# Intuitive conceptions of probability and the development of basic math skills

Gary L. Brase
Sherri Martinie
Carlos Castillo-Garsow
Kansas State University

*The idea of probabilities has been described as a "Janus-faced" concept, which can be thought of either in terms of frequencies or in terms of subjective confidence. This dualism contributes to debates about the nature of human rationality, and therefore the pedagogical assumptions and goals of education. For this reason, the present chapter explores the evidence regarding how quantitative information is intuitively understood in the human mind over the course of elementary school education. Are particular interpretations of probability equally weighted or does one interpretation predominate as mathematical concepts are being acquired? We find multiple, converging lines of evidence that indicate a frequency interpretation of probabilistic information is developmentally primary and privileged. This has implications for mathematics education, even before the introduction of actual probabilities, in areas such as learning fractions and decimals. Educational practices should work to bootstrap from these privileged representations (rather than fight them) and built towards a more inclusive and comprehensive model of probability knowledge. We conclude that a fundamental issue is not just whether students think about probabilities as a frequentist or as a subjectivist, but rather how they recognize when to be one versus the other.*

## The interpretation of probabilities

Mathematics may be the queen of the sciences, but this queen has a dirty little secret. Once in a while, she sneaks out and cavorts with the common people; arguing, gambling, and getting into fights. When she does this it is called *probabilities*. Far from the regal courtyards, the mathematics of probabilities developed as a way to evaluate what were good or bad gambles (Gigerenzer, Swijtink, Porter, Daston, Beatty, & Kruger, 1989). She engaged in ferocious debates about what she was and how she should behave (Gigerenzer, 1993). She behaves, in other words, in a very un-queenly manner.

In all seriousness, the concept of probabilities in mathematics really does lead a metaphorical double life. On the one hand, there is a conception of probabilities that involves using past events to statistically predict future outcomes. For example, a "30% chance of rain" is based on a model of past weather events (It has rained on 3 out of the 10 previous days similar to this one). We will call this the *frequentist* concept of probabilities because it is, at its core, based on frequencies of past events. On the other hand, there is a conception of probabilities that involves degrees of belief. For example, a "1% chance that the earth will be destroyed in 10 years" is based on a person's subjective beliefs, not a history of past events. We will call this the *subjectivist* concept of probabilities because it is, at its core, based on subjective belief states. Both concepts of probabilities can be legitimately claimed as correct, which led Hacking (1975) to refer to probability as a "Janus-faced" concept. These two conceptions of probability are sometimes easily distinguished, but at other times the distinction can be very subtle. For instance, a frequentist could claim that the probability of a six-sided die showing an even number is .5, based on a history of past rolls (i.e., of the prior rolls, half of them came up even). On the other hand, a subjectivist could claim the same answer by counting the sides, dividing by the number of sides with even numbers, assuming the die is fair, and predicting (without ever rolling the die once) that the probability of an even roll is .5.

The polysemous state of the probability concept is not constrained as just a mathematical annoyance. It also connects with disagreements about the nature of human rationality. After all, having two different conceptions of probability can lead to two people having different answers to the same question yet both believing they are rational and correct. This type of discrepancy fuels a debate about the nature of rationality in psychology research (see also the chapter by Ejersbo & Leron).

On one side of this debate about rationality is the *ecological rationality* view, which hews more closely to the frequentist conception of probability. This view is a more recent approach in psychology, although it has an intellectual history as long or longer than the alternative view (Gigerenzer et al., 1989). The ecological rationality view stresses the fit between the structure of the natural environment, which includes a frequentist focus on encountered objects, events, and locations, and the structure of the human mind. The emphasis here is on the rationality of making the best choice *under ecologically realistic circumstances.* Specifically, what is the best choice to make given the time and computational constraints of a situation and taking advantage of the inherent statistical properties of the natural world within which the choice is being made? (see Gigerenzer, Todd, & Group, 1999; also see chapters by Meder & Gigerenzer and by Martignon).

The other side of this debate about rationality is the *heuristics and biases* view, which is not as strongly affiliated with a particular view of probability but leans more towards the subjectivist conception. This view is more traditional in psychology, and in many ways is the current default approach, with seminal works by Tversky and Kahneman that go back to the 1970s (Kahneman & Tversky, 1973; Tversky & Kahneman, 1971, 1973, 1983). The heuristics and biases view emphasizes normative criteria for rationality, such that there is *a* correct answer and that answer is often (but not always) based on the subjectivist conception of probability. More recently, this heuristics and biases view has been shifting towards a dual process approach, in which people are thought to employ either deliberative styles of thought or intuitive styles of thought when reasoning (Kahneman, 2011).

Why is this situation important for mathematics education? Because any approach to education must take into account what the learner is bringing with them to the educational endeavor. In very simple terms, the ecological rationality view describes people as intuitive frequentists, whereas the heuristics and biases view describes people as somewhat muddled subjectivists. Effective teaching and learning needs to take into account what the learner already knows (and hence does not need to be taught), what needs to be taught, and what should not be taught (either because it is not the appropriate point in time or because it is inherently confusing or wrong). An accurate employment of these consideration depends on what the nature of the human mind is regarding mathematical thinking.

**Overview**

The bulk of this chapter addresses a basic question, derived from the ecological rationality approach as contrasted with the heuristics and biases approach: *Are frequentist interpretations of probability privileged in the human mind?* Whereas the heuristics and biases approach has been very well documented as an approach to the learning and teaching of probabilities (e.g., Jones & Thornton, 2005), this more recent ecological rationality approach is less well understood and appreciated. The basic question of this chapter is, in many ways, the central question regarding how these approaches hold implications for mathematics education. The heuristics and biases approach, in contrast to the ecological rationality approach, contends that both frequentist and subjectivists interpretations of probability are equally weighted as representations that the human mind can use for understanding probabilities. And, in fact, the subjectivist interpretations is often considered superior and thus to be promoted directly. These foundational assumptions color the pedagogical background assumptions, educational approaches, and expected learning outcomes of mathematics education. This chapter will therefore review the evidence that the ecological rationality approach, with its position that frequentist representations are cognitively privileged, is at least as valid as the heuristics and biases position.

We will use the method of converging operations to assess the evidence regarding this question (Garner, Hake, & Eriksen, 1956; Sternberg & Grigorenko, 2001; Sternberg & Grigorenko, 2001). In short, the following sections will look at different areas of research and assess what the balance of evidence within each area says about our question. (See also Schmitt and Pilcher (2004) for a similar approach). Specifically, we will be looking at evidence from evolutionary biology, information theory, phylogentic research, cognitive psychology, developmental psychology, cross-cultural research, and neuroscience.

**What is the Evolutionary "history" of Quantitative Information?**

Information from the natural environment is primarily frequencies – numbers of objects, occurrences of events, visits to locations – which can be tabulated in a binary fashion that produces basic, whole-number frequencies. These sequentially encountered frequencies can be easily and efficiently organized using a natural sampling framework. (Natural sampling structure refers to the non-normalized subset structure that can result from such a tracking and storage system; Kleiter, 1994) Furthermore, this would have included (and still does include) information about item frequencies of significant importance to survival and reproduction: patterns of food distribution, availability of potential mates, of coalition partners and of rivals, frequency of weather events, and frequency of events as indications of time passage. These and many other existing, potentially useful classes of information in the world constituted evolutionary selection pressures to attend to, remember, and think about (De Cruz, 2006; C R Gallistel, Gelman, & Cordes, 2006). Indeed, the entire field of optimal foraging theory in biology is predicated on the assumption that animals have been selected to track quantities of food and space (Stephens & Krebs, 1986; see section on comparative evidence). This idea was succinctly expressed by Beran (2008, p. 2):

> Any creature that can tell the difference between a tree with 10 pieces of fruit from another with only six pieces, or between two predators and three on the horizon, has a better chance of surviving and reproducing.

Although our modern, technologically advanced environment contains a great deal of quantitative information expressed in ways other than basic frequencies (e.g., a 60% chance of rain today), well over 99% of human evolutionary history did not include such numerical expressions. Over evolutionary history, information about the environment came largely from first-hand experience (even information provided by others would have been restricted to fairly recent events and those people with whom one actually lives). Experientially, single events either happen or they don't— either it will rain today or it will not. So one can observe that it rained on 6 out of the last 10 days with cold winds and dark clouds, and one can have a subjective confidence based on that history, but one cannot *observe* a 60% subjective confidence. Over evolutionary history, as individuals were able to observe the frequency with which events occur, this information was potentially available to be utilized in decision-making. Thus, if humans have adaptations for inductive reasoning, one might expect them to include procedures designed to take advantage of the existent frequency information in the environment (Hasher & Zacks, 1979; Hintzman & Stern, 1978).

What about the apparently non-frequency outputs that people generate, such as a subjective confidence that it will rain *today*? This is not a real dilemma, as it is about the end product of calculations that can easily be based on frequencies up until this point. By analogy, think about opening an image file on your computer. The fact that your computer can produce a picture does not mean that the representational format of that image inside the computer is also a picture. (In fact, it is not.)

**What are the Computational Properties of Different Quantitative Representations?**

Given that information, even if initially received in the format of frequencies, can be converted into other formats, it is necessary to assess the costs and benefits of different quantitative formats for mental representation. Different numerical representation formats have slightly different mathematical properties, each of which entail certain advantages and disadvantages. What are the computational advantages and disadvantages of frequencies and natural sampling relative to other formats?

Basic frequencies are in many ways the most foundational of numerical formats, both historically and because all other formats are computationally derivable from them. Unlike normalized formats (percentages, probabilities, and ratios), non-normalized frequencies preserve information about reference classes. For instance, *2 out of 3* and *1098 out of 1647* both produce the same normalized numbers (66.7%, .667, 2:5, 2:1), but the first frequency is a much smaller reference class than the later. This can be important in making judgments and decisions, because the size of the reference class provides unambiguous indications of the reliability and stability of the information. Following our above example, the addition of one more number can change *2 out of 3* into *3 out of 4* (drastically changing the nature of the information) whereas the addition of one more number changes *1098 out of 1647* very little (i.e., it is much more reliable and stable information).

Additionally, the storage of information as non-normalized frequencies is important for preserving a high degree of flexibility of categorization and subcategorization. For example, one may have seen 100 people fall sick with an unknown illness, and also observed a half-dozen possible diagnostic symptoms (as well as, of course, many non-ill people and their characteristics). Only with the original frequencies is it possible to reconfigure the organizational relationships between illness and symptoms to discover diagnostic patterns. (For example, the 100 sick people could be organized by one symptom, say 25 had a fever, and then reorganized to evaluate a different symptom, say 80 had swollen glands.) Information that has been already normalized to arbitrary reference classes (e.g., 100 for percentages, 1 for single-event probabilities, etc.) would have to be dismantled back into frequencies to be reconfigured in new ways (e.g., given that 25% of sick people had a fever, the ability to reorganize that information based on a different symptom is lost unless it is reconstituted to the original reference class)

Finally, only with frequencies can new information be easily and automatically incorporated with ongoing experiences. Once again, normalized numbers obscure the size of the reference class and make it impossible to tell how much (or how little) a single new observation will change the nature of the data. Frequencies are simply easier to work with because they conserve information about the base rates (Kleiter, 1994).

