

# A Proposed Theory of the Neurological Limitations of Cognitive Transfer

Michael Atherton (athe0007@comcast.net)  
University of Minnesota

## Abstract

Transfer of learning has a long and controversial history in educational and psychological research, often generating diametrically opposed perspectives. In this paper I propose a novel theory of transfer based on the percepts of evolutionary psychology and cognitive neuroscience. I define conceptual and operational definitions of the constituent components involved in transfer, outline a research program, and propose possible implications for educational theory and practice.

“Cognitive psychologists, and other people who should know better, continue to advocate a philosophy of education that is totally lacking in empirical support.” (Detterman, 1996, p. 16)

“Probably the single most prevalent claim advanced by the proponents of a new paradigm is that they can solve the problems that have led the old one to crisis. When it can legitimately be made, this claim is often the most effective one possible. In the area for which it is advanced the paradigm is known to be in trouble. That trouble has repeatedly been explored, and attempts to remove it have again and again proved vain.” (Kuhn, 1970, p. 153)

If any factor is relevant to quality education it is grounding of theory with empirical evidence. Historically, education has included a number of intractable problems and ongoing debates whose discussion has been made lucid by reducing the level of analysis through neuroscience methodology. In the past five years, the conflict between Phonics and Whole Language and related issues such as “dual path” has been clarified to some extent, by brain imaging (Noble & McCandliss, 2005; Shaywitz, et al., 2004). One of the major tenets of the new field of Educational Neuroscience is that with the ability to observe neurological structure and processes we can better constrain educational and cognitive theories (Atherton, 2002). In this paper I will outline a novel theoretical approach to one of the oldest and most contentious problems in educational research: transfer of learning.

Transfer of learning relates to the process of utilizing what we learn in one situation to another. The concept of transfer has been fundamental to educational policy

in the United States (Druckman & Bjork, 1994), but it is also notorious for a replicative consistency of null results. Except for infrequent successes, as a general rule, learners are unable to spontaneously apply knowledge or procedures acquired in one context to another. This paucity of effects has resulted in a number of responses from researchers: that we simply give up (Detterman, 1996), that we reconceptualize the problem (Bransford & Schwartz 1999), or that we better define and categorize what constitutes transfer (Barnett & Ceci, 2002). I agree with the latter approach: clear definitions and topological structure can help specify what constitutes successful transfer, but I also believe that by expanding the scope of analysis we can develop an understanding of its underlying dynamics and limitations without redefining the problem.

It is interesting that little attention has been paid by educators to the underlying neural architecture that supports or constrains cognitive transfer. Other than oblique references (Ansari, & Coch, 2006; Posner & Rothbart, 2005), I am unaware of any significant investigation into the neurological basis of the transfer of learning, although there are a number of studies investigating the foundations of analogical reasoning (Bunge, Wendelken, Badre, & Wagner, 2005; Geake & Hansen, 2005; Luo et al., 2003) and neuroscientists have also explored the phenomenon of intermanual transfer (Lange, Braun, & Godde, 2006; Schulze, Luders, & Jancke, 2002). A theory of transfer based on neurocognition may prove fruitful for educational theory and practice. Such a theory might be able to explain why transfer is so difficult to achieve and under what conditions transfer might be enhanced.

## How has Transfer been Defined?

Before we can discuss the nature of transfer, we must define what it means. This is not necessarily an easy task, given that there are many different perspectives of what is meant by transfer. We can begin with the popular conception: that what we learn can (and will) be applied (or transferred) successfully to novel situations. Hoffding (1892; as cited in Royer, Mestre, & Dufresne, 2005) defined transfer by specifying the links between external events and internal representations. According to Hoffding an event  $A$  is internalized as a representation  $a$ . The activation of  $a$ , may trigger an internal response  $b$ , which in turn results in some

external action,  $B$ . You can then analyze transfer as occurring with  $A$ - $a$ - $b$ - $B$  sequences.

