

Why EPIC was Wrong about Motor Feature Programming

David Kieras (kieras@umich.edu)

Electrical Engineering & Computer Science Department, University of Michigan
2260 Hayward Street, Ann Arbor, Michigan 48109 USA

Abstract

The EPIC computational cognitive architecture was among the first to propose representing motor movement constraints explicitly in the form of motor processors that implemented a specified time course for the preparation, initiation, and production of movements. A key feature of this proposal was that movements were specified in terms of features, and movement preparation time was linear with the number of features that had to be prepared before a movement was initiated. While successful in modeling many high-speed tasks involving choice reaction times with keypress responses, serious difficulties appeared in modeling high-speed visual search tasks involving saccades and mouse movements. A reappraisal of the basis for EPIC's assumptions requires a critical change: *visually aimed manual and ocular movements require no feature preparation time.*

Keywords: cognitive architecture; motor processing; motor features; S-R compatibility; spatial compatibility

Introduction

The EPIC architecture for human cognition and performance provides a general framework for simulating a human interacting with an environment to accomplish a task. Due to lack of space, the reader is referred to Kieras & Meyer (1997), Meyer & Kieras (1997), or Kieras (2004) for a more complete description of EPIC. Figure 1 provides an overview of the architecture, showing perceptual and motor processor peripherals surrounding a cognitive processor; all of the processors run in parallel with each other. To model human performance of a task, the cognitive processor is programmed with production rules that implement a strategy for performing the task. When the simulation is run, the

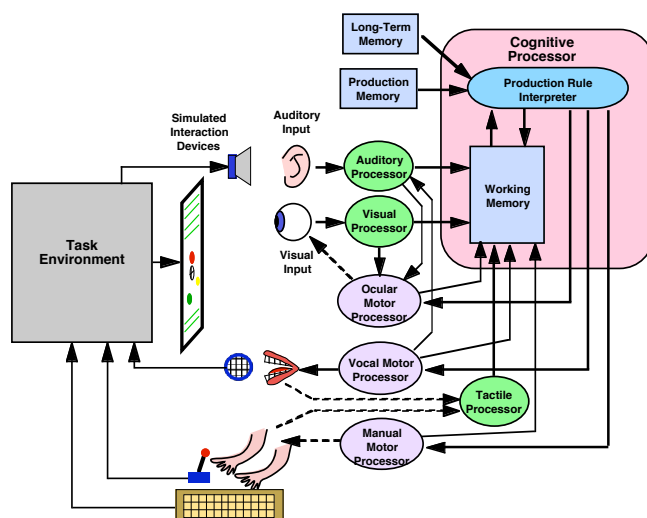


Figure 1. The overall structure of the EPIC architecture.

architecture generates the specific sequence of perceptual, cognitive, and motor events required to perform the task, within the constraints determined by the architecture and the task environment. Components of EPIC, especially the motor processors, have been incorporated into other cognitive architectures that use their own cognitive processor.

Motor Feature Preparation

Meyer and Kieras (1997) argued that a cognitive architecture must explicitly represent the constraints on motor activity in order to comprehensively account for task performance. They specified these constraints in the EPIC cognitive architecture in terms of motor processors that were equal in status to perceptual processors and the cognitive processor. These motor processors, one for each of the manual, ocular, and vocal motor modalities, accept symbolic movement commands from a production-rule cognitive processor, and then generate simulated movements that are inputs to a simulated task environment. Their characterization focussed on the temporal constraints, not on how muscle activity would be controlled, and can be summarized as follows:

1. Movements are described in terms of motor features, such as the direction and distance of a pointing movement, or the hand and finger used for a button-pushing movement. The type of movement, the *style*, was considered the dominant feature within each movement modality.

2. When a movement is commanded, the motor processor prepares each feature serially, requiring a constant time per feature, estimated as 50 ms. When all features have been prepared, the movement is initiated. After an initiation time delay (also estimated as 50 ms), the mechanical movement begins.

3. Once prepared, the features for a movement are retained in the motor processor. If a movement is repeated, its features do not have to be prepared, and the movement can be initiated immediately.

