DISENTANGLING MODELS OF EVIDENCE INTEGRATION

Andrei Teodorescu And Marius Usher
Tel-Aviv University
Email: andreite@post.tau.ac.il

Abstract

Decision making has been studied using a variety of experimental paradigms. Here, we focus on dynamically changing noisy perceptual stimuli that require the accumulation of evidence over time. Such decisions have often and successfully been accounted for by models implementing a psychophysically inspired sequential sampling framework. This framework represents a mechanistic approach to Signal Detection Theory where observers accumulate multiple samples of perceptual evidence to a predefined decision criterion. However, a multitude of such models exist which, despite their profound structural differences, all fit existing empirical data well. We propose an approach for comparing models which is based on isolating a specific model attribute to produce qualitative, rather than quantitative, predictions via computational simulations. Simulations demonstrate that, some models (mainly but not exclusively independent ones) speed up (due to statistical facilitation) while others slow down. Our results provide strong support for the presence of high level competition and against independent models of decision making.

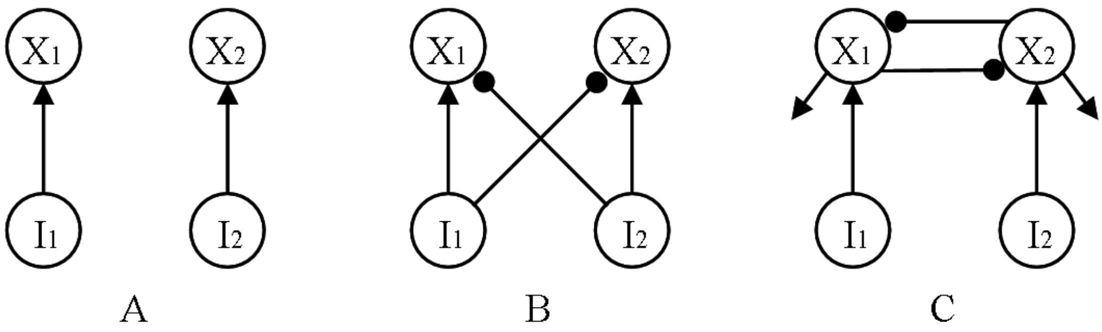
Perceptual decision processes are ubiquitous in all cognitive tasks as well as most daily activities. For example, one could be trying to decide if the traffic light, at a rapidly approaching intersection on a particularly foggy day, is red or green or whether the announcer at the train station just announced that her train will be leaving from platform B or from platform D. Research on fast perceptual decisions has led to the development of a number of reaction-time (RT; Luce, 1986; or as sometime labeled, *sequential sampling*) models, which share the assumption that evidence is integrated over time until enough has been accumulated for the formulation of a decision (Bogacz, Brown, Moehlis, Holmes & Cohen, 2006; Link & Heath, 1975; Ratcliff, 1978; Stone, 1960; Vickers, 1970).

Despite their similarities, these models still differ on fundamental assumptions about the nature of the accumulation process and its stopping rule. That is, whether the decision mechanism which is composed of the evidence integration stage and the termination rule, is *independent* or *competitive*. By independence we mean that each sample of evidence adds support for one of the alternatives without affecting the support for other alternatives. This is the case, for example in race models (Brown & Heathcote, 2008; Vickers, 1970; see illustration A in box), where each accumulator independently integrates the evidence supporting only one perceptual hypothesis and the decision process terminates when the first accumulator reaches a response criterion, irrespective of the state of the other accumulators at that time. Non-independent, *competitive* models, on the other hand, assume either that there is lateral inhibition between the evidence integrators (Usher & McClelland, 2001; Roe, Busemeyer, & Townsend, 2001; see illustration C in box) or that each sample of evidence adds support for one alternative and, at the same time, subtracts from the support for the other alternatives. This can be achieved via feed forward inhibition (Niwa & Ditterich, 2008; Roe, et al., 2001; see illustration B in box) or by maintaining a single tally (1-D variable) of subtracted evidence (Link & Heath, 1975; Ratcliff, 1978; Stone, 1960).

Box: Sequential Sampling models of choice-RT

Computer simulations were performed for four choice models, one independent race model (**Panel-A**; Brown & Heathcote, 2008; Smith & Vickers, 1988; Vickers, 1970), one input-competition model (**Panel-B**; Mazurek et al., 2003; Niwa & Ditterich, 2008) and two response competition models: 'max-next' diffusion (McMillen & Holmes, 2006; Ratcliff & Rouder, 1998) and LCA (**Panel-C**; Usher & McClelland, 2001). All models integrate noisy evidence (I1, I2,...) towards a response criterion.