We can extend this formulation to say that event  $A$  occurs in a context  $C_1$  and  $B$  occurs in  $C_2$ . Transfer then is the application of representation  $a$ , or some aspect of  $a$ , in  $C_2$ . Learning then occurs in a particular context  $C_1$  is applied (successfully) in a new context  $C_2$ . By relabelling terms and making them more meaningful I will develop a framework to help understand how different researchers have approached the concept of transfer. I will relabel,  $A$  as  $E_1$  (Event 1),  $B$  as  $E_2$  (Event 2);  $a$  as  $R_s$  (representation of the stimulus); and  $b$  as  $R_r$  (representation of the response). We now have sequences of the form:  $C_1$ - $E_1$ - $R_s$ - $R_r$ - $E_2$ - $C_2$ . Much of the confusion and controversy about transfer revolves around which of these factors are included in analyses and how they are defined. For the sake of this discussion I will define these terms in the following manner. A context,  $C_i$ , can be social, environmental, a domain of knowledge, or some combination thereof. An event,  $E_i$ , can be an environmental stimulus or a particular problem within a knowledge domain. A representation,  $R_i$ , can be an internal link between stimulus and response, a hypothetical internal conceptual schema or some configuration of brain states. This Context-Event-Representation (CER) model will help illustrate how the concept and analysis of transfer has evolved historically.

### Behaviorist Research

The debate about transfer can be traced back to the early 20<sup>th</sup> century. Prior to this time the dominant educational philosophies revolved around a concept called *mental discipline* (Ravitch, 2000). The central tenet of mental discipline was that intellectual skills were domain independent and that mental exertion in difficult subjects would strengthen mental abilities and that these abilities would then transfer between domains. This view was challenged by a series of experiments by Thorndike and Woodworth (1901), who showed that, for some domains (estimation of magnitude, letter identification, and memorization), transfer was dependent on the number of common perceptual elements between tasks ( $E_1$  and  $E_2$ ). Their formulation of this phenomenon was called the theory of *identical elements* and this perspective then became the basis for Behaviorist approaches to transfer for much of the 20<sup>th</sup> century. As characteristic of Behaviorist approaches, identical elements focuses on discernable features of events  $E_1$  and  $E_2$  and not on internal representations,  $R_s$  or  $R_r$ , or the broader contextual issues involving  $C_1$  and  $C_2$ .

Not long after the development of the theory of identical elements Judd (1908, as cited in Marton, 2006) broadened the discussion by showing that the learning of a general concept facilitated the transfer of learned skills. Specifically, Judd demonstrated that teaching the concept of optical refraction helped students to hit a target submerged underwater. Essentially, Judd extended his analysis to include an aspect of  $C_1$ , by explicitly teaching how light is refracted in water.

Royer, Mestre, and Dufresne (2005) delineate the evolution of transfer theories into three generations, defining these early behaviorally oriented perspectives as a first generation transfer paradigm. Behaviorist approaches dominated much of the research on transfer until the onset of the cognitive revolution in the fifties and early sixties.

### Cognitive Research

Cognitive approaches mark the beginning of second generation approach to transfer, in part as a reaction to the limitations of the Behaviorism. Using the CER model, the change in perspective can be characterized as an inclusion of representation ( $R_i$ ) as part of the analysis of transfer. The Cognitivists' internal representations of mental processing were formulated as abstract models such as production systems, semantic networks, or schemas (Gardner, 1985; Royer, Mestre, and Dufresne, 2005).

The dominant cognitive models used to analyze transfer have been semantic networks and schemas. A semantic network comprises a set of nodes and weighted links between nodes, where nodes represent concepts or data of varying degrees of complexity. Processing occurs as nodes become activated and then they in turn differentially activate other nodes according to the weight associated with their links. Transfer is thought to occur when an existing representation (a subset of interconnected nodes, call it  $R_1$ ) is co-activated with the nodes representing the target concept or skill ( $R_2$ ). The initial activation of  $R_1$  then facilitates the creation or enhancement of  $R_2$ .

