

Connected Mathematics - Building Concrete Relationships with Mathematical Knowledge

Uriel Joseph Wilensky

B.A., Brandeis University, 1977
M.A., Brandeis University, 1977

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

May 1993

© Massachusetts Institute of Technology 1993
All right reserved

Author.....

Program in Media Arts and Sciences

April 30, 1993

Certified by.....

Seymour Papert

Lego Professor of Epistemology and Learning Research

Thesis Supervisor

Accepted by.....

Stephen Benton

Chairman, Departmental Committee on Graduate Students

Program in Media Arts & Sciences

Connected Mathematics - Building Concrete Relationships with Mathematical Knowledge

Uriel Joseph Wilensky

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning
on April 30, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The context for this thesis is the conflict between two prevalent ways of viewing mathematics. The first way is to see mathematics as primarily a formal enterprise, concerned with working out the syntactic/formal consequences of its definitions. In the second view, mathematics is a creative enterprise concerned primarily with the construction of new entities and the negotiation of their meaning and value. Among teachers of mathematics the formal view dominates. The consequence for learners is a shallow brittle understanding of the mathematics they learn. Even for mathematics that they can do, in the sense of calculating an answer, they often can't explain why they're doing what they're doing, relate it to other mathematical ideas or operations, or connect the mathematics to any idea or problem they may encounter in their lives. The aim of this thesis is to develop alternative ways of teaching mathematics which strengthen the informal, intuitive and creative in mathematics.

This research develops an approach to learning mathematics called "connected mathematics" which emphasizes learners' negotiation of mathematical meaning. I have employed this approach by conducting in-depth mathematical interviews with adult learners. Through these interviews, I also infer the impact of the usual school approach to teaching mathematics, and my primary target will be the formal methods of instruction used in *university* mathematics classes – the litany of definition-theorem-proof which has remained virtually unchanged since the beginnings of universities. By contrast, in the connected mathematics approach, learners are encouraged to inquire into and construct meanings for mathematical concepts, make explicit connections between mathematical ideas and other pieces of knowledge both mathematical and non-mathematical, and by so doing foster more robust mathematical intuitions. Furthermore, learners are provided with an environment (conceptual, computational and social) in which they can explore the conceptual space of the mathematics and tie together ideas that don't seem to have any connection in a traditional curriculum. They are encouraged to develop multiple representations of mathematical objects and operations, to perceive the wide array of choices available in making these representations, and to develop connections among, and critiques of, these representations. They also make conjectures about what relations obtain between mathematical objects, and validate or disprove these conjectures through social argumentation and interaction in an environment open enough to permit learners to design projects around their own goals.

In this dissertation, I touch on three mathematical content areas that are often difficult for learners. My primary area of focus is mathematical probability. Pilot studies on fractions and recursion are also investigated. I show how traditional mathematics education approaches have led to "disconnectedness" in the mathematical knowledge of the learner, and how a "connected mathematics" approach can alleviate this difficulty. I revisit the concept of mathematical proof and show how, when examined from a connected mathematics perspective, proof is seen less as the source of mathematical certainty, and more as a way of connecting different parts of one's mathematical knowledge. I use the research of the psychologists Tversky & Kahneman and the responses to it by mathematics educators as a prism through which to illuminate the differences between these two disparate views of mathematics and their consequences for learning and for the development of mathematical intuitions. In contrast to claims made by educators, based on the research of Tversky & Kahneman, that we should mistrust our intuitions about probability, I shall claim that failures in developing good probabilistic intuitions, like other failures in mathematical understanding, are due to the lack of good learning environments and to the brittle formal methods used in mathematics education. As part of this effort, I develop a theory of mathematical learning which reveals the "messy" nature of mathematical understanding and provides an account of what it means to have concrete vs. abstract mathematical knowledge.

Thesis Supervisor: Seymour Papert
LEGO Professor of Learning Research
Media Arts and Sciences Section

This thesis describes research done at MIT's Media Laboratory. Support for this research was provided by LEGO Systems, Inc., the National Science Foundation (Grants # MDR-8751190, # TPE-8850449, and # MDR-9153719), Nintendo Co., Ltd, and the MacArthur Foundation (Grant # 874304). The ideas expressed here do not necessarily reflect those of the supporting agencies.