4. The motor processor can be commanded to prepare one or more movement features in advance; these are stored in the motor processor. When the movement is commanded, the previously prepared features do not have to be prepared, allowing the movement to be initiated sooner by the amount saved in preparation time.

5. The feature preparation mechanism is used for the motor processors in all modalities; the only difference is in the specific feature structure of different movements possible in each modality.

Meyer and Kieras based the motor processor assumptions on the available literature on motor control (see Rosenbaum, 1991 for an overview). Because the motor control area is seriously under-researched (Rosenbaum, 2005), the only

useful theoretical concept available was Rosenbaum's theory of motor feature programming, and so it was adopted. However, any cognitive architect has to go beyond the specific literature to some extent by simplifying and generalizing the empirical effects and available theory to produce a conceptually and practically manageable architecture. Uniformity of mechanism is a compelling first approximation as well. Meyer and Kieras therefore assumed that feature preparation held for all motor modalities and that the per-feature time was constant and uniform for all features and modalities. They also proposed the specific features for various movements and postulated dependencies between them. For example, manual pointing movement features could not be re-used in eye movement feature preparation; changing hands or fingers could reuse the remaining manual features, but changing the style of manual movement requires all features to be prepared.

Kieras and Meyer successfully constructed many EPIC models for high-speed choice reaction tasks, especially dual tasks, with these motor processors (summarized in Meyer & Kieras, 1999). Depending on the details of the task, motor feature preparation time often set a substantial constraint on other processes in the models, especially if features could be prepared in advance. However, since many experiments are done with the exact required response movement randomized over trials, the net effect of feature preparation usually is simply to produce an average preparation time that serves as a component in the overall latency of response. Also most of the modeled experiments involved button presses, typically using laboratory methodology in which the stimulus location is constant and the fingers are pre-positioned on the alternative response keys, meaning there is little or no eye or hand movement.

Since EPIC has been one of the few cognitive architectures that attempted to represent motor processes and constraints, even in highly abstracted form, its analysis of motor processing has been explicitly adopted in other architectures, in particular, the widely used ACT-R/PM and current ACT-R architectures (e.g. Anderson & Lebiere, 1998). Thus the status of EPIC's characterization of motor processing has broad relevance and concern to the cognitive architecture community as a whole. This paper presents why a major revision in this characterization is required: *visually aimed manual and ocular movements require no feature preparation time.*

Symptoms of the Problem

When models for high-speed visual search tasks were constructed, it proved to be extremely difficult to fit basic latency data given the constraints on ocular feature preparation. For example, models were constructed for Findlay's (1997) results for latency and accuracy in the first saccade in a conjunctive visual search task. Findlay observed that the latency of the first saccade was only about 250 ms, which was quite difficult to obtain in the EPIC architecture with its standard timing parameter values. Due to the syntax and semantics of the production rules, two production rule cycles are required to identify the target of the eye movement, for a total of 100 ms. A motor initiation requires 50 ms. An eye movement in the task required

preparing an average of one feature, for an additional 50 ms. The total is 200 ms, which leaves only 50 ms total for stimulus transduction and recognition, which seems implausibly short — 100 ms seems a more reasonable perceptual processing time.

In a more complex visual search task (such as in Kieras & Marshall, 2006), there are enough eye and hand movements that feature preparation time can sum to several hundred ms in the total RT. A more complex task strategy will also require more production rule firings to choose the next fixation target, making it even more difficult to fit the commonly observed 250 ms delay between successive saccades, even if multithreaded production rule strategies are deployed.

While the difficulty of programming a model is not normally grounds for rejecting a model, it is significant if the difficulty is due to a cognitive architectural feature. A cognitive architecture is supposed to capture the underlying mechanisms and processes of human activity; it is natural to expect that simple activities should have reasonably simple representations in the architecture. So undue difficulty in constructing a model for a straightforward task is a strong suggestion that the architecture is incorrect.

In the case of the visual search task modeling, it was observed that setting the feature preparation time to zero for aimed manual and ocular movements gave the strategy programming adequate "breathing room" in fitting the data. This led to a re-examination of empirical literature behind this basic feature of EPIC's motor processors to see if the original reasons for the motor feature programming were still justifiable.