Although defined differently by authors, a schema is generally an abstract representation of knowledge that may contain conceptual, factual, and causal relations (Fuchs, Fuchs, Finelli, Courtney, & Hamlett, 2004; Gick, & Holyoak, 1983). Schemas can also be represented as having either surface structure or deep structure. Surface structure might include the concept of identical perceptual elements or other relational features of problems, while deep structures can reflect more fundamental semantic relationships. Although, the underlying operations of schemas are rarely specified

(Haskell, 2001), transfer is thought to occur when events or contexts are structurally isomorphic, i.e., when problems share similar patterns or relations. Even with its expanded conceptualization and analysis of transfer Cognitive approaches failed to find strong empirical evidence of transfer.

### Third Generation Approaches

Royer, Mestre, and Dufresne (2005), identify a new generation of perspectives. In 1999, Bransford and Schwartz published an awarding winning article in which they expand the types of contexts ( $C_i$ ) in which we analyze transfer. Rather than just considering what people transfer within experimentally limited contexts, they broadened their perspective of transfer to include social interactions and inclusion of other sources of knowledge over extended time periods. Although their proposal was not necessarily novel, other theorists had proposed similar changes in perspective, it was particularly influential. Since that time other authors have posed additional questions about how we measure transfer (Hickey & Pellegrino, 2005) and whether we should accept transfer as a natural phenomenon or simply as another theory of learning (Carraher & Schlieman, 2002).

### Types of Transfer

While the identification of different generations of transfer theory provides us an understanding of how approaches have evolved, the conceptualization and classification of different categories of transfer is also necessary to understand what is meant by transfer.

**Near and Far Transfer.** What qualifies as transfer depends on a number of factors and the general concept of transfer can be broken down into several different types. Perhaps the broadest category is that between *near* transfer and *far* transfer. The degree of similarity in perceptual similarity of events,  $E_1$  and  $E_2$ , or in the similarity of underlying structure of contexts,  $C_1$  and  $C_2$  can differentiate between near and far transfer. Transfer between events in the same conceptual domain are often labeled as near transfer. Transfer between events in different domains or different contexts are labeled far transfer. For example, if the learning of addition problems facilitated the learning of subtraction problems it would be considered near transfer, however if learning music was accompanied by improved performance in arithmetic it would be categorized as far transfer. Implicit in this distinction between near and far transfer is the idea of conceptual distance, although the term is often not clearly defined.

In 2002 Barnett and Ceci developed a taxonomy of *far* transfer in order to help identify the extent to and in what contexts we should expect to see positive effects of transfer. Their taxonomy comprises two major dimensions: *Content* (what is transferred) and *Context* (from where and where to, as well as when). Each of these dimensions comprises a number of sub-dimensions, but a detailed description is beyond the scope of this discussions. What we can do is map the major dimensions within the framework of the CER model. Essentially, Barnett and Ceci's Content dimension is equivalent to various aspects of representation ( $R_i$ ) and the Context dimension captures aspects of events ( $E_i$ ) or contexts ( $C_i$ ). The *when* of their taxonomy places an emphasis on the importance of time between events ( $E_1 \leftrightarrow E_2$ ) or contexts ( $C_1 \leftrightarrow C_2$ ).

**Vertical and Lateral Transfer.** In 1965, Gagné also specified a distinction between what he called *vertical* and *lateral* transfer. Vertical transfer generally involves the benefits of the application of learning within a single domain to another application of greater complexity within the same domain, for example the transfer of learning between addition and multiplication. In terms of the CER model transfer would occur with the same conceptual domain within  $C_i$  or particular sets of events in  $E_i$ . Gagné broadly characterized lateral transfer as the generalization of learning across domains or events at a similar level of complexity (Gagné, 1965).