There most definitely are certain situations in which naturally sampled frequencies appear to be unable accomplish particular tasks. Two such situations that are often raised are: a) judgments about novel, one-time events, and b) determining if observations are statistically independent of one another (Over & Green, 2001). With respect to one-time events, one possibility is that people look to similar, though not identical, events for guidance in these situations[1]; that puts these cases back into the realm in which natural frequencies can be utilized. An alternative possibility is that, because singular events in one person's lifetime may have occurred repeatedly across the evolutionary history of a species, the mind might contain evolved behavioral predispositions regarding that type of event (assuming the event had inclusive fitness consequences). This would not be based on natural frequencies, but it also is clearly not free of evolutionary or ecologically rational processes either. Finally, it is possible that – as some people do find when faced with a novel, one-time event— one simply does not have clear or consistent guidance about what to do. The second situation (of determining statistical independence of observations) is, indeed, easier with normalized numbers. These normalized numbers, however, are quite obtainable as derived from frequencies.

---

[1] A related option is that people may be able to construct frequentist (or quasi-frequentist?) probabilities by counting hypothetical possibilities in the sample space: using theoretical frequencies, rather than actual frequencies. In other words, imagining the possible outcomes in a way that still respects constraints of the real world (see Shepard, 1984 for a similar idea within the field of perception). For example, a judgment about how likely it is to roll 10 or greater than 10 on the sum of two 4 sided die and one 8 sided die is probably a novel one-time event, but one can figure it without too much effort. (either by using subjectivist methods of expected value, or by counting in the hypothetical frequencies within the relevant sample space). The expected value solution would be that I expect to roll an average of 9.5 on the sum of the dice, so rolling a 10 or greater should be 50%. Investigating how people, in fact, deal with situations such as this is important but not adequately studied at this time. The most immediate point for our purposes here is that these situations posed by Over and Green (2001) are far from conclusive.

There is no *a priori* reason why these insights cannot be derived from the frequency information contained within naturally sampled frequencies, following an algorithmic transformation. There is, however, an insurmountable problem if only the percentage information is available. Because information about the reference class size is gone, one cannot say whether or not the information (e.g., 80%) is based on 800 out of 1000 or based on 4 out of 5.

**What about other Animals?**

Many of the initial considerations of an ecological rationality approach apply not only to humans but to other animals as well (e.g., the fit between the structure of the natural environment and the structure of the [non-human] mind; time and computational constraints of a situation; the inherent statistical properties of the natural world within which the choice is being made). It is therefore important to look at phylogenetic evidence for how quantities are represented in non-human animals. This includes information from comparative psychology, primatology, physical anthropology, and paleontology.

A number of evolutionary biologists have investigated behaviors that involve judgment under uncertainty in nonhuman animals in order to test various mathematical models from optimal foraging theory (e.g., Stephens & Krebs, 1986). Their experiments have demonstrated that a wide array of animals – even ones with truly minuscule nervous systems, such as bumblebees – make judgments under uncertainty during foraging that manifest exactly the kind of well-calibrated statistical induction that humans have seemingly struggled with on paper-and-pencil tasks (e.g., Gallistel, 1990; Real, 1991; Real & Caraco, 1986; Staddon, 1988). How can this be?

At a basic level, a wide array of non-human animals (including salamanders, rats, various types of birds, dolphins, monkeys, and apes; see Beran, 2008 for an overview) are sensitive to the quantitative properties of various natural events. These are, to be sure, experienced frequencies of objects, events, and locations, rather than explicit numerical notations (most animals do not understand written numbers). But even in the realm of explicit numerical notation, non-human primates have been shown to be capable of simple mathematic operations (e.g., matching, addition, and subtraction of small quantities) using abstract stimuli such as the number of dots on a computer screen (Beran & Rumbaugh, 2001; Boysen & Berntson, 1989; Brannon & Terrace, 2000, 2002; Jessica F Cantlon & Brannon, 2007; Inoue & Matsuzawa, 2007; Matsuzawa, 1985; Tomonaga & Matsuzawa, 2002).

What is one to make of the contradiction between these impressive non-human animal performances and the often seemingly error-prone human judgment and decision making abilities? A key to resolving this apparent paradox is to realize that these other animals (including even bumblebees) made great judgments under uncertainty because they were tested under conditions that were ecologically valid for their species (Tooby & Cosmides, 1992). The ecological rationality approach simply claims that the same should hold true for humans.

**What is the Cognitive Psychology of Quantitative Information?**

There are several areas of research within Psychology that provide relevant evidence about how people encode, represent, and use quantitative information. This is also, perhaps not coincidentally, the area in which the disagreements between ecological rationality and heuristics and biases views get played out most often and most assiduously. This section is divided into parts, discussing the research on how well people can track the occurrence of frequencies, how people evaluate different numerical formats, how people perform on complex statistical tasks when given information in different numerical formats, and how adding different pictorial aids can influence statistical reasoning.

### Frequency Tracking

Although frequency judgments in naturalistic studies of autobiographical memory have been inaccurate in many studies, laboratory studies that more carefully control for the multitude of potentially confounding factors endemic to naturalistic studies have found remarkably accurate frequency judgments (for a review, see Hasher & Zacks, 2002). Even studies using classic tasks that are supposed to demonstrate the fallibility of frequency estimation abilities (e.g, by people having to resort to the availability heuristic; Tversky & Kahneman, 1973), have been found to nevertheless maintain accurate relative frequencies (Sedlmeier, Hertwig, & Gigerenzer, 1998). This work has also led to a more subtle point that the ecological rationality perspective does not predict perfect or optimal frequency tracking in terms of explicit numbers, but rather that frequency tracking should be expected to *satisfice*: to be as accurate as necessary under normal ecological circumstances (i.e., realistic item and computational limits) so as to produce good judgments and decisions; see Simon, 1956).

### Evaluating Different Numerical Formats

Another way to evaluate and understand quantitative information is to directly ask them. Brase (2002) found that when people are asked to evaluate the clarity and understandability of different statistical statements, frequencies (both simple frequencies and relative frequency percentages) are seen as clearer and easier to understand than single-event probabilities. This research also found that simple statements of large-scale frequencies (e.g., millions or billions) lead to systematic distortions in the persuasive impact on both attitudes and potential behaviors, both for very low base rates (e.g., expressing a 1% base rate in terms of absolute numbers of the world populations, which leads to greater relative impact than other formats) and for very high base rates (e.g., expressing a change from 6 billion to 7 billion, which leads to less relative impact than other formats).

Wang (Wang, 1996a, 1996b; Wang & Johnston, 1995; Wang, Simons, & Brédart, 2001) similarly found that using large-scale frequency contexts led to systematic differences in behavior. In this case, when people were asked to make decisions about a small-group or family-sized populations (fewer than 100 people) the traditionally observed framing effects of decision malleability (i.e., shifting towards riskier outcomes when framed as gains) disappeared (Tversky & Kahneman, 1981; Wang, 1996a, 1996b; Wang & Johnston, 1995). The explanation for this pattern of sudden consistency stemmed from the ecological rationality insight that these smaller population sizes are on scales of magnitude with which humans have directly and recurrently dealt over their personal history.

Statistical Reasoning with Different Numerical Formats

A classic area of research on statistical reasoning is that of Bayesian inference (determining the posterior probability of an event, given some new information to be combined with an initial base rate; for example, how likely a patient has a disease given a positive test result), with the classical finding that people very rarely perform well (i.e., reach the correct Bayesian inference; e.g., Casscells, Schoenberger, & Grayboys, 1978). This classical finding, however, was based on tasks in which the information was presented in probabilistic formats (most often, percentages and single event probabilities).

When people are given information as naturally sampled frequencies, however, they are more likely to generate normatively correct statistical inferences (e.g., Gigerenzer & Hoffrage, 1995). Research has demonstrated that naturally sampled frequencies (often shortened to just "natural frequencies") can be used to improve the communication and understanding of medical statistics by physicians and patients, judges and jurors, and even children (Hoffrage, Kurzenhauser, & Gigerenzer, 2005; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000; Zhu & Gigerenzer, 2006).

Disputes still exist about the theoretical implications of these effects for our human cognitive architecture, and specifically whether or not it implies that the human mind is inherently predisposed to process natural frequencies (Brase, 2008). Distinguishing between these two theoretical interpretations is difficult because the statistical reasoning tasks traditionally used in this research involve exactly the sorts of computations that get simpler with the use of natural frequencies (Gigerenzer & Hoffrage, 1995).

The claims that certain formats - most often percentages or chances—are just as clear and easy to understand as frequencies has been a consistent critique of the ecological rationality view. It is also, however, a critique that rests on very specific and tenuous arguments. Specifically, the claim that chances (e.g., a person has 20 chances out of 100) are probabilities and are therefore importantly different from frequencies (e.g., 20 people out of 100) was promulgated by Girotto & Gonzalez (2001) and argued against by ecological rationality proponents (Brase, 2002b; Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). The crux of the claim is that such chances refer to a single event (e.g., a person) and therefore must be a single-event probability. The counter-argument is that such statements are really about the multiple chances (be they real or hypothetical), and therefore these chances can be just as validly defined as frequencies. Recent research to clarify this issue (Brase, 2008) found that research participants –when asked—did report differing interpretations of the "chances" numbers in Bayesian reasoning tasks (some interpreted them as frequencies; others as single-event probabilities). These participants, however, performed better when given equivalent frequencies as compared to isomorphic tasks that involved chances. Furthermore, the participants who were given tasks that used the "chances" wording performed better if they had adopted a frequentist interpretation of those chances. Thus, the assertion that chances are probabilities, and not frequencies, is something that not only many researchers reject but many research participants also reject. The assertion that performance with chances is equivalent to performance with natural frequencies simply does not survive serious scrutiny.