*I went for a walk over the dunes again this morning
to the sea,
then turned right along
the surf*

*rounded a naked headland
and returned*

along the inlet shore:

*it was muggy sunny, the wind from the sea steady and high,
crisp in the running sand,
some breakthroughs of sun
but after a bit*

continuous overcast:

*the walk liberating, I was released from forms,
from the perpendiculars,
straight lines, blocks, boxes, binds
of thought
into the hues, shadings, rises, flowing bends and blends
of sight:*

*I allow myself eddies of meaning
yield to a direction of significance
running
like a stream through the geography of my work:
you can find
in my sayings*

*swerves of action
like the inlet's cutting edge:
there are dunes of motion,
organizations of grass, white sandy paths of remembrance
in the overall wandering of mirroring mind:*

*but Overall is beyond me: is the sum of these events
I cannot draw, the ledger I cannot keep, the accounting
beyond the account:*

*in nature there are few sharp lines: there are areas of
primrose
more or less dispersed;
disorderly orders of bayberry; between the rows
of dunes,
irregular swamps of reeds,
though not reeds alone, but grass, bayberry, yarrow, all . . .
predominantly reeds:*

*I have reached no conclusions, have erected no boundaries,
shutting out and shutting in, separating inside
from outside: I have
drawn no lines:
as*

*manifold events of sand
change the dune's shape that will not be the same shape
tomorrow,*

*so I am willing to go along, to accept
the becoming
thought, to stake off no beginnings or ends, establish
no walls:*

*by transitions the land falls from grassy dunes to creek
to undercreek: but there are no lines, though
change in that transition is clear
as any sharpness: but "sharpness" spread out,
allowed to occur over a wider range
than mental lines can keep:*

*the moon was full last night: today, low tide was low:
black shoals of mussels exposed to the risk
of air
and, earlier, of sun,
waved in and out with the waterline, waterline inexact,
caught always in the event of change:
a young mottled gull stood free on the shoals
and ate*

*to vomiting: another gull, squawking possession, cracked a crab,
picked out the entrails, swallowed the soft-shelled legs, a ruddy
turnstone running in to snatch leftover bits:*

*risk is full: every living thing in
siege: the demand is life, to keep life: the small
white blacklegged egret, how beautiful, quietly stalks and spears
the shallows, darts to shore
to stab –what? I couldn't
see against the black mudflats –a frightened
fiddler crab?*

*the news to my left over the dunes and
reeds and bayberry clumps was
fall: thousands of tree swallows
gathering for flight:
an order held
in constant change: a congregation
rich with entropy: nevertheless, separable, noticeable
as one event,
not chaos: preparations for
flight from winter,
cheet, cheet, cheet, cheet, wings rifling the green clumps,
beaks
at the bayberries
a perception full of wind, flight, curve,*

*sound:
the possibility of rule as the sum of rulelessness:
the "field" of action
with moving, incalculable center:*

*in the smaller view, order tight with shape:
blue tiny flowers on a leafless weed: carapace of crab:
snail shell:
pulsations of order*

*in the bellies of minnows: orders swallowed,
broken down, transferred through membranes
to strengthen larger orders: but in the large view, no
lines or changeless shapes: the working in and out, together
and against, of millions of events: this,
so that I make
no form of
formlessness:*

*orders as summaries, as outcomes of actions override
or in some way result, not predictably (seeing me gain
the top of a dune,
the swallows
could take flight –some other fields of bayberry
could enter fall
berryless) and there is serenity:*

*no arranged terror: no forcing of image, plan,
or thought:
no propaganda, no humbling of reality to precept:*

*terror pervades but is not arranged, all possibilities
of escape open: no route shut, except in
the sudden loss of all routes:*

*I see narrow orders, limited tightness, but will
not run to that easy victory:
still around the loser, wider forces work:
I will try
to fasten into order enlarging grasps of disorder, widening
scope, but enjoying the freedom that
Scope eludes my grasp, that there is no finality of vision,
that I have perceived nothing completely,
that tomorrow a new walk is a new walk.*

Corson's Inlet

- A.R. Ammons

Acknowledgments

As many have said before me, a dissertation is not a solitary work but a collaboration of a lifetime of influences.

My greatest intellectual debt is to my advisor Seymour Papert, who through his writing, his probing criticisms which so often went to the heart of the matter, his ferocious intellectual energy, has changed my life. I have learned from him immeasurably. Thank you Seymour for seeing me and for so often knowing what I mean by what I say before I do.

David Chen has been an encouragement and a support from the moment I met him two years ago. When I was discouraged, he believed in me yet still gently nudged me to pay careful attention to my educational vision. His concern for my intellectual companionship touched me and, by introducing me to Walter Stroup, he initiated a fruitful collaboration.