**Literal and Figural Transfer.** In the late Seventies, Royer (1979) identified a structural difference in transfer by differentiating between *literal* and *figural* transfer. Literal transfer uses intact knowledge learned in an original context and applied in a novel context. Figural transfer applies some aspects of the learning in the first context to a novel context. Since these types of transfer refer generally to the structure of knowledge, in the CER model they would be most closely related to how representations ( $R_x$ ) are applied between  $E_1$  and  $E_2$  or  $C_1$  and  $C_2$ .

**Low-road and High-road Transfer.** In the Eighties Salomon and Perkins (1984;1989) proposed a distinction between what they termed *low-road* and *high-road* transfer. Low-road transfer refers to learning that is characterized by extensive and varied practice and is applied in new contexts automatically. High-road transfer involves awareness and intentional application of previous learning in a novel context. While low-road transfer seems by its definition to be classified as near transfer, high-road transfer would not necessarily encompass all instances of far transfer; that is, far transfer can occur without intention or conscious awareness.

**Positive and Negative Transfer.** Perkins and Salomon (1992) also make a distinction between what they term *positive* and *negative* transfer. Positive transfer occurs when previous learning enhances performance in a new context ( $C_2$ ). Negative transfer occurs when previous learning inhibits learning in a new context. Negative transfer would appear to be very similar to the more traditional psychological construct of proactive interference. In terms of the CER model relate to the valence of transfer effects in  $C_2$ .

### The Problem with Transfer

Although the problem with transfer has traditionally been thought of as a scarcity of confirmatory evidence, possibly the true nature of the problem is really the a priori assumption of its existence. Most researchers and educators do not think of transfer as a theoretical construct of learning, rather its existence is accepted as a matter of faith. So what we have had is a theoretical construct in search of effects and these effects have been notoriously absent, inconsistent, or unreliable in situations in which they would normally be expected (Detterman, 1996). Several authors have attempted to solve this problem by broadening the scope of analysis and expanding the contexts ( $C_i$ ) in which we define and study transfer, but what if the problem is more fundamental? What if our assumption of the ubiquity of transfer is too sanguine? Suppose that rather than beginning with an assumption of its existence, we consider how transfer might have evolved from an evolutionary perspective? This will be a starting point for developing an alternative theoretical framework for transfer. This avenue is not evidence, instead it is intended to lay a plausible framework which poses questions and provides directions for future research.

### Transfer in an Evolutionary Context

While it has been difficult to demonstrate the effects of transfer, there is another structurally similar phenomenon that has been extensively studied and documented for more than a century (Ghirlanda & Enquist, 2003). *Generalization* is the tendency to respond to a range of stimuli, after learning a particular one. The primary differences between generalization and transfer are that in transfer there is greater degree of complexity in contexts and mental representations. In generalization studies normally only one stimuli is presented (although it might be varied along several dimensions) and other contextual factors are controlled, while in transfer research what is presented is often a complex set of interrelated stimuli (knowledge) and most often their presentation occurs in rich contextual environments. We can think of the study of

generalization as being limited to  $E_1 - E_2$  sequences, whereas most studies of transfer are also interested in contexts ( $C_i$ ), representations ( $R_i$ ), or both. While these two phenomena differ in complexity we can relate them metaphorically to form hypotheses about similarities of their function.

In generalization, responses to a stimulus are recorded along two dimensions and often take on a normal (Gaussian) or other symmetric distribution with a certain variance. The amount of variance of these curves can differ based on a number of different factors. Suppose we were recording the startle responses of an antelope in the field to a range of auditory stimuli. Depending on both the evolutionary and behavioral history of the animal these responses might afford different levels of survival value. Cues associated with predators would have high survival value, so we might expect that the variance of responses to be greater, however if the animal reacted to too wide a range of stimuli and expended too much energy doing so long term survival value would be reduced. So, it would seem that, depending on a set of interacting variables, there would be an optimal degree of variance which would provide the greatest probability of survival. In effect, the task of the animal is to find an optimal variance for the response distribution. This variance will, in some sense, be evolutionary bounded by the animal's genotype. Although researchers in evolutionary psychology consider this a plausible conjecture, a detailed theory has not yet been formalized (K. Cheng, personal communication, September 4, 2006).