The issue of how people consider percentages is, like the issue with chances, confused by the ambiguous nature of the format. Generally, percentages can be more formally considered normalized relative frequencies – that is, they are an expression of frequency out of a normalized quantity (100). Thus, 37% means "37 out of 100". This definition of percentages places it within the family of frequencies, but not as naturally sampled information. Any claim that percentages are better (or worse) than frequencies has to be implicitly considering percentages as being defined in a way slightly different from that definition (e.g., considering percentages as expressions of subjective confidence). There is a second problem with the claim that percentages (of some conception) are just as clear and easy to understand as frequencies; it is based on a faulty reading of the research. Barbey and Sloman (2007, p 249) argued for the status of percentages with perhaps the most forceful argument:

> …single numerical statements… have a natural sampling structure, and, therefore, we refer to Brase's "simple frequencies" as natural frequencies in the following discussion. Percentages express single-event probabilities in that they are normalized to an arbitrary reference class (e.g., 100) and can refer to the likelihood of a single-event (Brase, 2002a; Gigerenzer & Hoffrage, 1995). We therefore examine whether natural frequencies are understood more easily and have a greater impact on judgment than percentages

As pointed out by several commentators (e.g., Brase, 2007) this interpretation is only possible due to a remarkable combination of incorrect reinterpretations (for instance, re-labeling simple frequencies as natural frequencies), incomplete definitions (of percentages, which *can* refer to single events, but in the more usual case refer to relative frequencies), and (for the particular case of the Brase, 2002a results) the omission of experimental results that would have invalidated the claim

Statistical reasoning with Natural Frequencies and with Pictures

The use of pictorial representations to aid in judgment and reasoning tasks has been an area of some recent interest. The ecological rationality explanation for performance improvements due to supplementary pictures is that pictures further push participants "to represent the information in the problem as numbers of discrete, countable individuals" (Cosmides & Tooby, 1996, p 33). Other theorists, from a more heuristics and biases position, have explained the pictorial representation boost as a result of the pictures making "nested set relations transparent" (Barbey & Sloman, 2007, p 248; Sloman, Over, Slovak, & Stibel, 2003). Barbey and Sloman also point out that "abstracted, pictorial representation (e.g., Euler circles and tree diagrams) have been shown to improve performance across different deductive inference tasks such as categorical syllogisms and the THOG task, as well as categorical inductive reasoning tasks" (p. 251), which they count as a finding in favor of dual processes theories even though it could legitimately be explained from either perspective.

Recent research has pitted these two interpretations of what aspects of pictorial representations are producing facilitation in judgment under uncertainty tasks. By comparing different types of pictorial representations (Euler circles, rows of icons, and circles with dots in them), Brase (2009) was able to establish that the best and most consistent facilitation due to pictorial representations was from the use of discrete, countable icons. This lends greater weight to the ecological rationality interpretation of pictorial representation facilitation, and thus the privileged representation status of frequencies (and frequency-like pictorial presentations)

**What about Infants and Children?**

If, as the ecological rationality position proposes, people have an inherent tendency to encode, store, and use quantitative information in the form of frequencies then this should be apparent even in the earliest stages of cognitive development. Evidence actually exists of this specific progression across development (in typical environments, in the absences of developmental disorders).

Developmental work has demonstrated surprising quantitative abilities in infants, so long as the information is presented in ecologically realistic, frequentist ways. (It is interesting to note some clear similarities between the infant cognition methodologies and the comparative psychology methodologies described above.) Using methods such as the habituation paradigm (e.g., Xu & Spelke, 2000) and visually presented objects and events, it is possible to "ask" infants questions such as, "If there were two objects, and one is removed, should there be anything left?" (i.e., 2-1=?). Infants within these paradigms are remarkably good at doing simple addition, subtraction, and comparisons (Charles R Gallistel & Gelman, 1992; Gilmore, McCarthy, & Spelke, 2007; McCrink & Wynn, 2004; Van Marle & Wynn, 2011; K Wynn, 1998; Karen Wynn, 1998; F Xu & Spelke, 2000). Infants can also perform matching to samples (Fei Xu & Garcia, 2008).

Such clear, consistent, and dramatic results have led many researchers to propose that there is a core "number sense" that is part of the normally developing human endowment (Feigenson, Dehaene, & Spelke, 2004; Lipton & Spelke, 2003; Spelke & Kinzler, 2007; for dissenting views, see Mix, Huttenlocher, & Levine, 2002; Mix & Sandhofer, 1998). As children get older, this basic number sense appears to become elaborated upon through education (Sophian, 2000). When we reach K-6 level children, as discussed above, use of natural frequencies in Bayesian reasoning tasks becomes relevant. When information is presented in the form of natural frequencies, even 5$^{th}$ and 6$^{th}$ graders can successfully solve simple Bayesian reasoning problems (Zhu & Gigerenzer, 2006).

**Are Quantitative Abilities Consistent across Cultures?**

Cross-cultural lines of evidence for a trait hinge on the twin issues of universality and variation. The psychological and physiological bases of a human trait should be reliably developing in virtually all cultures, but at the same time there can be variations in the expression of traits (i.e., ecology-dependent variability and facultative/conditional adaptations). For example, it is cross-culturally universal that females are capable of bearing children, but there is variation in cultural circumstances that affect rates of actual per-individual children borne.

Deliberate cross-cultural research to evaluate ecological rationality explanations for statistical reasoning have not been done. Much of the research on frequency representations in statistical reasoning has spanned several countries and cultures (United States, Germany, England, etc.), but these have been relatively homogenous in that they are all technologically advanced, Western cultures. Research on Asian/Western differences in general mathematic skills has focused on the acquisition of academic mathematical abilities more generally in students. It remains to be seen what variations exist between Asian and Western cultures when ecologically rational tasks are systematically used. The ecological rationality perspective predicts that a number of the already established differences between these cultures may be reduced when using such tasks.

Although general cross-cultural research can be tremendously helpful, more targeted research on hunter-gatherer societies is particularly important. Hunter-gatherer societies provide an approximation of ancestral human environments, both including information and environments that are no longer common in industrialized societies and excluding information and environments that have only recently become common. It is therefore important to specifically look at hunter-gatherer societies (e.g., via cultural anthropology, human ethology, and human behavioral ecology) as a potential line of evidence.

Comparisons of numerical representation abilities across technologically advanced cultures and hunter-gatherer (or slash and burn horticulturalist) societies are complicated by the radically different educational systems across these cultures. In technologically advanced cultures there is a vast array of *explicit* quantitative information –numbers—in the surrounding environment. Numbers are found on telephones, computers, digital clocks, timers, product labels, newspapers, magazines, television, radio, the internet, and so forth. These explicit numerical representations are largely or entirely absent in hunter-gatherer cultures. Nevertheless, and perhaps remarkably so, mental representations of quantity are ubiquitous across cultures of all types. Brown (1991) summarizes the general state of numerical representation across all cultures thus:

> People may have a very elementary system of numbers and yet have a full-blown ability to count (which allows them very quickly to adopt complex number systems when they become available and prove useful). Theoretically, [Hall, 1975] argues that numbers and counting could be absent among a given people, particularly if they had no need to count. And yet the ability to count is universal as an innate (and presumably specific) capacity of the human mind. (p 46)

One implication of this situation is that the traditional paper-and-pencil statistical reasoning tasks are unsuitable for research with most hunter-gatherer populations.

The range and scale of explicit numbers, however, pales in comparison with the amount of quantitative information that is *implicit* in the world around us: all the objects, events, and locations in the world that can be perceived as multiple, discrete items are potential contributors to implicit quantitative information. If one can track either how often something occurs or how many there are (consciously or subconsciously), it is a source of quantitative information.

All of this quantitative information that exists around us is potential fodder for individuals to use in order to understand and interact with the environment. From lab animals that experience different schedules of reinforcement in operant conditioning (tracking the frequency of rewards along with the frequency of the operant behavior), to wild animals that adjust foraging behaviors in response to the frequency with which food is found in different locations, to people who notice the frequency of cars that go by as they walk along the street (perhaps considering when to cross), individuals track and use quantitative information for successfully navigating the world –even in the absence of explicit numbers.

There does appear to be one culture in which there seems to be a profound lack of any explicit, or even implicit, number system. Members of the Pirahã tribe – a group of South American hunter-gatherers— seem to have no concept of numbering and counting, even when explicitly tested for the existence of such an ability (Frank, Everett, Fedorenko, & Gibson, 2008; P. F. Gordon, 2004). The Pirahã furthermore show profound difficulties in tracking and maintaining visual representations of quantities above about 4 items. This suggests that an ability to track and utilize frequencies of objects, events, and locations may not actually be a universal human ability, under evolutionarily representative circumstances.

These case of the Pirahã tribe presents an intriguing situation, and further work with the Pirahã is certainly warranted although this is difficult given that there are only about 300 members of this tribe. What is particularly perplexing is the long list of abilities that they apparently fail to exhibit: They reportedly have no distinct words for colors, no written language, don't sleep for more than two hours at a time, communicate almost as much by singing, whistling and humming as by normal speech, frequently change their names (because they believe spirits regularly take them over and intrinsically change who they are), do not believe that outsiders understand their language even after they have just carried on conversations with them, have no creation myths, tell no fictional stories and have no art. One can easily get the impression that there is either a more general and pervasive cognitive issue involved with the Pirahã, or that there may be something unusual going on in the social or methodological details of researcher/ Pirahã interactions. One of the foremost experts on the Pirahã has ascribed a large number of these characteristics to a feature of of Pirahã culture that "constrains communication to nonabstract subjects which fall within the immediate experience of interlocutors" (Everett, 2005, p. 621). What is clear is that more, and better controlled studies are needed on the quantitative representational abilities (including implicit representations) of non-industrialized populations (Casasanto, 2005; P. Gordon, 2005; see also Dehaene, Izard, Pica, & Spelke, 2006 for related research on a very similar group of Amazonian hunter-gatherers)

**Is there a Neuroscience of Quantitative Representation?**

The claim that the mind is designed to more effectively take in, process, and understand frequency representations of quantitative information directly implies that there should be specific neuroanotomical structures that are dedicated to performing this function. Although the judgment under uncertainty literature had not generally referenced the neuroimaging literature, there is a wealth of support for this claim.

Both functional neuroimaging studies and studies of patients with selective brain damage indicate that the intraparietal sulcus (IPS) and prefrontal cortex (PFC) are key areas related to the processing and tracking of quantitative information (see Butterworth, 1999; Dehaene, 1997 for reviews). Furthermore, there is evidence that numerical processing ability, in those locations, develops early in the lifespan (at least by three months of age; Izard, Dehaene-Lambertz, & Dehaene, 2008). Other work has demonstrated that analogous brain regions are responsible for numerical information processing in non-human animals (e.g., Diester & Nieder, 2007).

There is still debate as to the particular roles of these different areas, but at this point it appears that the PFC is more involved in the mapping of symbolic representations with numerical concepts, whereas the IPS is more directly involved in the tracking of numerical information in earlier stages. With increasing age and numerical proficiency, the understanding of numerical information (whether symbolic numbers, numbers words, or analog stimuli) appears to rely less on the PFC and more directly on the parietal cortex (D. Ansari, Garcia, Lucas, Hamon, & Dhital, 2005; J.F. Cantlon, Brannon, Carter, & Pelphrey, 2006; Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Piazza, Pinel, Le Bihan, & Dehaene, 2007; Rivera, Reiss, Eckert, & Menon, 2005).