Reappraising the Literature

Manual movement feature preparation

The seminal experimental demonstration of motor feature preparation is Rosenbaum (1980) in which participants made button-press movements in response to precues and cues. The experimental task is diagrammed in Figure 2. On each trial, the participant received a precue which specified some of the putative features of the movement, then a cue, which specified the exact movement, whereupon the participant made the response movement. The latency of the initiation of the movement was recorded. More specifically, as shown in Figure 2, the response buttons were a set of eight buttons arranged in two parallel rows, one on the right, and one on the left, running forwards towards the display, and rearwards towards the participant. The two center buttons in each row were the "home" buttons; at the start of the trial, the participant held each home button down with their left- and right-hand index fingers, and in response to the cue, moved one of the fingers to the response button; the time of release of the home button is the RT. The buttons were color-coded; the participant was practiced in associating the color codes with the physical location of the buttons. The response cue was a colored disk appearing on the display that designated which button to press.

The precue was presented on the display before the cue, and consisted of three letters, one for each putative feature of the movement which Rosenbaum described as *Direction*,

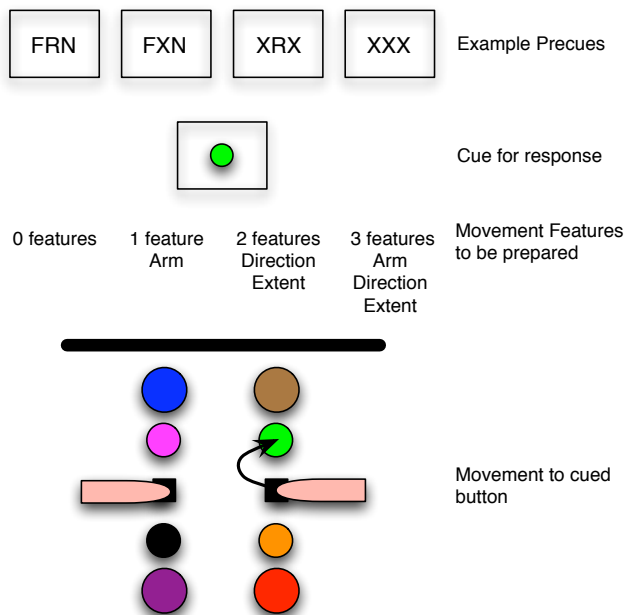


Figure 2. The Rosenbaum task. First a precue appears – four example are shown. Then appears a color-coded cue designating the button to be pressed. Depending on the precue, some number of movement features must be prepared, then the participant moves the left or right index finger from the home button to the designated button. The response buttons are hidden from the participant's view.

Arm, and Extent. For example, FRN stood for *forward, right, near*, which completely specifies the features of the movement to be made; in contrast, XRX specifies only a single feature, *right*. The participant was practiced in interpreting the precues. The logic of the paradigm is that when the cue appeared, the participant would have to prepare the remaining features before the movement could be initiated. Figure 2 shows additional examples of precues that vary the number of features that would have to be prepared before the movement could be initiated. The more features needing to be prepared, the greater the latency should be, and the results confirmed the prediction: zero, one, two and three features produced latencies of about 300, 450, 550, and 700 ms respectively. However, there were subtle and confusing specific-feature effects: different features appeared to require different times to prepare (ranging from 150 to 200 ms), and some features required different times depending on their values; for example, near movements were initiated faster than far movements, and more so if more features had to be prepared. Thus while demonstrating feature preparation time effects, the effects as presented were a complex mixture of general and specific effects.

Theorists of choice RT paradigms usually postulate a *response selection* stage of processing in which the stimulus is mapped to the response to be made. Motor feature preparation would follow response selection and should be a distinct process. However, it is clear that performing this

task requires some complex mappings - first from the letter codes to movement features, and then from the cue color to the button. An immediate question that arises is the extent to which the effect of the precue is actually a response selection effect - maybe the precue is assisting response selection, not movement preparation.