Given that it is possible that generalization is evolutionarily constrained, it would seem equally plausible that transfer is similarly limited. While applying learning in contexts in which it has application is an adaptive, applying it in contexts in which it is not might often be maladaptive. When we think about transfer in education we assume that knowledge will be applied in contexts in which it is useful, but to know when knowledge is useful and not dangerous, or at least improvident, would require an all seeing oracle to predict which particular knowledge is germane. The problem with novel contexts is that, by definition, they are different from anything we have seen or known before. While an organism might develop general heuristics for dealing with future events, it is impossible to determine specifically what learning will transfer successfully. Thus, at least in relation to far transfer in novel contexts, our expectation for a high degree of transfer may be unreasonable. These arguments suggest the following hypothesis.

**Hypothesis 1:** Cognitive transfer is constrained by the uncertainty of evolutionary adaptation.

While this hypothesis might not be easy to prove, it is testable. The evidence to support it could be generated by genetic studies using correlational techniques, or experimentally with transgenic animals, or possibly with computer models of evolutionary dynamics. Regardless, this hypothesis is only intended to suggest the plausibility of limitations placed on transfer by genotype and neuroanatomy.

### Neuroanatomy

Even if you do not accept the conjecture that transfer is limited by evolutionary demands, it is still possible that transfer could be limited as an indirect consequence of the evolution of the human brain. The brain is not fully connected, at least in terms of being equipotent. Specialized perceptual, cognitive, and emotional functions are localized to particular regions and these local structures are not equally interconnected; well defined pathways exist between different regions. Then, is it somehow possible that this neural architecture controls or limits how transfer occurs? I believe it has to.

It is interesting that although mental models and mental representations have played a role in the study of transfer, neuroscience has not. Robert Haskell states “Currently, there exists no neurological theory of transfer...” (2001, p. xix). Haskell has identified a diverse set of perspectives of how evolution and brain structure might support transfer and has also outlined a provisional explanation of the neurological basis of transfer, however Haskell’s explanation is different than the one is presented in this paper. Haskell assumes that there is a neurological foundation that has evolved to support transfer. This approach again assumes that transfer is a well defined, delineated, and ubiquitous construct, rather than an ethereal cognitive phenomenon.

Haskell suggests that neural architectures function to filter data and detect similarities and invariants. These invariants or equivalences allow the detection of small nuances which make possible “spontaneous and significant transfer” (2001, p. 201). Haskell cites a number of concepts to help support this conclusion, including brain lateralization, neural networks, and neural modules. However, details of how these mechanisms might function to produce transfer are not specified. Haskell suggests that a rich database of knowledge that is highly interrelated and interconnected facilitates transfer and that “deep” processing and “deep” learning can enhance transfer. The problem with this argument is the same as that of

with *levels of processing* where “deep” and “level” are not clearly defined and we are left without a concise conceptualization of the underlying mechanisms. In contrast, the goal of this paper is to provide a theoretical framework to initiate research into the neurocognitive mechanisms and limitations of transfer.

### The Neuroanatomy of Cognitive Transfer

As cited earlier, it’s not that we don’t have any empirical evidence of transfer, it is just that the evidence we do have does not match well with our expectations for educational outcomes. The phenomenon of generalization is ubiquitous and consistent within certain constraints, but is too restricted to be thought of as transfer in educational contexts. Gagné’s concept of *vertical* transfer is also a universal characteristic of instruction: previous learning facilitates future learning in the same context. There are also well documented examples of transfer between specific domains such as math and music. What is it about these examples that provide a basis for transfer and what is lacking in the situations in which we expect to see transfer but fail to? The gist of this paper is that the problem with transfer lays primarily within the limitations of neural representations ( $R_i$ ), not context or pedagogy.