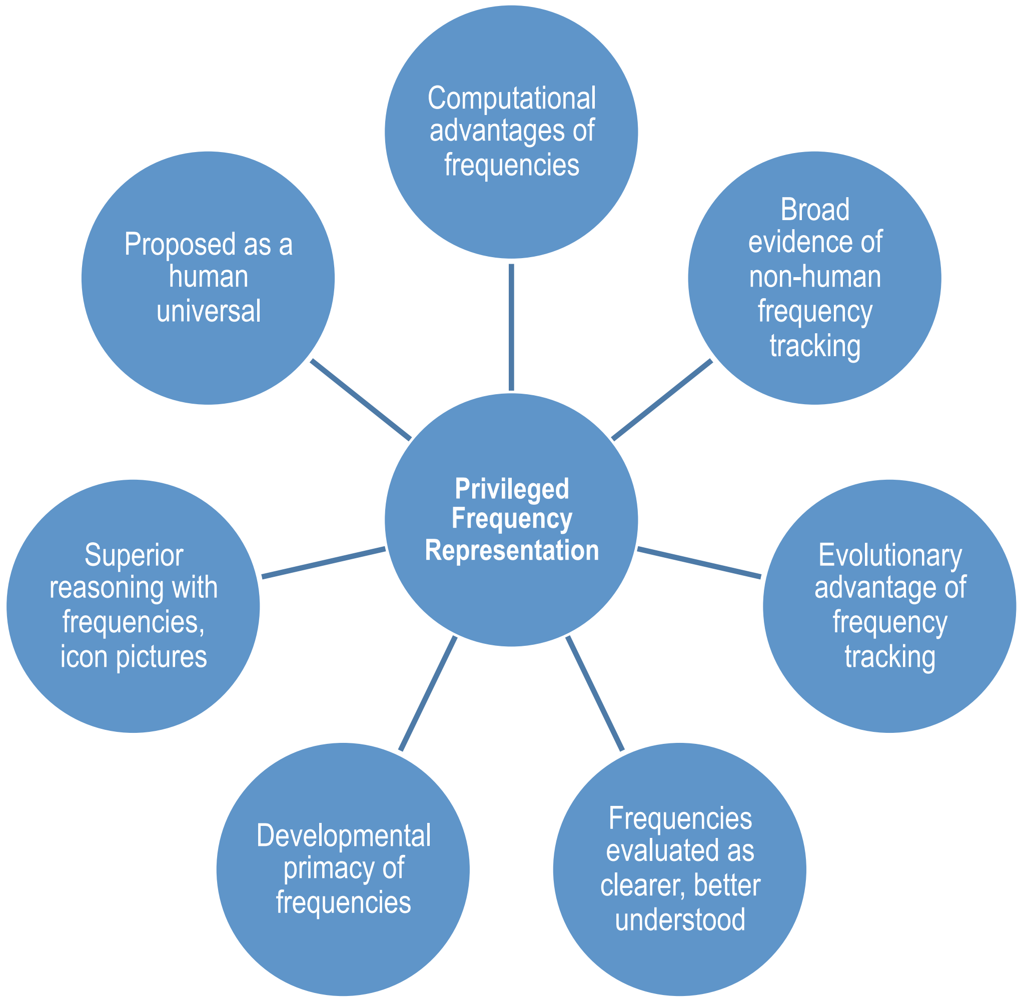
There is also debate at this point about the specific nature of these frequency tracking abilities; whether they are composed of one basic number system or two basic systems (one for small sets and another for large sets; Beran, 2007; J.F. Cantlon & Brannon, 2006; L. Feigenson, Carey, & Hauser, 2002; Lisa Feigenson et al., 2004; Fei Xu, 2003). What is notable for present purposes is that the idea of neurological underpinnings for tracking and storing numerical quantities, established as a fact, is now the background assumption for this debate (Daniel Ansari, 2008; Jacob & Nieder, 2009).

**A Summary of the Evidence**

The preceding sections outlined affirmative evidence for the question of if frequentist interpretations of probability privileged in the human mind, as suggested by an ecological rationality perspective. Specifically, this privileged frequentist thesis derives quite strong theoretical support from evolutionary theory, information theory, cognitive psychology, developmental psychology, cross-cultural research, and neuroscience (see Figure).

The converging lines of evidence shown in the figure helps to identify and organize various bodies of evidence applied to this topic, connecting initially disparate phenomena into integrated scientific understanding, and allowing concise comparisons between the views. These are useful accomplishments for researchers, for scientific communication, and for teaching. (See Schmitt and Pilcher's 2004 model for a similar approach to organizing multiple converging lines of evidence). In summary, when assessed in terms of multiple, independent, and converging lines of evidence, the privileged frequentist hypothesis appears to be a very good scientific description of reality.

*Figure 1: Summary of the Ecological Rationality view of human numerical representation, using the Convergence framework.*




**What does this all mean for Mathematics Education?**

What are the educational implications of having a mind built to preferentially think about frequencies, rather than other numerical formats? First, and most obviously, it implies that other numerical formats should often be problematically interpreted through a frequentist lens. The first part of this section reviews some of the evidence that these phenomena actually occur. We then look at how these phenomena can conceptualized and applied in the service of better mathematics education, and finally we conclude with a summary of an ecologically rational pedagogical model of mathematics education.

<u>Pervasive Frequency Interpretations</u>

First of all, learning of frequencies is developmentally primary (i.e., it happens first), it is relatively unproblematic, and it is therefore the foundation for other numerical concepts. Of course, children can also learn other numerical representations such as percentages, fractions, decimals, and even single-event probabilities. In contrast to basic frequencies, though, these representations require more effort to be learned (Charles R Gallistel & Gelman, 1992), and in the process of learning these other formats there are persistent problems with children trying –incorrectly— to understand these other representations as frequencies. Research indicates that students need to be competent in the four basic operations of whole numbers and have a solid understanding of measurement for them to be prepared to learn rational numbers (Behr & Post, 1988). Rational numbers are the first such numbers experienced by students that are not based on a counting algorithm. This shift is a challenge for many students. Initially, whole number knowledge inhibits learning rational numbers because children over generalize their counting principles and because whole numbers have a "next" number whereas rational numbers do not (Behr et al., 1984; Carraher, 1996; Gelman & Mack, 1992).

For example, there is a long-documented tendency for children to develop "buggy algorithms" in their understanding of fractions and decimals (J. S. Brown & Burton, 1978; Gould, 2002; Lachance & Confrey, 2001; Resnick et al., 1989). A couple of the most common example of these "bugs" are:

a) The rational number addition bug (a.k.a. the freshman bug), which involves adding the numerators and the denominators of two fractions when trying to combine them (e.g., 1/3 + 1/2 = 2/5; Silver, 1986).
b) The decimal version of the rational number addition bug is sometimes called ‘‘Benny's bug'' (Erlwanger, 1973). An example of this is when a student encounters a decimal math problems such as ‘‘0.2 + 2.0 = ?'' and responds with "0.4". What they have done is add together the integers as whole numbers and then placed them after a decimal (presumably because they recognize that this is a decimal problem).

What these bugs have in common, other than their frequency of occurrence, is that they reveal a tendency to treat numbers –even when expressed as fractions and decimals– as frequency counts. If "1/3+1/2=" were referring to actual counts (e.g., 1 out of 3 apples are green, another 1 out of 2 apples are also green, and we put all the apples together in a pile, the answer is indeed 2 out of 5 green apples). There are very specific properties of both fractions and decimals that preclude that interpretation (Brase, 2002c), but those properties do not seem to be as intuitively obvious as the frequentist interpretation and they are too often not clearly taught as the explicit properties of these numerical formats. For example, Kieren (1976) identified seven interpretations of fractions that formed the foundation for solid rational number knowledge. These were ultimately condensed to five distinct, yet interconnected, interpretations of fractions: Part/whole, measure, operator, quotient, and ratio (Kieren, 1988; Lamon, 2001). Clements & Del Campo (1990,, p. 186) note that:

> At various times of their schooling, children are told that the fraction 1/3, for instance, is concerned with each and all of the following: (a) sharing a continuous quantity between three people, (b) sharing 12 (say) discrete objects between three people, (c) dividing the number 1 by the number 3, (d) a ratio of quantities, (e) a 1 for 3 replacement operator, (f) a rational number equal to 2/6, 3/9, and so forth, and (g) a decimal fraction of 0.333.

Decimals are usually taught after fractions in mathematics curricula, and thereafter becomes the predominant format (Hiebert & Wearne, 1986; Hiebert, Wearne, & Taber, 1991). Yet the occurrence of "Benny's bug" indicates that even students who have learned fractions are still predisposed to frequentist interpretations of numbers. In fact, the learning of decimal fractions seems to lead to a proliferation of buggy conceptions (Brase, 2002c; Moloney & Stacey, 1997; Nesher & Peled, 1986; Resnick et al., 1989; Sackur-Grisvard and Leonard, 1985;Steinle, 2004), all of which can be understood as attempts to apply a frequentist view to decimals:

a) Adopting the rule that ''more digits means bigger'' (e.g., 0.1234 is larger than 0.32) generally occurs because students are using a judgment method that was successful for whole numbers
b) Adopting the rule that ''more digits means smaller'' can be a reaction to learning that the first rule is wrong (e.g., 0.4321 is smaller than 0.23), or it can result from the use of place value names to decide on the size of decimals. That is, they recognize that one tenth is larger than one hundredth and reason that any number of tenths is going to be larger than any number of hundredths. They can also respond with this rule if they are confusing notation for decimals with that for negative numbers. When comparing 0.3 and 0.4 they will reason that -3 is greater than -4.
c) Adopting a rule that attaching zeros to the right of a decimal number increases the size of that number. (e.g., 0.8 < 0.80 < 0.800) by treating the decimal part as a whole number
d) Adopting a rule of ignoring zeros on the left (e.g., 0.8 = 0.08 = 0.008) in a way similar to the "more digits means bigger" rule with a variation. Just as the number 8 is not changed by placing a zero in front of it (08), the zeros after the decimal point are also disregarded making 0.8 the same as 0.08 just as 8 is the same as 08.
e) Adopting a rule of ignoring decimals, thus lining up digits on the right rather than lining up the decimal points (e.g., Benny's bug: 0.2 + 4 = 0.6, and variants: 0.07 + 0.4 = 0.11, 6×0.4 = 24, and 42 ÷ 0.6 = 7)

As Hiebert and Wearne note (1986, pp. 204–205):

> Extending concepts of whole numbers into referents that are appropriate for decimal fraction symbols is a delicate process. Students must recognize the features of whole numbers that are similar to decimal fractions and those that are unique to whole numbers...The evidence suggests that many students have trouble selecting the features of whole number that can be generalized... Most errors can be accounted for by assuming the student ignores the decimal point and treat the numbers as whole numbers.''