The 'max-next' diffusion involves the independent accumulation of evidence for each alternative, like the race model, however the decision is triggered when the difference between the two leading accumulators reaches the response criterion ('max-next' stopping rule); the binary diffusion model (Ratcliff & Rouder, 1998) is a particular instances of the max-next diffusion, for n=2. The FFI diffusion implements feed forward inhibition in such a way that each accumulator integrates its own supporting evidence minus the average of all other (non-supporting) evidence. In the LCA, evidence for each alternative is accumulated competitively via lateral inhibition (each accumulator is suppressed by all the other accumulators in proportion to their activation) and imperfectly via information leakage, producing a model that interpolates between race and diffusion models (see Fig 2 (top)).



These two distinct classes of evidence integration and process termination mechanisms can lead to opposed predictions if evidence (or inputs) for the different alternatives are independently manipulated. To see this, consider first the independent model as it manifests in an independent race between two runners (i.e., the runners cannot assist or hinder each other in any way). Now, consider two such races: [1] a race between a fast runner (F) and a slow runner (S); [2] a race between the same fast runner (F) and a medium runner (M). On average, finishing times for race [2] would be faster than for race [1]. This happens since runner (F) is just as fast in both races but runner (M) is faster than runner (S). So, runner (F) loses more of his slower runs to runner (M) than to runner (S) resulting in a speedup of overall finishing times. This phenomenon is aptly named statistical facilitation (Luce, 1986; Raab, 1962; Townsend & Nozawa, 1995) and is most evident in independent processes. Let us now introduce competition into the race. Assume that, as the runner who is behind gets closer to the leader, her ability to slow the first runner down (say, by pulling on her shirt), improves. In contrast to the independent race, now race [2] will result in slower, rather than faster, finishing times, since the medium runner (M) has more opportunity than the slow runner (S) to hinder the fast runner (F). Following from this analogy, one can predict that manipulating task difficulty by increasing the perceptual evidence only for the weak alternative (i.e., replacing the slow runner with a medium runner) should provide a good way to distinguish independent from competitive models.

The aim of this study is to disentangle the tight cluster of independent and competitive models through a combined computational and experimental approach. To this end, we use a task in which, participants are required to choose out of four flickering lights the one that is flickering fastest. At each time frame, the stimulus provides independent evidence for the four perceptual alternatives. We manipulate the amount of evidence in this task, in a way that resembles replacing runner (S) with runner (M) in the example above. Thus, in all conditions the target stimulus is kept constant while the difficulty of the task is determined only by the strength of the non-target stimuli. As the flickering rate of the principal non-target (the second fastest flickering light) is increased, the task becomes more difficult. However, in addition to increasing difficulty which usually results in delayed RTs, this manipulation has the added effect of increasing statistical facilitation which should result in expedited RTs. As we show in the simulations, the conflict between the expected increase in RT due to higher difficulty and RT acceleration due to statistical facilitation, results in qualitatively different predictions for competitive and independent models.

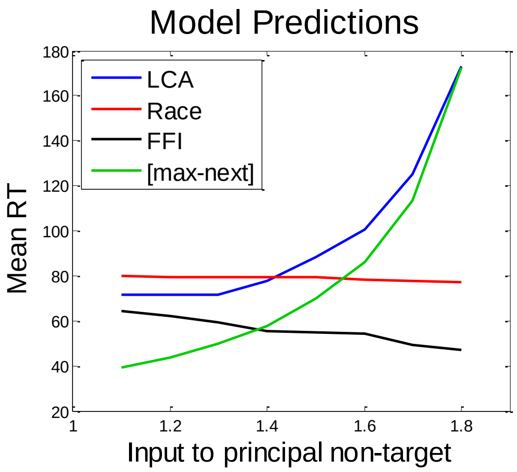


Figure 1: Simulation: Mean-RT for three choice models, as a function of the input strength of the brightest non-target (I2); I1 was kept constant at 2, and normalization was maintained for the other units. Response-criteria for the three models were set so that the accuracy for each level of I2 was nearly identical across models.