I’ll begin this section by providing a logical and conceptual argument to support my next hypothesis. If for the moment we lay aside any dualist conceptions of the division between mind and brain and instead take a material or physicalist perspective, then cognitive processing necessarily requires some type of activity in the brain. And given considerable empirical evidence, it is reasonable to assume that this activity is distributed between different areas primarily by axonal and synaptic connections between neurons. Therefore, if we expect one representation of knowledge to affect another it would seem plausible that the regions containing these representations be connected in some way and secondly, that they be active within some common timeframe. We might then infer that these conditions may place some limit on cognitive processing and transfer and that the following hypothesis is reasonable:

**Hypothesis 2:** The probability of transfer is a function of the co-activation of brain regions subserving previous and current contexts and the degree of interconnectivity between these regions.

In other words, if brain regions are not connected we should not expect to see any transfer between the cognitive functions they support and secondly, if these regions are connected but never activated

simultaneously within a timeframe which permits one to affect the other we should not expect to see transfer. Which regions are connected and the degree to which they are connected is still being researched, but enough is known to make this a verifiable hypothesis. The weighting of these two factors, connectivity and co-activation, and their effect on transfer will have to be established empirically, but what is proposed here is that connectivity and co-activation are necessary conditions, although the complete picture is sure to be more complex.

Theoretical verification is often constrained by the lack of operational definitions, so it is important to make clear what is meant by activation and connectivity. Here activation is defined as some type of neural processing above baseline. Activation can be measured by several inter-correlated factors such as neural spikes or cellular metabolic processes. Co-activation can be established with the statistical methods associated with imaging techniques such as magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), or positron emission tomography (PET). It is likely that often the activation accompanying transfer will be below the threshold detectable by current brain imaging technologies and that some may involve very small populations of neurons. However, the general phenomenon of activation accompanying transfer should be prevalent enough to be statically significant. There are several ways to measure the connectivity between areas in the brain. Although the most accurate might be to count synaptic connections it is not practical, so counting axioms or the density of axonal tracks might be more reasonable. Diffusion tensor imaging (DTI) provides a means to measure the degree of connectivity between regions. MRI and DTI can be performed using the same equipment and provides a way to measure both functional correlation of activity and the degree of connectivity.

### **Empirical Evidence**

While there isn't direct experimental evidence of the validity of Hypothesis 2, there are studies that suggest that it might be. As mentioned previously there are measurable transfer effects in some domains such as motor skill and bilingual instruction. While a comprehensive review is beyond the scope of this paper, we can consider the evidence between several domains; in particular transfer between learning in music and other cognitive skills. It is an unfortunate that one study relating transfer in music has become the archetypical example of the overgeneralization of neuroscience in education and development (Bangerter & Heath, 2004), because it tends to diminish the

importance of other studies. The original study on the "Mozart Effect" (Rauscher, Shaw, & Ky, 1993) dealt only with short term listening to classical music and related it to small performance changes in spatial reasoning (many questions have been raised about the adequacy of controls; Chabris, C. F, 1999), but music involves a number of neurological processes: fine motor skills; auditory processing; mathematic ability in counting, fractions, and repetitive patterns; visual spatial ability in reading and writing musical notation; and executive processing in the parsing of musical compositions.

One of the first questions that might be asked about the neuroscience of music is the difference between nature and nurture. While there is no clear evidence that individuals are genetically differentiated by structural neural-anatomical differences (Norton, Winner, Cronin, Overy, Lee, & Schlaug, 2005), it is clear that there are structural differences between adult musicians and non-musicians and there also appears to be training and developmental effects on neurological structures (Schlaug, Norton, Overy, & Winner, 2005; Schmithorst & Wilke, 2002). Thus, these anatomical differences might fit hypotheses presented here by inferring the possibility of more highly connected brain regions.