The transition from basic frequencies to normalized frequencies is a matter of complex changes in perspective. Early stages in the transition involve tracking two frequencies that are repeatedly added. So that "2 out of 3" is equivalent to "4 out of 6" because a repetition of three and a repetition of two generates an equivalent value. Similarly 37% can be interpreted as 37 out of 100, or 370 out of 1000. The "m parts out n parts" that describes additive thinking leads students to think that the numerator is a part of the denominator. Fractions such as 3/2 are thus difficult to interpret because this involves thinking of 3 as part of 2 (Thompson & Saldanha, 2003). Students must learn to interpret fractions as a single number rather than thinking of them as two numbers (Kerslake, 1986). As a student's schooling progresses, and fractions begin to be treated as numbers themselves, a key transition is that students move away from "so many out of so many" and begin to think of fractions multiplicatively rather than additively (Thompson, 2004; Vergnaud 1994). This process begins with the coordination of partitioning, multiplication, and measurement (interpreting “A is m/n times as large as B” as meaning that A is m times as large as an $n^{th}$ part of B), a perspective now a critical part of the Common Core State Standards in Mathematics. It ends when a student’s understanding of fractions is fully integrated with conceptions of proportion and the real number system so that a student can imagine a fraction as a smooth proportion: a relationship between two quantities that change continuously though all real values, but are coordinated in such a way that one is always a constant number of times as large as the other (Castillo-Garsow, 2010; Thompson & Thompson 1992) just as an integer is the result of all possible subtractions that generate that integer (Thompson 1994).

In the case of Ann (Thompon & Thompson; 1994, 1996), this distinction between adding on (additive reasoning), and cutting up (multiplicative reasoning) can be seen to have a direct impact on the ability of a student to solve problems of distance, time, and speed. Ann thought of speed as a repeated distance: a speed of 20 feet per second meant repeatedly adding 20 feet every second. Ann was able to solve unknown time problems by counting the number of times the 20 feet speed-length fit into a total travel distance of 100 feet. However, the problems of finding an unknown speed from a distance and time had her stumped, because she did not have a speed-length to count. It was not until a researcher intervened to build a meaning of speed based on the cutting up of accumulated distance and accumulated time (Cutting up 7 seconds into 7 pieces and 100 feet into 7 pieces to arrive at a speed in feet per second) that Ann was able to solve these problems. We can see in this example a case of the student making a transition from additive reasoning to multiplicative reasoning and building a new understanding of rate and fractional equivalence. She has come to imagine that 100 feet in 7 seconds is the same as 100/7 feet in 1 second, and in doing so she has built equivalence between basic natural frequencies.

## Making sense of the bugs

The key insight from an ecological rationality approach is that these problematic phenomena have a clear and comprehensible explanation. Whereas earlier explanations have tended to view these buggy procedures as developing spontaneously due to overgeneralizations of earlier learned mathematics (whole numbers) to new topics (decimals/fractions; e.g., Resnick et al., 1989), the ecological rationality perspective explains these phenomena as a part of how the human mind is designed to work. This also, of course, helps to explain the frequency and pervasiveness of these buggy procedures across classrooms, topics, and generations of students. What has been described as a “whole number bias” (e.g., Ni, 2005) is not merely an annoyance and a buggy bias to be stomped out; it is a design feature of the human mind. As such, pedagogical techniques will be more fruitful (and less frustrating) is they take those design features into account and work with them in the teaching of mathematics.

A number of the considerations and lines of evidence in the preceding sections have led to theoretical models for the development of mathematics. For instance, Geary (1995) proposes conceptualizing mathematics into two fundamental types: biologically primary mathematical abilities and secondary abilities which develop on top of those primary abilities. This has also been extended to, for instance, the consideration of mathematical disabilities (e.g., dyscalculia; Geary, 2007; Geary, Hoard, & Hamson, 1999). Given the history of a heuristics and biases perspective dominating ideas in mathematics education is it perhaps not too surprising that this more ecologically rational model has met with some resistance (e.g., Glassman, 1996).

The Common Core State Standards (CCSS) places a lot of attention on fractions while, some would say, snubbing decimals. In grades 4 and 5, the symbolic notation of decimals is viewed as a notation for a particular type of fraction (specifically with base-ten denominators). While this marks a clear attempt to avoid the misconception that decimals are something totally different, a concerted effort to help students see this will be necessary. Assisting students as they come to terms with the complexity of decimals is often undervalued and underemphasized. Although students struggle with rational numbers, in general, decimals are one of the greatest challenges. Results from the Third International Mathematics and Science Study (TIMSS) indicate that students do not perform as well on questions involving decimals compared to those involving fractions (Glasgow, Ragan, Fields, Reys, & Wasman, 2000). The decimal representation is compounded in complexity by the merging of whole number knowledge and common fractions with very specific kinds of units. In addition, decimals can be viewed as both continuous and discrete.

The impact of whole number (i.e., frequency) thinking on students' development of decimal number knowledge may also relate to the nature of the knowledge. Students' prior knowledge may be predominately procedural, which may account for misconceptions applied to decimal numbers (Hiebert & Wearne, 1985). Research suggests the same is true for the conceptual understanding for fractional operations. Students have a procedural knowledge of operations rather than an understanding of fundamental concepts that form the basis for fractions (Mack, 1990). Often students inappropriately apply rules, which can result in the right answer for the wrong reason. This reinforces the inappropriate use of the rule, and the error rules persist and procedural flaws are not corrected (Hiebert & Wearne, 1985; Steinle & Stacey, 2004b). Students who are instructed using only procedural methods tend to regress in performance over time (Woodward, Howard & Battle, 1997; Steinle & Stacey, 2004a). In addition, knowledge developed in one representation (e.g., decimals) does not necessarily transfer to all representations of rational numbers (Vamvakoussi & Vosniadou, 2004).

Kieren (1980) suggests that partitioning is fundamental to the meaningful construction of rational number just as counting is the basis to understanding the construction of whole number. In fact, the construction of initial fraction concepts hinges on the coordination of counting and partitioning schemes (Mack, 1993). The notion that partitioning results in a quantity that is represented by a new number is fundamental to rational number. For many children, coordinating their ideas to reach this level takes time and experience. For example, Mack (1995) found that third and fourth graders use their understanding of partitioning to connect operations on fractions to their prior knowledge of whole numbers. This enables them to solve addition and subtraction problems with a common denominator by separating the numerator and denominator and thinking only about the number of pieces combined or removed. Although this process assists them with the operation, it limits their conceptions about fractions by treating them as whole numbers. The same is true with decimal numbers where children separate parts from the number and treat them as a whole number. Instruction on decimals as an extension of the whole number system often occurs without an adequate understanding of place-value concepts that enable them to work with whole numbers (Fuson, 1990). As a result, students often attempt to write too many digits into a column. Extending the place value structure to include digits to the right of the decimal point presents additional challenges. Without understanding the value of the columns, some students think the further away from the decimal point, the larger the value of the digits (e.g., 0.350 is larger than 0.41 because 350 is larger than 41).

Applying whole number properties inappropriately also occurs because students fail to understand the symbolic representation of fractions, which is often prematurely introduced. Research indicates that students base their informal knowledge of fractions on partitioning units and then treat the parts as whole numbers (Ball, 1993; D'Ambrosio and Mewborn, 1994; Mack, 1990, 1995; Streefland, 1991). Without a meaningful understanding for fraction symbols, students acquire misconceptions from attempting to apply rules and operations for whole numbers to fractions. Mack (1995) indicates that, with time and direct effort, students can separate whole number from rational number constructs and develop a meaningful understanding of how fractions and decimals are represented symbolically. Instruction must begin with a focus on student transition from additive to multiplicative reasoning and move students from seeing fractions and decimals as two number to seeing them as a single value (Kent, Arnosky, & McMonagle, 2002; Sowder et al., 1998). Students can then build relationships among fractions, ratios and proportions (Sowder et al., 1998). Fraction concepts and the relationship among various forms of rational numbers can be explained through the use of a combination of representations such as verbal statements, images/pictures, concrete materials, and real world examples, and finally written symbols. Students have to develop appropriate images, actions, and language to set the stage for formal work with rational numbers. (Kieren, 1988, Lesh et al., 1983).

The entrenched tendency of students to over-apply frequentist whole number knowledge is has been recognized in education research and often interpreted as being due to primitive schema that are deeply embedded and difficult to modify (McNeil & Alibali, 2005). On this interpretation, it is the persistence of the over-applications and resulting misconceptions, the resistance to change despite instruction, that presents a problem (Harel & Sowder, 2005; Stienle, 2004 Zazkis & Chernoff, 2008). The experiences that contradict previous knowledge and create cognitive conflict are often avoided when students adopt "coping strategies" such as annexing zeroes to equalize the length of decimal numbers in comparison situations or blindly using left to right digit comparison procedures. Students often adopt these procedures temporarily but then later regress to behaviors consistent with their prior schema (Siegler, 2000; Steinle, 2004).

Another example is descriptions of the "stickiness" of additive reasoning (rather than considering that this may be an inherent predisposition towards frequentist interpretations). Because students are familiar with the concept of addition, and they learn to rely on it early and often, this is assumed to be the reason why it is often very resistant to change. Indeed, additive reasoning is used in a qualitative, intuitive way, not just by seven-year-olds, but by students from the ages of eleven through sixteen who have been taught something about proportions (Hart, 1988). Furthermore, research shows that children do not "grow out" of erroneous addition methods (Thornton & Fuller, 1981) and that this additive error may actually cause a delay in the development of multiplicative thinking (Markovits & Hershokowitz, 1997) which can disrupt work with rational number and ultimately with more advanced topics such as slope and probability. For example, students inappropriately apply additive thinking when judging the equivalency of two fractions with different numerators and denominators (Behr et al., 1984).

Zazkis and Chernoff (2008) suggest that students who wrestle with counter-examples that enable them to personally experience potential cognitive conflict, are more likely to respond with new learning than those where expert opinion is simply staged for them. Contributing to the complexity of rational numbers are the concepts of equivalence and that rational numbers can be represented in various forms and can look very different but mean the same thing.

We would be remiss if we didn't address percent as an equivalent form of rational number. Percent is a particular way to quantify multiplicative relationships. According to Parker and Leinhardt (1995) it "is a comparative index number, an intensive quantity, a fraction or ratio, a statistic or a function" (p. 444). Throughout all of these interpretations, it is "an alternative language used to describe a proportional relationship" (Parker & Leinhardt, 1995, p. 445). Together these interpretations create the full concept of percent and are essential understandings in order to solve a wide variety of problems involving percent (Risacher, 1992). As with the other interpretations of rational number, students often use incorrect rules and procedures related to percentages (Gay, 1997). This may be a direct consequence of students studying from a curriculum that emphasizes rules and procedures (Gay, 1997; Hiebert, 1984; Rittle-Johnson, Siegler & Alibali, 2001). When they are not sure what to do, students will revert to rules and procedures from concepts that are more familiar, more intuitive, and/or resistant to change, such as whole number (Risacher, 1992).