One drawback of this procedure, however, is that both independent and competitive models can incorporate a pre-decisional normalization procedure (Brown & Heathcote, 2008; Usher & McClelland, 2001) which can represent either a feature of the stimulus or an early perceptual mechanism. Such a normalization, which is most often implemented by maintaining the sum of the inputs constant, leads to a decrease in the strength of the target as a result of the increase in non-target strength in the difficult condition. This inevitably leads to a slow-down of fast, correct responses. The problem is that if, empirically, such difficulty manipulations result in a slow-down, one can then only conclude that the choice mechanism is subject to some kind of competition. Nevertheless, this competition can be interpreted either as response competition between accumulators or as input competition via the normalization of evidence prior to the decision process. We aim to further distinguish between these different competitive mechanisms, by co-varying the brightness of the remaining two non-targets. We can now, make the task difficult, by increasing the principal non-target towards the target, while at the same time maintain the normalization by lowering the brightness of the remaining non-targets. This additional manipulation makes both independent and competitive models invariant to the inclusion of a pre-decisional normalization and eliminates the ambiguity regarding the competitive mechanism responsible for the observed behavior.

Simulations were run to formally evaluate the effect of increasing task difficulty via the augmentation of supporting evidence for the strongest non-target, on the mean-RT of a pure independent (*race*) model, and three competitive models: *[max-next] diffusion*, *FFI diffusion* and *LCA* (see Box). As one can see in Figure 1 (top), as the evidence strength for the main non-target (I2 - x axis in Fig 1) increases, the independent race and FFI diffusion models predict a speedup of RT and the response competition models (LCA and [max-next] diffusion), predict a slowdown in RT with increasing task difficulty. While the observed speedup effect is due to statistical facilitation, the slowdown effect is the result of increased competition between the target and the principal non-target. Our experiment was designed to test these contradictory predictions.

Method

Five Tel-Aviv University students, all female, participated in the study as part of their Introduction to Psychology course requirements. Each participant was tested in three to four sessions no more than four days apart. Each session was about 45 minutes long. All participants had normal or corrected to normal vision.

All stimuli in this experiment were presented on a Samsung SyncMaster 943b LCD monitor. The stimuli were composed of four homogenous, round, white patches on a black background. The alternatives were positioned at the four corners of an imaginary square relative to a central plus sign fixation point. Each patch flickered (abrupt transitions from a gray level of 1 to 0 and vice versa) randomly and independently of the other patches over the course of each trial.

For the 'easy' condition the mean interval between consecutive 'on' frames was 11.11 frame for the target, and 33.33 frames for the distractors. For the ‘hard’ condition the mean interval between consecutive 'on' frames was 11.11 frames for the target, 20.00 for the main distractor and 60.00 frames for the secondary distractor. To achieve stochasticly dynamic stimuli, these intervals were continuously perturbed by augmenting them with a uniformly distributed variable in the (-2, +2) range. Refresh rate was set at 60Hz (16.6ms per frame). Thus for the 'easy' condition average flicker rates were 5.45 frames per second (fps) and 1.82fps for the target and distractors respectively. For the 'hard' condition average flicker rates were 5.45fps, 3.00fps and 1.20fps for the target, main distractor and secondary distractor respectively. The first 'on' frame of each alternative was determined randomly on each trial (uniform distribution between 1 and 10) to prevent repetitive circular (periodic) patterns.

All trials were randomly assigned to one of two possible conditions: 'easy' and 'difficult'. Each trial began with a large fixation cross at the center of the screen which stayed on for one second then turned into a smaller one just as the target and distractor stimuli appear. Stimuli stayed on until the subject made his response. After the participants' response came a short 1sec ISI followed by the next trial. The fixation cross remained on the screen throughout the experiment but it briefly increased in size during each of the ISIs and came back to its original size at its end in order to draw attention back to the center of the screen before the beginning of the next trial.

Answers were given via the right number pad of a standard keyboard. The 1, 3, 7 and 9 keys represented the lower left, lower right, upper left and upper right patches respectively. Subjects were asked to place two fingers from the right hand on the 3 and 9 keys and two fingers from the left hand on the 1 and 7 keys.

Participants were presented with blocks of 20 trials. Each block consisted of an equal number of trials from each condition. After each block there was a self-timed intermission to allow the subject to rest its eyes and prepare for the next trial. During each of these breaks the subjects were presented on the screen with their average accuracy for the last block. Breaks were ended by pressing any key on the keyboard.

Subjects were instructed to try and maximize both accuracy and response time such that if they reached 100% accuracy they should try to respond faster. Participants were also instructed to keep their eyes focused on the fixation cross throughout the trial though in the absence of an eye tracker there was no way to verify that they actually complied with this request. The experiment was held in a partially darkened room.