To eliminate the possibility that purely cognitive response selection effects were responsible for the latency effects, Rosenbaum conducted a second, *decision task*, experiment in the participant viewed the precue and then the cue, and rather than making the response movement, made a vocal response for whether or not the cue was valid (consistent with) the precue. The resulting RTs showed a strong effect of the number of precued features, but no effect of the specific features or feature values. Even though the primary effect of number of features was present in both experiments, Rosenbaum claimed that the lack of feature-specific effects meant that response selection effects were not responsible for the differences in movement RTs. However, this argument is hard to understand – the feature preparation concept would not seem to require feature-specific effects, which in any case are hard to explain. Additionally, the logic of deciding which movement to make would seem to overlap a lot with deciding whether the movement could be made. The present author correlated the mean decision RT with the mean movement RT for each precue condition, and discovered that 91% of the variance in movement RT is accounted for by the decision RT. This strongly suggests, contrary to Rosenbaum's claim, that most of the movement RT is accounted for by some form of response selection process, even if there are specific feature effects.

Response Selection Effects: S-R compatibility

A long-studied aspect of response selection is *S-R compatibility*, which can be described as the ease with which the mapping from stimulus to response can be made. See Proctor & Vu (2006) for a recent review, and Rosenbaum & Newell (1987) or John, Rosenbloom, & Newell (1985) for computational model accounts of some forms of S-R compatibility. One feature of Rosenbaum's task is that both the precues and the cues would require a complex mapping to the actual response movements. Goodman and Kelso (1980) examined this issue in a critical but usually overlooked response to Rosenbaum. They first replicated Rosenbaum's results using color words or number labels for the target buttons. In a second study they used a precue and cue display, diagrammed in Figure 3, consisting of an array of lights in the same spatial arrangement as the response buttons. Precues were indicated by illuminating the lights corresponding to the buttons consistent with the precued movement features. For example, the FRN precue would illuminate the single light for the button corresponding to the three features, while XRX precue would illuminate all the lights for the right-hand side of the button set. The cue would then consist of the single light for the to-be-pressed button. This presentation has an especially powerful form of S-R compatibility termed *spatial compatibility* – the spatial properties of the stimulus map directly to the spatial properties of the response.

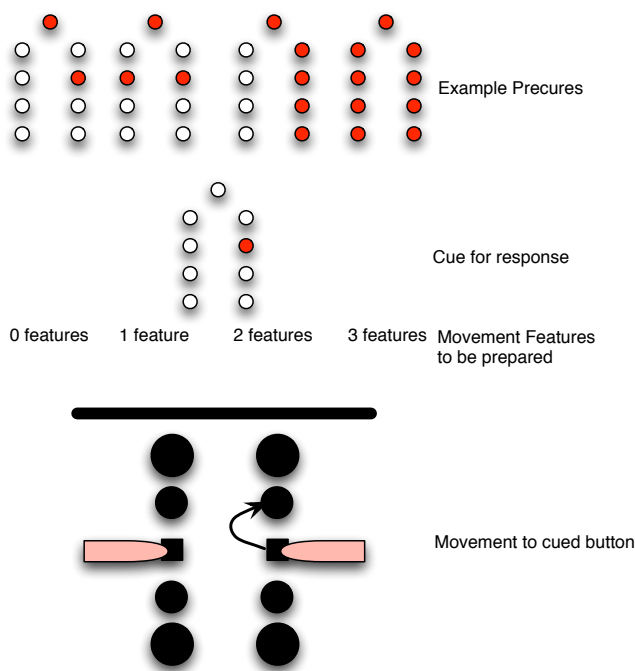


Figure 3. The Goodman & Kelso version of the Rosenbaum task. First a precue appears on an array of lights that matches the layout of the response buttons. The top light indicates a precue (vs. cue) display. Then in the same array appears a cue designating the button to be pressed. Depending on the precue, some number of movement features must be prepared, then the participant moves the left or right index finger from the home button to the designated button.

Compared to Rosenbaum's and the replication presentations, this spatially-compatible presentation of precue and cue information drastically reduced the effect of number of precued features. In contrast to the 300-700 ms range of Rosenbaum's latencies, the range was only about 250-350 ms.

S-R compatibility is normally assumed to be a response-selection process; there is nothing in the movement feature concept that suggests S-R compatibility would be involved. That is, once the response has been selected, the spatial similarity of the stimulus to the response should be irrelevant to computing the movement features. This is a further strong suggestion that Rosenbaum's effects were actually response-selection effects rather than movement preparation effects.