Common errors working with percent have been identified in research (Parker & Leinhardt, 1995; Risacher, 1992). First, students tend to ignore the percent sign and treat the percent as a whole number. Second, they follow what is referred to as the "numerator rule" where they exchange the percent sign on the right with a decimal on the left. Third, they implement the "times table", also known as a "random algorithm" (Parker & Leinhardt, 1995: Parker, 1994, Payne & Allinger, 1984; Risacher, 1992). In addition to problems that arise from overgeneralizing whole number rules and procedures, misconceptions also result from students' limited instruction of percent being part of a whole. Researchers regard the interpretation of a percent as a fractional part of a whole to be of primary importance before working percent problems (Allinger & Payne, 1986). Yet, students who are over reliant on part-whole notions find percentages greater than one hundred problematic since in their mind the part cannot exceed the whole (Parker & Leinhardt, 1995).

Focusing on this controversy, Moss & Case (1999) report on an experimental curriculum that basically introduced the rational number subconstructs of fraction, decimal, percent in reverse. The curriculum begins rational number instruction with percent in a linear measurement context. It then extends that with instruction with decimals to two places then to three and one places. Finally, instruction leads to fraction notation. Results indicate that students using this curriculum model had a deeper understanding of rational number, less reliance on whole number knowledge, and made more frequent references to proportional reasoning concepts. Moss & Case (1999) found that introducing decimals before fractions led to better educational outcomes than the more traditional sequence of teaching decimals after fractions. In addition, the frequency of the "shorter-is-larger" misconception was higher in countries where fractions were taught before decimals, such as United States and Israel (Nesher & Peled , 1986; Resnick et al., 1989) compared to countries like France where decimal fractions expressed in place-value system are taught before other fractions (Resnick et al., 1989; Sackur-Grisvard & Leonard, 1985).

Extending to Probability

Concepts of probability develop in alignment with other ideas in mathematics such as whole number knowledge (counting, addition/subtraction, multiplication/division), fractions (part-whole thinking, equivalent fractions, fraction operations, proportional reasoning), and data (counting/classifying, organization of data, descriptive statistics, data representation and comparison, and sample size). Research into students' ability to compare two probabilities was initiated with work done by Piaget and Inhelder (1951), and other researchers (Ficshbein, 1975; Fischbein & Gazit, 1984; Canizares et al, 1997) have continued the study of students' abilities to reason with probabilities. Comparing probabilities is based on the comparison of two fractions; therefore, proportional reasoning provides the foundation for probabilistic reasoning. Where they differ is that comparing fractions refers to a certain event, whereas comparing (subjective) probabilities involve various amounts of confidence. This degree of uncertainty influences a student's answer because their intuition weighs in on their decision. Intuition is something one considers likely based on instinctive feelings, which may or may not coincide with scientific reasoning, and defines the subjective elements that students assign to probabilities.

In the report *Children's understanding of probability,* Bryant & Nunes (2012) provide a review of four "cognitive demands of understanding probability".

- Understanding randomness
- Working out the sample space
- Comparing and quantifying probabilities
- Understanding correlation (or relationships between events)

The report elaborates on the evidence in each area and highlights studies that are relevant to teaching and learning.

Jones et al. (1997, 1999) evaluated the thinking of 3rd-grade students in an educational setting to create a "Probabilistic Thinking Framework," which describes the levels of reasoning stages related to key constructs (sample space, probability of an event, probability comparisons, and conditional probability). It is interesting to note that these key constructs maintain a striking resemblance to the four cognitive demands made on children when learning probability. Jones and colleagues (1999) also identified four levels of probabilistic thinking. Level 1 is tied to subjective thinking. Level 2 recognizes the transition between subjective thinking and "naïve quantitative thinking" (p. 490). Level 3 engages informal quantitative thinking. Finally, level 4 capitalizes on numerical reasoning. The characteristics within each of the levels of development provide guidance for the design of instructional activities and the selection of instructional strategies that will capitalize on where students are in their reasoning in this sequence of development. Results of this study provide evidence that instruction can influence the learning of probability (Jones et. al, 1999).

Way (2003) similarly found three developmental stages of reasoning with probability tasks, along with two distinct transitional stages. In this research, children (age 4 to 12) were asked to make choices regarding probability tasks and to explain their reasoning. Two types of random number generators were used: Discrete items with up to four colors (numerical form) and spinners with up to four colors (spatial form). Comparisons were drawn within sample space by looking at the colors on one spinner and considering which color is more likely. Comparisons were also made between sample spaces by comparing two spinners and judging with spinner gives the better chance at a specified color (Way, 1996, 2003). Key characteristics for each of the age-related stages of development are described in the table below (see Table 1). Children around the age of nine years had built some of the concepts that provide a foundation for probability and demonstrate learning from participation in probability reasoning tasks. At this age, students are likely to be benefit from instruction that enables them to make connections among concepts (both within the topic of probability and across mathematical ideas that align with probability) and further expand their early numerical strategies into more sophisticated proportional thinking.

*Table 1: Description of Stages of Development (from Way, 2003)*

| **Stage of Development** | **Age** | **Key characteristics** |
|---|---|---|
| **1: Non-Probabilistic Thinking** | Average: 5 years 8 months<br>Range: 4 years 3 months to 8 years 2 months | • Minimal understanding of randomness<br>• Reliance on visual comparison<br>• Inability to order likelihood |
| **Transition from Non-Probabilistic to Emergent Thinking** | Average: 7 years 9 months<br>Range: 5 years 9 months to 11 years 1 months | • Equal number of characteristics from stage 1 and stage 2 |
| **2: Emergent Probabilistic Thinking** | Average: 9 years 2 months<br>Range: 6 years 11 months to 12 years 2 months | • Recognition of sample space structure<br>• Ordering of likelihood through visual comparison or estimation of number<br>• Addition/subtraction strategies used in comparison<br>• Concepts of equal likelihood and impossibility |
| **Transition from Emergent to Quantification** | Average: 9 years 5 months<br>Range: 7 years 6 months to 11 years 8 months | • Emergent thinking dominates but there exists evidence of the initiation of quantification |
| **3: Quantification of Probability** | Average: 11 years 3 months<br>Range: 9 years 1 months to 12 years 7 months | • Numerical comparisons<br>• Doubling and halving<br>• Proportional thinking<br>• Quantification of probability emerging or present |

The classification of probabilistic thinking to levels or stages implies a sequential nature that develops slowly over an extended period of time. Batanero & Diaz (2011) argue that the various meanings of probability must be addressed in school mathematics progressively at various levels, beginning with intuitive ideas and a subjective view of probability as a "degree of belief" and building up to formal definitions and approaches. The variability in students' reasoning with probability marks the importance of designing probability tasks and instructional programs in alignment with their cognitive ability and current understandings. It also highlights the significance of formative assessment techniques used to identify and diagnose students' errors and misconceptions as a solid place to start when designing instruction. Hence, teacher training related to probability becomes a concern. Teachers need to be aware of the common errors and misconceptions that students possess and they need to be properly prepared to develop strategies to help students confront them and push their thinking forward (Fischbein & Gazit, 1984).

Batanero & Diaz (2011) contend that effective teaching and learning of probability in schools will hinge on proper preparation of teachers. They cite several reasons for the difficulty teachers have in teaching probability. Hill, Rowan, & Ball (2005) found that teachers' mathematics preparation and the resulting mathematics knowledge for teaching positively predicted gains in student achievement. Concerns abound that many current teacher-training programs do not sufficiently prepare teachers to teach probability (Franklin & Mewborn, 2006). Research indicates that the content knowledge of both pre-service and inservice teachers at elementary and secondary levels is weak. In the late 1980's, the National Center for Research on Teacher Education reported that elementary and secondary teachers were unable to explain their reasoning or why algorithms they used worked (RAND, 2003). Instead, their explanations were procedural in nature and lacked conceptual understanding, which is consistent with weak understanding going back to the teachers' own education in schools. A study of seventy secondary mathematics pre-service teachers from two universities found significant content weaknesses (Wilburne & Long, 2010), and other research found content knowledge to generally be lacking in conceptual depth (Bryan, 2011). Research has documented the struggle of preservice teachers to identify the source of students' misconceptions and the challenge of finding ways other than the recitation of rules or procedures to eliminate errors and misconceptions (Kilic, 2011).

This is consistent with research on inservice teacher knowledge of probability, which was found to be often limited to procedural of formula-based knowledge and to lack conceptual depth. In one study "most of the teachers primarily had intuitive, informal notions of probability but these later evolved into the classical, frequentist, subjective and mathematical conceptions as they build their conceptual understanding of probability" (Reston, 2012 p. 9). Teachers will generally teach as they were taught, so if they are going to incorporate new ideas into their instruction they need to experience those ideas as students. It is already documented that engaging teachers in an inquiry-based approach, which involved strategies that provided for the confrontation of misconceptions, promoted conceptual development of probability and enhanced their pedagogical skills (Reston, 2012). Albert (2006) argues that subjective probability is often ignored in school curriculum and should play a larger role. The subjective viewpoint is the most general, incorporates the classical and frequentist viewpoints, and expands the definition of probability to include those that cannot be computed and events that cannot be repeated under the same conditions. In his earlier work Albert (2003) found that college students were marked by overall confusion about the three viewpoints of probability. It is critical to establish different perspectives that enable students to gain confidence in their abilities to reason probabilistically (Albert, 2006). According to Shaughnessy (1992), the type of task to be investigated and the type of problem to be solved should drive the viewpoint that is taken in a particular context. Without coherent instruction on probability, students leave the K-12 school system with a jumbled perspective of probability that reflects a mix of the classical, frequentist, and subjectivist viewpoints. Without an understanding of these distinct viewpoints we will continue to see evidence of children and adults inappropriately applying these models of probability.

What exactly should an ecologically rational pedagogy for mathematics education look like, including the development of probabilistic thinking? In general, educational practices should work to bootstrap from these privileged representations (rather than fight them) and built towards a more inclusive and comprehensive model of mathematics knowledge. In summary, we suggest that the following educational plan would be more effective and less difficult (for both students and teachers):

a) *Start with frequencies*: This is the natural starting point and should be recognized as such, returning to this starting point as necessary when building up to some more advanced concept.
b) *Introduce percentages and decimals before fractions*: Given the partitioning nature of these numerical formats, relative to whole numbers, it makes more sense to start with formats that use standard units of partitioning first (i.e., percentages can be thought of as 100 parts; decimals can be thought of as successive partitions of 10). This can actually help to lay the groundwork of the much more partition-complex situation represented by fractions
c) *Explain percentages, decimals, and fractions in terms of how they differ from frequencies*: Teachers should expect that many students will adopt frequentist interpretations of these numerical formats. Rather than pretend that those are arbitrarily or definitively wrong, without any explanation, teachers should be allowed and encouraged to address common bugs (i.e., frequentist interpretations that are flawed within these format contexts) in an explicit and transparent way. This can lead to student not just seeing the surface similarities between frequencies and other formats, but also understanding that there are further properties of non-frequency formats.
d) *Finally, probabilities should be introduced in a similar manner*: Probabilities based on a frequentist interpretation are valid (just like, in some contexts frequency-based percentages, decimals and fractions can exist). This can be recognized, acknowledged as a real perspective, but then students can be progressively moved to a more extensive and complete understanding of what probabilities can be (either frequentist or subjectivist interpretations). Accomplishing this can lead to students who not only engage in correct probabilistic reasoning, but who understand why and how they did so.