Mitchel Resnick has been a source of many a stimulating conversation. He has taught me to view the world through the lens of emergence and a powerful lens it is. Our collaborations have been a source of growth and joy.

Aaron Brandes has been a companion - in every sense of the word. As my colleague, my officemate, my friend, he has always been there. Without his support - emotional, intellectual and physical, I would have been immeasurably poorer. His pouring over thesis drafts at the last possible hour was a labor of love.

Idit Harel, also, was a great support. She gave generously of her time to help me to sort out the academic world in which I had landed. Her amazingly quick and insightful comments on my writing kept me going when times were rough.

In addition to all these, I would like to thank the many others that have helped me during my years at MIT:

The teachers of the Hennigan School especially Gilda Keefe; the teachers in the SWL Workshop, especially Maria Lizano, Lynn Marshall, Lois Oxman, and Kip Perkins, and; members of the epistemology and learning group -- Edith Ackerman for tuning me in to the nuances of Piaget; Mario Bourgoïn for technical help; Amy Bruckman for the health of my fish; Chris Dagnon for speedy and faithful transcriptions; Sarah Dickinson for a fresh and funny perspective; Michelle Evard for comments on drafts of this thesis; Aaron Falbel for Illich; Greg Gargarian for making me laugh; Wanda Gleason for making the office a much more pleasant place; Nira Granott for asking real questions; Paula Hooper for honesty, Summermath and hugs; Yasmin Kafai for a collaboration in the air; Katherine Lefebvre for poetic inspiration; Kevin McGee for a fruitful collaboration on Recursion, Chaos and Fractals; Jayashree Ramadas for caring about the inquiry; Judy Sachter for honesty and treasure chests; Randy Sargent for being interested in my puzzles and interesting me in his responses; Ricki Goldman Segall for warmth,

Geertz and hypertext; Alan Shaw for a great smile and fun conversations; Carol Sperry for a Lithuania connection; and Carol Strohecker for psychological insights and great dinners.

Marvin Minsky and Daniel Dennett exerted strong influences on my intellectual growth. Marvin's humor kept me laughing through many a winter night. Dan's fundamental philosophical questionings of my assumptions prevented me from overlooking and overstepping.

I also wish to thank the "LME people" who provided me with a sustaining community outside of MIT. Richard Noss, Liddy Neville, Celia Hoyles and David Williams were especially supportive. I also benefited greatly from conversations with Marion Barfurth, Kathy Crawford, Andi diSessa, Laurie Edwards, Paul Goldenberg, Brian Harvey, Uri Leron, and Rina Zazkis.

Three friends have helped me greatly throughout my MIT career. David Frydman stayed up all night several times to hear me talk through my ideas, sat beside me as I wrote, and was unstintingly generous. Josh Mitteldorf never ceased to remind me that I could "do it" if I wanted to but that it may not be worthwhile. Nevertheless, his last minute modem transfers filled with corrections to my thesis text were done meticulously and with probity. David Rosenthal thoroughly read every word I wrote and continued to think it was great. He buoyed me countless times when my spirits flagged. His hour-long calls from Japan kept me going as deadlines came and went.

Walter Stroup, though of relative short acquaintance, has been a wonderful collaborator and I feel sure our collaboration will continue. His support and suggestions for my defense were invaluable.

Many other friends helped, cheerleaded, and coaxed me along the way. I'd like to thank Ellie Baker, Franklin Davis, Gary Drescher, Roger Frye, Karen Gladstone, Phil Greenspun, Fausta Hammarlund, Stuart Kelter, Doug MacRae, Arnie Manos, Chris Martell, Brij Masand, Aja, Howie, Jonah and Sue Morningstar, Patricia Myerson, Rose Ann Negele, Jenny Patterson, Alan Robinson, Ken Rodman, Eric Roberts, Kathy Shorr, Janet Sinclair, Dan Towler, Joanne Tuller, Natalie & John Tyler, Michael Vickers, and Leora Zeitlin.

The interviewees in this study were the best a researcher can hope for -- interested, committed, reliable and fun.

Carol Gilligan helped me understand more deeply than ever before how profound are a person's actual words. This is a lesson I will continue to learn. I wish I had translated this lesson better in the present work. Mark Tappan gave generously of his time and reminded me always of the moral dimension of educational research.

Deborah Schifter and Jim Hammerman at Summermath helped stoke the fire.

My gratitude to the poetry series folk: Pam Alexander, Linda Conte, Dave Kung, Ben Lowengard, Joseph Stampleman, Steve Tapscott, and Ruth Whitman. They made sure the muses visited MIT; and to Friday grads who nurtured the inner beauty.