Response Selection Effects: Hick's Law

A second major aspect of response selection difficulty is the number of possible responses in the selection set, long codified as Hick's Law (Hick, 1952), which states that the RT in a choice reaction task is proportional to \log_2 of the number of alternative (possible) responses. One consequence is that if the number of possible responses is held constant, then the RT should be constant. This would

take ordinary response selection effects out of the picture, leaving only motor feature programming to produce RT differences. Thus, Goodman and Kelso (1980) conducted an additional experiment that held the number of possible responses constant at two. Using the same compatible display, they precued two possible responses by precuing both possible values of a single feature, such as illuminating the lights for both right and left forward near buttons. Also included were ambiguous precues that illuminated two lights, but which had no feature values in common, such as left-rearward-far and right-forward-near. These results were highly persuasive: the movement RTs were virtually identical at about 300 ms for all feature precue types, including the ambiguous precues. Apparently the specific movement features involved were irrelevant; what matters is only the difficulty of response selection, governed in this case by the number of possible responses.

It is also an old result that Hick's Law effects disappear in the presence of high S-R compatibility (Teichner & Krebs, 1974). Goodman & Kelso's highly compatible presentation of cues and precues drastically reduced the putative feature programming effects, and when the number of alternatives was held constant, they disappeared altogether.

This suggests that other aimed movement tasks in which S-R compatibility is manipulated might shed light on whether movement feature preparation is involved. That is, if S-R compatibility results in no Hick's Law effects, then there would be no response selection effects to be confused with feature preparation, and then perhaps other evidence of motor feature preparation would be visible, such as a movement latency long enough to have "room" for something like 50 ms or more per feature, and evidence of feature reuse, as described above, in which a repeated movement could be initiated more quickly.

Dassonville, Lewis, Foster, and Ashe (1999) had participants make joystick movements to place a cursor on visible targets arranged in a circle around the starting position, with various cues that differed in compatibility. In highly compatible mappings, there was no effect of the number of possible targets (no Hick's Law effect) and a latency of only about 300 ms. If the cue/response was repeated, the second response was substantially faster in the incompatible mappings, but not in the compatible mappings.

Wright, Marino, Belovsky, and Chubb (2007) had participants move a physical stylus from a starting point to one of several target pads arranged in an arc. The movement target and response cue was indicated by illuminating the pad, a perfectly compatible S-R mapping. There was no Hick's Law effect of the number of targets, the latencies were about 250 ms, and there was little or no effect of repetitions.

These results all point to the same conclusion: The motor feature preparation hypothesis states that features should require substantial time to prepare before a movement could be initiated and then could be reused in subsequent movements. Instead the effects are due to response selection effects described by Hick's Law, and when these effects are removed by highly compatible specifications of movement targets, reuse effects disappear, and the movement is launched so rapidly that there is no time to spare from other

aspects of EPIC's architecture for feature programming to occur.

The same story for eye movements

A parallel story appears in the case of eye movements. Another key demonstration of motor feature programming was Abram & Jonides (1988) who applied the Rosenbaum feature preparation paradigm to saccade preparation. The movement task, illustrated in Figure 4, consisted of a saccade to one of four targets, two on each side of the fixation point at different distances. The precue and cue were shown in four circles, two on each side of the fixation point, inside the actual targets. The saccade responses required were *anti-saccades* — the eye movement had to be made in the *opposite* direction from the precue or cued direction, an S-R incompatibility. The saccade latencies

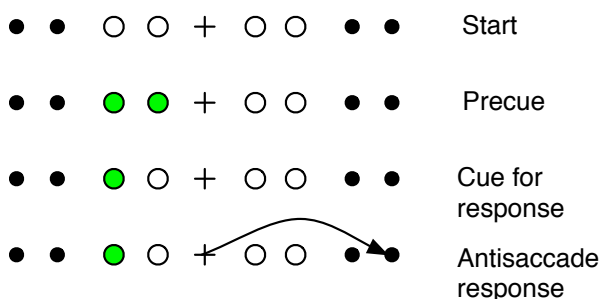


Figure 4. The Abram and Jonides task. The trial starts with the participant fixating the central cross. A precue is then presented showing e.g. the direction and both possible extents of the movement. Then a cue appears designating the actual movement target, one of the four small outer circles. The participant responds by fixating the target at the same distance but opposite direction as the cue.