While the existence of structural difference between musicians and non-musicians might provide the possibility of connectivity, it is necessary to show a relationship between different cognitive domains that accompany musical training. Transfer effects have been demonstrated between music ability and elements of language processing (Anvari, Trainor, Woodside, & Levy, 2002; Gromko, 2005; Slevc & Miyake, 2006), now recent research is beginning to show how language and music processing occur in overlapping regions in the brain (Brown, Martinez, & Parsons, 2006; Koelsch, Kasper, Sammler, Schulze, Gunter, & Friederici, 2004). Although research results do not illustrate as consistent a relationship between mathematics and music, a brain imaging study using fMRI suggest a correlation between math proficiency and training in music that may be linked to improved working memory performance (Schmithorst & Holland, 2004). What is needed at this point are studies which specifically look for overlapping activation when transfer is consistently observed between different domains.

### **Educational Implications**

If it can be shown that Hypothesis 2 is true, there are several possible implications for education. Most importantly, we should not expect to see equal amounts of transfer in all cognitive contexts and domains. In terms of enhancing the possibility of transfer, we should

consider insuring that there are connections between different domains and that we co-activate representations in memory between which we hope to enhance transfer. Possibly, as Haskell (2001) and others have suggested, broad knowledge across multiple domains may increase the likelihood of transfer and as suggested here, this facilitation might occur because of an increase in neural connections which might then increase the probability of co-activation between regions. Priming, hints, advanced organizers, and guided discovery might have advantages over unstructured discovery methods which leave connections and co-activation to their level of natural occurrence.

A theory of transfer based on evidence generated by research into the accuracy of Hypothesis 2 might also provide explanatory unity to the now divergent types of transfer as outlined above. It is tempting to speculate that one difference between near and far transfer might be reflected in actual distance in terms of neural connections or the bandwidth of the channel in relation to the density of white matter or numbers of synapses. Vertical transfer could be described in terms of adding to the complexity of existing connections, whereas lateral transfer by establishing new connections between separate regions. Literal and figural transfer could involve differences between which processing occurs primarily within areas subserving existing representations compared to forming connections between areas that had not previously been highly connected. Low-road and High-road transfer may be differentiated by differences in executive processing in frontal regions of the brain and low-road transfer by automated processing in posterior brain regions (see Atherton, 2001). Positive and negative transfer might be related to cognitive processes such as proactive and retroactive interference and neurological or neural processes such as competition between excitation and inhibition.

### Conclusion

The intent of this paper is to attempt to provide a plausible explanation for a problem that has long stymied educational researchers. It is hoped that the proposed hypotheses will simulate debate and research and if this explanation is verified empirically that it will show how it is possible for Educational Neuroscience can make a significant contribution to the theoretical foundations of education. If we can better ground theory with the reality of neurocognitive constraints we may be able to move beyond debates that restrained progress in education.

### References

- Atherton, M. (2002, August). A neurocognitive model for students and educators. Paper presented at the annual meeting of the Cognitive Science Society, Fairfax, VA
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, 83, 111-130.
- Bangerter, A., & Heath, C. (2004). The Mozart effect: Tracking the evolution of a scientific legend. *British Journal of Social Psychology*, 43, 605-623.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612-637.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61-100.
- Brown, S., Martinez, M. J., Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European Journal of Neurosciences*, 23, 2791-2803.
- Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex*, 15, 239-249.
- Carraher, D., & Schlieman, A. D. (2002). The transfer dilemma. *The Journal of the Learning Sciences*, 11, 1-24.
- Chabris, C. F. (1999). Prelude or requiem for the 'Mozart effect'? *Nature*, 400, 826-827.
- Detterman, D. K. (1996). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R.J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1-24). Norwood, NJ: Ablex.
- Druckman, D., & Bjork, R. A. (1994). *Learning, remembering, believing: Enhancing human performance*. Washington, DC: National Academy Press.
- Fuchs, L. S., Fuchs, D., Finelli, R., Courtney, S. J., & Hamlett, C. L. (2004). Expanding schema-based transfer instruction to help third graders solve real-life mathematical problems. *American Educational Research Journal*, 41, 419-445.
- Gagné, R. M. (1965). *The conditions of learning*. New York: Holt Rinehart & Winston.
- Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books