Results and Discussion

As can be seen in Figure 2 and according to the predictions of competitive models, participants were slower to respond in the ‘difficult’ condition ($M=1.38s$, $SD=0.10$) than in the easy condition ($M=1.28s$, $SD=0.12$; $t(4)=3.5$, $p<0.05$). On the other hand, independent models are unable to predict the slowdown effects observed since they do not incorporate a competitive element.

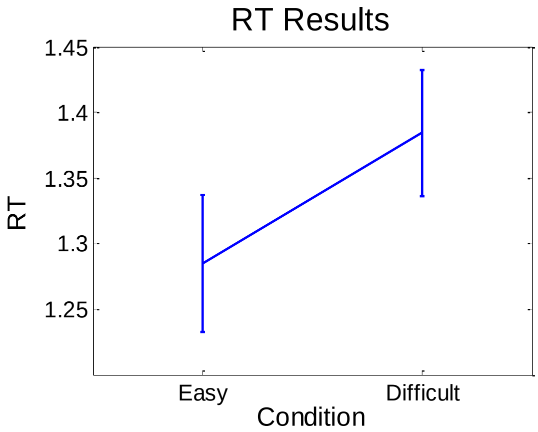


Figure 2: RT results for the 'easy' and 'difficult' conditions.

The aim of this study was to distinguish between models of perceptual choice that differ with regards to their implementation of competitive interactions in the accumulation process. To do so, we developed an experimental paradigm that produces diverging predictions for the different models. The experiment involved a manipulation of difficulty via the manipulation of evidence strength to a week alternative while maintaining the total sum of evidence constant. We found that participants compensate for task difficulty by increasing decision-time. In contradiction with this compensation pattern, the race model, predicts a speedup, rather than a compensatory slowdown of RTs for the more difficult stimuli. This is a result of statistical facilitation between two independent accumulation processes. Therefore, independent accumulation models are not able account for the data presented here.

The FFI diffusion model also fails to account for the data. This is, however, due to a different characteristic of the model. In the FFI, the momentary inhibition felt by any given accumulator is calculated as the *mean* of the other inputs on that time step. Note that the since target strength does not change and all other inputs are continuously normalized, the inhibition felt by the target (mean of non-target strengths) remains constant. Therefore, in the FFI model the rate of evidence accumulation in favor of the target is not affected by the difficulty manipulation. The apparent invariance of the FFI model to our manipulation renders the competitive element of the model ineffective and leaves the model vulnerable to the opposite (speedup) effect of statistical facilitation. This is why we observe a speedup of RT for the FFI in Figure 1.

Both the LCA and diffusion models have a competitive response termination mechanism, which produces a compensatory slowdown in difficult conditions. When difficulty is increased by adding evidence favoring a weak (non-target) alternative, decision time slows down due to stronger inhibition of the target by the, now more strongly activated, principal non-target accumulator.

Our results provide strong support for competitive interactions beyond a simple normalization of input strength. In addition, competitive mechanisms inhibition on one alternative is generated as a function of the mean of the other alternatives seem to be unable to capture the effects of our manipulation. Future research is needed to examine if competition is a general property of the decision system and whether it depends on task demands (accuracy vs. speed emphasis) and types of stimuli. One possibility is that race models could underlie mechanisms, like fast eye-movement (saccade) control, where speed may be prioritized over accuracy (Ludwig, Gilchrist, McSorley & Baddeley 2005). In addition, manipulations that produce diverging predictions are needed to distinguish between types of response competition for both two and multi alternative choice tasks.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, 113, 700-765.

Brown, S.D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57, 153-178.

Link, S.W., & Heath, R.A. (1975). A sequential theory of psychological discrimination. Psychometrika, 40(1), 77-105.

Luce R.D. (1986). Response Times: Their Role in Inferring Elementary Mental Organization. Oxford, UK: Oxford Univ. Press.

Ludwig, C.J.H., Gilchrist, I.D., McSorley, E., & Baddeley, R.J. (2005). The temporal impulse response underlying saccadic decisions. *Journal of Neuroscience*, 25, 9907–9912.

Mazurek, M.E., Roitman, J.D., Ditterich, J., & Shadlen, M.N. (2003). A role for neural integration in perceptual decision making. *Cerebral Cortex*, 13, 1257.

McMillen, T., & Holmes, P. (2006). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, *50*(1), 30–57.

Niwa, M., & Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *Journal of Neuroscience*, *28*(17), 4435

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Ratcliff, R., & Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347-356.

Raab, D.H. (1962b). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24, 574–590

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108(2), 370-392.

Smith P.L., Vickers D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*. 32, 135–168.

Stone, M. (1960). Models for reaction time. *Psychometrika*, 25(3), 251-260.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical

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