Perception & Psychophysics* **43** (1988), 575–91.
- [Tre86] Anne Treisman, *Features and objects in visual processing*, *Scientific American* **255** (1986), no. 5, 114B–125.
- [VSG89] Norma Van Surdam Graham, *Visual pattern analyzers*, Oxford Psychology Series, Oxford University Press, 1989.



Brett Jefferson

Credits

Opening image is courtesy of Ryzhi via Getty.

Figures 1–4 are courtesy of Brett Jefferson.

Photo of Brett Jefferson is courtesy of Andrea Starr, Pacific Northwest National Laboratory.