- Geake, J. G., & Hansen, P. C. (2005). Neural correlates of intelligence as revealed by fMRI of fluid analogies. *Neuroimage*, 26, 555-564.
- Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, 66, 15-36.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research in Music Education*, 53, 199-209.
- Haskell, R. E. (2001). *Transfer of learning: Cognition, instruction and reasoning*. New York: Academic Press.
- Hickey, D. T., & Pellegrino, J. W. (2005). Theory, level, and function: Three dimensions for understanding transfer and student assessment. In J. P. Mestre (Ed.), *Transfer of learning: From a modern multidisciplinary perspective*. Greenwich, CT: Information Age Publishing.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, 7, 302-307.
- Kuhn T. S. (1970). *The structure of scientific revolutions*, (2nd. ed.). Chicago: Chicago University Press.
- Lange, K., Braun, C., & Godde, B. (2006). Coordinate processing during the left-to-right hand transfer investigated by EEG. *Experimental Brain Research*, 168, 547-556.
- Luo, Q., Perry, C., Peng, D. L., Jin, Z., Xu, D., Ding, G. S., & Xu, S. Y. (2003). The neural substrate of analogical reasoning: an fMRI study. *Cognitive Brain Research*, 3, 527-534.
- Marton, F. (2006). Sameness and difference in transfer. *The Journal of the Learning Sciences*, 15, 499-535.
- Noble, K. G., & McCandliss, B. D. (2005). Reading development and impairment: Behavioral, social, and neurobiological factors. *Journal of Developmental and Behavioral Pediatrics*, 26, 370-378.
- Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., & Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain and Cognition*, 57, 124-134.
- Perkins, D. N., & Salomon, G. (1992). *Transfer of learning*. *International encyclopedia of education* (2nd ed.). Oxford, UK: Pergamon Press.
- Posner, M. I., & Rothbart, M. K. (2005). Influencing brain networks: implications for education. *Trends in Cognitive Sciences*, 9, 99-103.
- Rauscher, F. H., Shaw, G. L., & Ky, K.N. (1993). Music and spatial task-performance. *Nature*, 365, 611-611.
- Ravitch, D. (2000). *Left back: A century of failed school reforms*. New York: Simon & Schuster.
- Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist*, 14, 53-69.
- Royer, J. M., Mestre, J. P., & Dufresne, R. J. (2005). Introduction: Framing the transfer problem. In J. P. Mestre (Ed.), *Transfer of learning: From a modern multidisciplinary perspective*. Greenwich, CT: Information Age Publishing.
- Salomon, G., & Perkins, D. N. (1984, August). *Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon*. Paper presented at the Conference on Thinking, Harvard Graduate School of Education, Cambridge, MA.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, 24(2), 113-142.
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Neurosciences and Music II: From Perception to Performance*. *Annals of the New York Academy of Sciences*, 1060, 219-230.
- Schmithorst, V. J., & Holland, S. K. (2004). The effect of musical training on the neural correlates of math processing: a functional magnetic resonance imaging study in humans. *Neuroscience Letters*, 354, 193-196.
- Schmithorst, V. J., & Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neuroscience Letters*, 321, 57-60.
- Schulze, K., Luders, E., & Jancke, L. (2002). Intermanual transfer in a simple motor task. *Cortex*, 38, 805-815.
- Shaywitz, B.A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., Holahan, J. M., Marchione, K. E., Fletcher, J. M., Lyon, G. R., & Gore, J. C. (2004) Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, 55, 926-933.
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, 17, 675-681.
- Thorndike, E. L. & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions (I). *Psychological Review*, 8, 247-261.