increased by about 50 ms per feature preparation required, consistent with the feature preparation model. But in a second experiment, they required compatible response saccades, and held the number of precued alternatives constant at two. While there were some feature-specific differences, the saccade latency was basically constant across number of precued features, corresponding to the Goodman & Kelso (1980) results with the number of possible responses held constant. Again the feature-preparation effect seems to be confounded with a response selection effect.

What if the cue and response are more compatible? As illustrated in Figure 5, Crawford and Mueller (1990) used targets that were six lights, three on each side of the fixation point. A precue consisted of a background illumination around the possible target; the cue was illuminating the target light itself; in response, the participant made an eye movement to the illuminated target. The precue locations were either the same as the target (valid), different from the target (invalid), or at the fixation point (neutral), and presented either 100 ms or 500 ms before the cue. The results were very short latencies (about 250 ms), a small benefit of valid or neutral precues if the precue-cue delay was short, and no effect at all if it was long. Such an effect

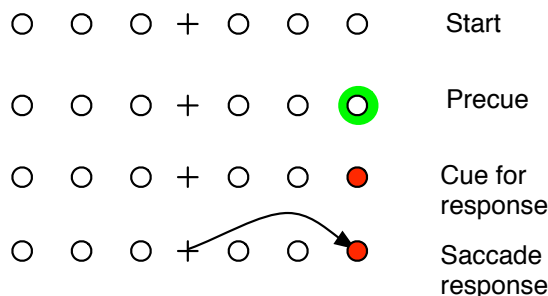


Figure 5. The Crawford and Mueller task. The trial starts with the participant fixating the central cross. A precue is then presented, e.g. a valid cue designating the future movement target. Then a cue appears designating both the actual movement target and acting as the stimulus for the movement. The participant responds by fixating the target.

would not be expected from the motor feature preparation concept - if anything, the benefit of the precue should be larger with more time. Rather the delay results suggest some low-level visual effect on saccade initiation.

Additional studies further clarify the compatibility effects for eye movements. Lee, Keller, and Heinen (2005) had participants make eye movements to memorized color-coded locations in a circular array given a color cue, not unlike Rosenbaum's approach. Hick's Law effects were observed. Kveraga, Berryhill, and Hughes (2002) and Kveraga, Boucher, and Hughes (2005) used targets arranged in a circle or semicircle, and the movement cue was co-located with the target, producing no Hick's Law effect. However, if anti-saccades or key presses were required to the same stimuli, Hick's Law effects were obtained.

The results for eye movements point to the same conclusion as for aimed manual movements: Effects suggesting motor feature preparation for eye movements are better explained as response selection effects accounted for by Hick's Law, and when these effects are removed by highly compatible specifications of movement targets, there is no evidence of feature preparation and the movement is launched so rapidly that there is no time to spare for feature programming to occur.

Conclusion

It was wrong

Empirically, once the target has been visually identified, an aimed manual movement or eye movement can be quickly launched to it without any S-R translation or motor feature programming delays; there is no evidence of feature programming effects. In terms of the EPIC architecture, once the production rules have identified the target of a movement as an object currently visible, and passed the identity of that object to the motor processor in a movement command, the movement will be initiated without any feature programming time. There seems to be no reason to maintain feature preparation delays for aimed movements in the architecture at the cost of making the models substantially more difficult to fit to important classes of

data, especially in the high-speed performance tasks that motivated the design of EPIC. This original feature of EPIC was simply an incorrect overgeneralization. Fortunately, the solution is simple: set the per-feature preparation time for saccades and aimed manual movements to zero.

Implications for previous models

What effect does this change have on previous models built with EPIC? As mentioned earlier, most of the models in the original Meyer & Kieras (1997, 1999) work used keypress responses, which are not affected by this correction because they would not seem to be aimed manual movements (but see Welford, 1971).