## References


Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport*, *16*, 1769-1773.

Ansari, Daniel. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, *9*(4), 278-291. doi:10.1038/nrn2334

Ball, D. (1993). Halves, pieces and twoths: constructing and using representational contexts in teaching fractions. In T.P.Carpenter, E. Fennema and T.A Romberg (Eds.) *Rational numbers: An integration of research*. Hillsdale, Lawrence Erlbaum Associates, NJ, 157-195.

Batanero, C. & Diaz, C. (2011). Training school teachers to teach probability: Reflections and challenges. *Chilean Journal of Statistics*. Retrieved August 2012 from: http://chjs.deuv.cl/iFirst_art/ChJS010202.pdf

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*(3), 241-254.

Behr, M.J.; Wachsmuth, I., Post, T.R., & Lesh, R. (1984). Order and equivalence of rational numbers: A clinical teaching experiment. *Journal for Research in Mathematics Education, 15*, 323-341.

Beran, M. J. (2007). Rhesus monkeys (Macaca mulatta) enumerate sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*, 42-54.

Beran, M. J. (2008). The evolutionary and developmental foundations of mathematics. *PLoS biology*, *6*(2), e19. doi:10.1371/journal.pbio.0060019

Beran, M. J., & Rumbaugh, D. M. (2001). "Constructive" enumeration by chimpanzees (Pan troglodytes). *Animal Cognition*, *4*, 81-89.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a champanzee (Pan troglodytes). *Journal of Comparative Psychology*, *103*, 23-31.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (Macaca mulatta). *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 31-49.

Brannon, E. M., & Terrace, H. S. (2002). The evolution and ontogeny of ordinal numerical ability. In M. Bekoff, C. Allen, & G. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 197-204). Cambridge, MA, US: MIT Press.

Brase, G. L. (2002a). Which statistical formats facilitate what decisions? The perception and influence of different statistical information formats. *Journal of Behavioral Decision Making*, *15*(5), 381-401.

Brase, G. L. (2002b). Ecological and evolutionary validity: Comments on Johnson-Laird, Legrenzi, Girotto, Legrenzi, and Caverni's (1999) mental-model theory of extensional reasoning. *Psychological review*, *109*(4), 722-728.

Brase, G. L. (2002c). "' Bugs'" built into the system : How privileged representations influence mathematical reasoning across the lifespan. *Learning and Individual Differences*, *12*, 391-409.

Brase, G. L. (2007). Omissions, conflations, and false dichotomies: Conceptual and empirical problems with the Barbey & Sloman account. *Behavioral and Brain Sciences*, *30*(3), 258-259.

Brase, G. L. (2008). Frequency interpretation of ambiguous statistical information facilitates Bayesian reasoning. *Psychonomic Bulletin and Review*, *15*(2), 284-289.

Brase, G. L. (2009). Pictorial Representations in Statistical Reasoning. *Applied Cognitive Psychology*, *23*(3), 369-381.

Brown, D. (1991). *Human Universals*. New York: McGraw-Hill.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, *2*(2), 155-192.

Bryan, T.J. (2011). The Conceptual Knowledge of Preservice Secondary Mathematics Teachers: How Well Do They Know the Subject matter they will teach? *Issues in the Undergraduate Preparation of Teachers: The Journa*l, Vol 1.

Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.

Canizares, M.J., Batareno, C., Serrano, L. & Ortiz, J.J. (1997). Subjective elements in children's comparison of probabilities. In E. Pehkonen (Ed.). *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education (PME)*. University of Helsinki.

Cantlon, Jessica F, & Brannon, E. M. (2007). Basic math in monkeys and college students. *PLoS biology*, *5*(12), e328. doi:10.1371/journal.pbio.0050328

Cantlon, J.F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401-406.

Cantlon, J.F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, *4*, e125.

Casasanto, D. (2005). Crying "Whorf." *Science*, *307*, 1721-1722.

Casscells, W., Schoenberger, A., Grayboys, T., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, *299*(18), 999-1000. doi:10.1056/NEJM197811022991808

Castillo-Garsow, C. W. (2010). *Teaching the Verhulst model: A teaching experiment in covariational reasoning and exponential growth*. PhD thesis, Arizona State University, Tempe, AZ.

Clements, M. A., & Del Campo, G. (1990). How natural is fraction knowledge? In L. P. Steffe & T. Wood (Eds.), *Transforming Chiildren's Mathematics Education: International Perspectives* (pp. 181-188). Hillsdale, NJ: Larwrence Erlbaum.

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron*, *53*(2), 307-314. doi:10.1016/j.neuron.2006.12.025

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*(1), 1-73.

D'Ambrosio, B.S. & Mewborn, D.S. (1994). Children's constructions of fractions and their implications for classroom instruction. *Journal of Research in Childhood Education*, *8,* 150-161.

De Cruz, H. (2006). Why are some numerical concepts more successful than others? An evolutionary perspective on the history of number concepts. *Evolution and Human Behavior*, *27*(4), 306-323.
Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press.
Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science*, *311*(5759), 381-384. doi:10.1126/science.1121739
Diester, I., & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLoS Biology*, *5*, e294.
Erlwanger, S. H. (1973). Benny's conception of rules and answers in IPI mathematics. *Journal of Children's Mathematical Behavior*, *1*(2), 7-26.
Everett, D. L. (2005). Cultural Constraints on Grammar and Cognition in Piraha Another Look at the Design Features of Human Language. *Current Anthropology*, *46*(4), 621-646.
Feigenson, L., Carey, S., & Hauser, M. D. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, *13*, 150-156.
Feigenson, Lisa, Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307-314. doi:10.1016/j.tics.2004.05.002
Fischbein, E. (1975). The Intuitive Sources Of Probabilistic Thinking In Children. Dordrecht: Reidel.
Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics*, 15, 1-24.
Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819-24. doi:10.1016/j.cognition.2008.04.007
Fuson, K.C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction* 7(4), 343-403.
Gallistel, C R, Gelman, R., & Cordes, S. (2006). The cultural and evolutionary history of the real numbers. In S. C. Levinson & J. Pierre (Eds.), *Evolution and culture: A Fyssen Foundation symposium* (pp. 247-274). Cambridge, MA: MIT Press.
Gallistel, Charles R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.
Gallistel, Charles R, & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.
Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*(3), 149-159.
Gay, A. S., & Aichele, D. B. (1997). Middle school students' understanding of number sense related to percent. *School Science and Mathematics*, *97*, 27-36.
Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist*, *50*(1), 24-37.
Geary, D. C. (2007). An evolutionary perspective on learning disability in mathematics. *Developmental Neuropsychology*, *32*(1), 471-519.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology*, *74*(3), 213-239.

Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A Handbook for Data Analysis in the Behavioral Sciences* (pp. 311-339). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*(4), 684-704. doi:10.1037/0033-295X.102.4.684

Gigerenzer, G., Swijtink, Z. G., Porter, T. M., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance: How probability changed science and everyday life*. New York, NY: Cambridge University Press.

Gigerenzer, G., Todd, P. M., & Group, T. A. B. C. R. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*(7144), 589-91. doi:10.1038/nature05850

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition*, *78*(3), 247-76.

Glasgow, R., Ragan, G., Fields, W. M. (2000). **The decimal dilemma.** Teaching Children Mathematics, 7(2), 89.

Glassman, M. (1996). The argument for constructivism. *American Psychologist*, *51*(3), 264-265.

Gordon, P. (2005). Author's response to "Crying Whorf." *Science*, *307*, 1722.

Gordon, P. F. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496-499.

Gould, P. (2002). Year 6 Students ' Methods of Comparing the Size of Fractions, 393-400.

Hacking, I. (1975). *The emergence of probability: A philosophical study of early ideas about probability induction and statistical inference*. Cambridge, MA: Cambridge University Press.

Harel, G., & Sowder, L. (2005). Advanced mathematical-thinking at any age: Its nature and its development, *Mathematical Thinking and Learning, 7*, 27-50

Hart, K. M. (1981). *Children's Understanding of Mathematics: 11–16. The CSMS Mathematics Team*. John Murray, London:

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356-388.

Hasher, L., & Zacks, R. T. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency Processing and Cognition* (pp. 21-36). Oxford: Oxford University Press.

Hiebert, J. (1984). Children's Mathematics Learning: The struggle to link form and understanding. *The Elementary School Journal*, *84*(5), 497-511.

Hiebert, J., & Wearne, D. (1985). A model of students' decimal computation procedures. *Cognition and Instruction* 2 (3&4), 175-205.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*.

Hiebert, J., Wearne, D., & Taber, S. (1991). Fourth graders' gradual construction of decimal fractions during instruction using different physical representations. *The Elementary School Journal*, *91*, 321-341.

Hill, H.C., Rowan, B. & Ball, D.L. (2005) Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. American Educational Research Journal, 42(2), 371-406.

Hintzman, D. L., & Stern, L. D. (1978). Contextual Variability and Memory for Frequency. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(5), 539-549.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: what natural frequencies are and what they are not. *Cognition*, *84*(3), 343-52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12044739

Hoffrage, U., Kurzenhauser, S., & Gigerenzer, G. (2005). Understanding the Results of Medical Tests: Why the Representation of Statistical Information Matters. In R. Bibace, J. D. Laird, K. L. Noller, & J. Valsiner (Eds.), *Science and medicine in dialogue: Thinking through particulars and universals* (pp. 83-98). Westport, CT: Praeger Publishers/Greenwood Publishing Group.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, *290*(5500), 2261-2262.

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17*(23), R1004-5. doi:10.1016/j.cub.2007.10.027

Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct cerebral pathways fro object identity and number in human infants. *PLoS Biology*, *6*, e11.

Jacob, S. N., & Nieder, A. (2009). Notation-independent representation of fractions in the human parietal cortex. *Journal of Neuroscience*, *29*(14), 4652-7. doi:10.1523/JNEUROSCI.0651-09.2009

Jones, G., Langrall, C., Thornton, C. & Mogill, T. (1997). A framework for assessing and nurturing young children's thinking in probability. Educational Studies in Mathematics, 32, 101-125.

Jones, G., Langrall, C., Thornton, C. & Mogill, T. (1999). Using probabilistic thinking to inform instruction. Journal for Research in Mathematics Education, 30(5), 487-519.