Furthermore, because experimental results are typically aggregated over specific response movements, the net effect is that previous models using aimed manual movements or eye movements have a variable component of response time that instead of being due to movement preparation, has to be reattributed to stimulus encoding or response selection. At this point the theoretical implications appear to be minor.

Should feature preparation be discarded for keypress movements as well? Unfortunately, this question cannot be easily answered because the motor control literature remains so sparse (Rosenbaum, 2005) that we are still in the earliest stages of our theoretical understanding of how movements are performed. An interim heuristic would be to assess whether keystroke feature preparation can be replaced by changing the response selection strategy.

Acknowledgment

This work was supported by the Office of Naval Research, under Grant No. N00014-06-1-0034.

References

- Abrams, R.A., & Jonides, J. (1988). Programming saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 428-443.
- Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Crawford, T.J., & Muller, H.J. (1992). Spatial and temporal effects of spatial attention on human saccadic eye movements. *Vision Research*, **32**, 293-304.
- Dassonville, P., Lewis, S.M., Foster, H.E., & Ashe, J. (1999). Choice and stimulus-response compatibility affect duration of response selection. *Cognitive Brain Research*, **7**, 235-240.
- Findlay, J. (1997). Saccade target selection during visual search. *Vision Research*, **37**, 617-631.
- Goodman, D., & Kelso, J.A.S. (1980). Are movements prepared in parts? Not under compatible (naturalized) conditions. *Journal of Experimental Psychology: General*, **109**, 475-495.
- Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, **4**, 11-26.
- John, B. E., Rosenbloom, P. S., & Newell, A. (1985). A theory of stimulus-response compatibility applied to human-computer interaction. In *Proceedings of CHI 1985*, New York: ACM. pp. 212-219.
- Kieras, D.E. (2004). EPIC Architecture Principles of Operation. Web publication available at <ftp://www.eecs.umich.edu/people/kieras/EPIC/EPICPrinOp.pdf>
- Kieras, D.E. (2007). The control of cognition. In W. Gray (Ed.), *Integrated models of cognitive systems*. (pp. 327 - 355). Oxford University Press.
- Kieras, D.E., & Marshall, S.P. (2006). Visual Availability and Fixation Memory in Modeling Visual Search using the EPIC Architecture. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 423-428.
- Kieras, D. & Meyer, D.E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, **12**, 391-438.
- Kveraga, K., Berryhill, M., & Hughes, H.C. (2006). Directional uncertainty in visually guided pointing. *Perceptual and Motor Skills*, **102**, 125-132.
- Kveraga, K., Boucher, L., & Hughes, H.C. (2002). Saccades operate in violation of Hick's law. *Experimental Brain Research*, **146**, 307-314.
- Lee, K.M., Keller, E.L., Heinen, S.J. (2005). Properties of saccades generated as a choice response. *Experimental Brain Research*, **162**, 278-286.
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, **104**, 3-65.
- Meyer, D. E., & Kieras, D. E. (1999). Precursor to a practical unified theory of cognition and action: Some lessons from computational modeling of human multiple-task performance. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII*. (pp. 15-88) Cambridge, MA: M.I.T. Press.
- Proctor, R.W., & Vu, K.L. (2006). *Stimulus-response compatibility principles: Data, theory, and application*. New York: CRC Press.
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, **109**, 444-474.
- Rosenbaum, D. A. (1991). *Human motor control*. New York, Academic Press.
- Rosenbaum, D.A. (2005). The Cinderella of psychology; The neglect of motor control in the science of mental life and behavior. *American Psychologist*, **50**(4), 308-317.
- Rosenbloom, P.S., & Newell, A. (1987). An integrated model of stimulus-response compatibility and practice. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol, 21, pp 1-52). New York:Academic Press.
- Teichner, W.H., & Krebs, M.J. (1974). Laws of visual choice reaction time. *Psychological Review*, **81**, 75-98.
- Welford, A.T. (1971). What is the basis of choice reaction-time? *Ergonomics*, **14**, 679-693.
- Wright, C.E., Marino, V.F., Belovsky, S.A., & Chubb, C. (2007). Visually guided, aimed movements can be unaffected by stimulus-response uncertainty. *Experimental Brain Research*, **179**, 475-496.