Jones, G. A., & Thornton, C. A. (2005). An overview of research into the learning and teaching of probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 65-92). New York, NY: Springer.

Kahneman, D, & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237-251.

Kahneman, Daniel. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.

Kahneman, Daniel, & Frederick, S. (2001). 2 . Representativeness Revisited : Attribute Substitution in Intuitive Judgment.

Kent, L. B., Arnosky, J., & McMonagle J. (2002). Using representational contexts to support multiplicative reasoning. In B. Litwiller & G. Bright (Eds.), *Making sense* *of fractions, ratios, and proportions: 2002 yearbook* (pp. 145–152). Reston, VA: National Council of Teachers of Mathematics.

Kieren, T.E. (1980). The rational number construct: Its elements and mechanisms. In T. Kieran (Ed.), *Recent research on number learning* (pp. 125-149). Columbus, Ohio: ERIC-SMEAC.

Kieren, T.E. (1988) Personal knowledge of rational numbers: Its intuitive and formal the middle grades (pp. 162-181). Reston, VA: Author.

Kilic, H. (2011). Preservice Secondary Mathematics Teachers' Knowledge of Students. Turkish Online Journal of Qualitative Inquiry, 2(2).

Kleiter, G. (1994). Natural sampling: Rationality without base rates. In G. H. Fischer & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 375-388). New York: Springer.

Lachance, A., & Confrey, J. (2001). Helping students build a path of understanding from ratio and proportion to decimal notation. *Journal of Mathematical Behavior*, *20*(4), 503-526.

Lesh, R., Post, T. R., & Behr, M. (1988). Proportional Reasoning. In J. Heibert & M. Behr (eds.), *Number Concepts and Operations in the Middle Grades*. National Council of Teachers of Mathematics, Reston, pp. 93–118.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396-401. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12930467

Mack, N.K. (1990) Learning fractions with understanding: building on informal knowledge, *Journal for Research in Mathematics Education, 21*(1), 16-33.

Mack, N.K.(1993) Learning rational numbers with understanding: The case of informal knowledge. In T.P.Carpenter, E. Fennema and T.A Romberg (Eds.), *Rational numbers: An integration of research* (pp. 85-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Markovits, Z. & Hershkowitz, R. (1997). Relative and absolute thinking in visual estimation processes. Educational *Studies in Mathematics, 32,* 29-47.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, *315*, 57-59.

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776-781. doi:10.1111/j.0956-7976.2004.00755.x

McNeil, N. M. & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood.* New York: Oxford University Press.

Mix, K. S., & Sandhofer, C. M. (2007). Do we need a number sense? In M. J. Roberts (Ed.). *Integrating the mind*. Hove, UK: Psychology Press.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: a new model and an experimental curriculum. *Journal for Research in Mathematics Education*, *30*, 122-147.

Nesher, P. & Peled, I. (1986). Shifts in reasoning. *Educational Studies in Mathematics*, *17,* 67-79.

Ni, Y. (2005). Teaching and Learning Fraction and Rational Numbers: The Origins and Implications of Whole Number Bias, (776109343). doi:10.1207/s15326985ep4001

Over, D. E., & Green, D. W. (2001). Contingency, causation, and adaptive inference. *Psychological Review*, *108*(3), 682-684.

Parker, M. & Leinhardt, G. (1995). Percent: A privileged proportion. Review of Educational Research, 65(4), 421-481.

Payne, J. N., & Allinger, G. D. (1984). *Insights into teaching percent to general mathematics students.* Unpublished manuscript, Montana State University, Deparment of Mathematical Sciences, Bozeman.

Piaget, J. & Inhelder, B. (1975) The Origin of the Idea of Chance in Children. London: Routledge & Kegan Paul. (Original work published 1951).

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293-305. doi:10.1016/j.neuron.2006.11.022

RAND Mathematics Study Panel [RAND MSP]. (2003). *Mathematical proficiency for all students: Towards a strategic development program in mathematics education*. Santa Monica, CA: RAND Corporation. (MR-1643.0-OERI).

Real, L. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science*, *253*, 980-986.

Real, L., & Caraco, T. (1986). Risk and foraging in stochastic environments: Theory and evidence. *Annual Review of Ecology and Systematics*, *17*, 371-390.

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education*, *20*, 8-27.

Reston, E (2012). *Exploring inservice elementary mathematics teachers' conceptions of probability through inductive teaching and learning methods.* 12th International Congress on Mathematics Education, Soeul, Korea.

Risacher, B. F. (1992). Knowledge growth of percent during the middle school years. *Dissertation Abstracts International*, 54 (03), 853A.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, *15*, 1779-1790.

Schmitt, D. P., & Pilcher, J. J. (2004). Evaluating evidence of psychological adaptation: How do we know one when we see one? *Psychological Science*, *15*(10), 643-649. doi:10.1111/j.0956-7976.2004.00734.x

Sedlmeier, P., Hertwig, R., & Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *Journal of*

*Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 754-770. . doi:10.1037/0278-7393.24.3.754

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In *Handbook of Research on Mathematics Teaching and Learning.* D.A. Grouws (Ed.). pp. 465-494. New York: Macmillan.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91, 417-447,*

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. (pp. 181–198). Hillsdale, NJ: Lawrence Erlbaum Associates.
Erlbaum Associates.

Simon, H. A. (1956). Rational choice and the structure of environments. *Psychological Review*, *63*, 129-138.

Sloman, S. a., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, *91*(2), 296-309. doi:10.1016/S0749-5978(03)00021-9

Sophian, C. (2000). From objects to quantities: Developments in preschool children's judgments about aggregate amount. *Developmental Psychology*, *36*(6), 724-730.

Sowder, J., (1997). Place value as the key to teaching decimal operations. *Teaching Children Mathematics 3*(8) 448-453.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89-96. doi:10.1111/j.1467-7687.2007.00569.x

Staddon, J. F. R. (1988). Learning as inference. In R. C. Bolles & M. D. Beecher (Eds.), *Evoution and Learning* (pp. 59-77). Hillsdale, NJ: Erlbaum.

Steinle, V. (2004). *Changes with age in students' misconceptions of decimals numbers*. University of Melbourne.

Steinle, V., & Stacey, K. (2004a). A longitudinal study of students understanding of decimal notation: an overview and refined results. In I. Putt, R. Faragher & M. McLean (Eds.), *Proceedings of the 27th annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 541-48). Townsville: MERGA.

Steinle, V., & Stacey, K. (2004b). Persistence of decimal misconceptions and readiness to move to expertise, *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 225-232). Bergen: PME.Thornton & Fuller, 1981

Stephens, D., & Krebs, J. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.

Sternberg, R J, & Grigorenko, E. L. (2001). Unified psychology. *The American Psychologist*, *56*(12), 1069-79.

Sternberg, Robert J, & Grigorenko, E. I. (2001). The (Mis)organization of Psychology. *APS Observer*, *14*(1).

Thompson, P. W. (1994). The development of the concept of speed and its relationship to concepts of rate. In Harel, G. and Confrey, J., editors, *The development of*

*multiplicative reasoning in the learning of mathematics,* pages 181–234. SUNY Press, Albany, NY.
Thompson, P. W. (2004). *Fractions and multiplicative reasoning*, pages 95–114. National Council of Teachers of Mathematics, Reston, VA.
Thompson, P. W. and Thompson, A. G. (1992). *Images of rate*. San Francisco, CA.
Thompson, A. G. and Thompson, P. W. (1996). Talking about rates conceptually, part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27,* 2–24.
Thompson, P. W. and Thompson, A. G. (1994). Talking about rates conceptually, part I: A teacher's struggle. *Journal for Research in Mathematics Education, 25,* 279–303.
Tomonaga, M., & Matsuzawa, T. (2002). Enumeration of briefly presented items by the chimpanzee (Pan troglodytes) and humans (Homo sapiens). *Animal Learning and Behavior*, *30*, 143-157.
Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 19-136). Oxford: Oxford University Press.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105-110.
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(207-232).
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453-458.
Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293-315.
Vamvakoussi, X. & Vosniadou, S. **(2004***). Understanding the structure of rational numbers: A conceptual change approach.* Στο L. Verschaffel και S. Vosniadou (Eds.), Conceptual Change in Mathematics Learning and Teaching, Special Issue of *Learning and Instruction, 14,* pp.453-467
Van Marle, K., & Wynn, K. (2011). Tracking and quantifying objects and non-cohesive substances. *Developmental Science*, *14*(3), 502-15. doi:10.1111/j.1467-7687.2010.00998.x
Wang, X. T. (1996a). Domain-specific rationality in human choices: Violations of utility axioms and social contexts. *Cognition*, *60*(1), 31-63. doi:10.1016/0010-0277(95)00700-8
Wang, X. T. (1996b). Framing effects: Dynamics and task domains. *Organizational Behavior and Human Decision Processes*, *68*, 145-157.
Wang, X. T., & Johnston, V. S. (1995). Perceived social context and risk preference: A re-examination of framing effects in a life-death decision problem. *Journal of Behavioral Decision Making*, *8*, 279-293.
Wang, X. T., Simons, F., & Brédart, S. (2001). Social cues and verbal framing in risky choice. *Journal of Behavioral Decision Making*, *14*(1), 1-15. doi:10.1002/1099-0771(200101)14:1<1::AID-BDM361>3.3.CO;2-E
Way, J. (1996). Children's strategies for comparing two types of random generators. In L, Puig & A, Guitierrez (Eds.), Proceedings of the Twentieth International Conference for the Psychology of Mathematics Education. (Vol. 4, pp. 419-426) Valencia,

Spain.
Way, J. (2003). The development of young children's notions of probability. *CERME 3*, Italy.
Wilburne, J. M & Long, M. (2010). Secondary Pre-service Teachers' Content Knowledge for State Assessments: Implications for Mathematics Education Programs. IUMPST: The Journal. Vol 1 (Content Knowledge). www.k-12prep.math.ttu.edu
Woodward, J., Howard, L., Battle, R. (1997). Learning subtraction in the third and fourth grade: What develops over time (Tech. Report 97-2). Tacoma, WA: University of Puget Sound.
Wynn, K. (1998). Numerical competence in infants. In C. Donlan (Ed.), *The development of mathematical skills. Studies in developmental psychology*.
Wynn, Karen. (1998). An evolved capacity for number. In D. Cummins Dellarosa & C. Allen (Eds.), *The Evolution of Mind* (pp. 107-126). New York, NY: Oxford University Press.
Xu, F, & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1-B11.
Xu, Fei. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15-B25. doi:10.1016/S0010-0277(03)00050-7
Xu, Fei, & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(13), 5012-5. doi:10.1073/pnas.0704450105
Zazkis, R. & Chernoff, E. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics, 68*(3), 195-208.
Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, *98*(3), 287. doi:10.1016/j.cognition.2004.12.003