Thanks to Archie Ammons and Mary Oliver whose poetry kept me sane.

My mother, Sarah Heller Wilensky, taught me to ask the deep foundational questions and my father, Mordecai Wilensky, reminded me not to overlook the historical facts in attempting to answer them. Their support was unflagging, but more importantly, both of them instilled in me a deep love of knowledge for its own sake - a gift of incomparable value.

Finally, none of this work would have come to fruition without the constant support of Donna Woods. She believed in me always, even when I was sure my work was completely without value, stayed up with me when I was anxious, took over household responsibilities when I was overwhelmed by work, and she did all this cheerfully, without complaint. I can only hope I will have the opportunity to do the same for her.

“Only Connect”

E.M. Forster

Preface

Throughout this thesis, pseudonyms have been substituted for the actual names of the participants in the research.

The linear format of a book is not ideally suited to convey the multiple connections inherent in this research. A hypertext-like medium would have been better, but universities are not quite yet at the stage where a hypertext document would be acceptable as a thesis. I have endeavored, therefore, to “linearize” the ideas. I hope not too much has been lost in the process.

Table of Contents

Abstract.....	2
Acknowledgments.....	8
Preface.....	12
Overview of the Thesis Document.....	16
Chapter I - Introduction.....	17
Chapter II - A New Paradigm for Thinking about Mathematics	26
Epistemology of Mathematics - a Short History.....	28
The Ontological Status of Mathematical Objects	28
The Formalist Response to the Crisis	31
The Failure of Formalism	33
Post-modern View of Mathematics and Science	33
Proof.....	34
A New View of Proof - Connected Proof.....	37
Proof and Pedagogy	39
The Computer Comes to Mathematics.....	40
New Paradigm, New Pedagogy	42
Chapter III - Related Work	43
School Math	43
Constructionism	46
Constructivist Alternatives to School Math.....	49
Chapter IV - Concrete Learning.....	52
Concretion.....	53
Standard Definitions of Concrete.....	54
Critiques of the Standard View.....	55

Towards a New Definition of Concrete	57
Consequences of the New View	58
Mathematical Concepts are Societies of Agents.....	60
Paradox.....	64
Concepts are Messy	66
Covering up is Damaging	68
Chapter V - Fractional Division.....	70
Fractions Setting	70
Discussion.....	71
Chapter VI - Recursion	77
Recursion Setting	77
The Workshop.....	78
Discussion.....	80
Chapter VII - Probability: Settings and Rationales.....	88
Why Probability?	88
The Subject Everyone Loves to Hate.....	88
Unresolved Philosophical Underpinnings.....	89
Epistemology of Probability	89
Tversky & Kahneman.....	93
Emergent phenomena.....	98
Probability Setting.....	101
Chapter VIII - Probability Interviews	104
Normal Distributions - Why is Normal Normal?.....	105
Divorce.....	110
Monty Hall	120
Summary	132
Circle Chords	134

Envelope Paradox	144
Dale	145
Barry.....	148
Mark	152
Gordon	157
Additional Starlogo Programs.....	166
Binomial and Normal Distributions Revisited.....	166
Particles in Collision - *LogoGas.....	172
Chapter IX - Conclusion	180
Summary and Discussion.....	180
Response to Tversky and Kahneman.....	183
Obstacles to Learning Probability.....	189
Future Directions.....	192
Appendix A - Probability Interview.....	194
References.....	201

Overview of the Thesis Document

Chapter I is an introduction to the research -- a summary of the research problem and the methods used to approach it. Chapter II gives an account of a changing paradigm of discourse about mathematics and the role of proof within that paradigm. Chapter III gives an outline of school mathematics and describes constructivist theory and practice in response to traditional mathematics education. It is broken up into three sub-sections: school math, constructivist psychology/constructionism and alternatives to school mathematics education. In Chapter IV, I investigate the claim that mathematics learning should revalue the concrete and propose new theoretical descriptions of concrete and abstract knowledge. Chapter V describes a pilot study which investigates division of fractions. Chapter VI presents another pilot study, a description of a workshop on recursion. In Chapter VII, I describe the methods used to collect the data for the probability study and present extensive motivation for the selection of probability as an area in which to try a connected mathematics approach. Chapter VIII presents the probability interviews and weaves together theoretical material and interview text. Chapter IX summarizes the results of this research, draws conclusions for mathematics learning and pedagogy, uses data from this research to respond to the claims of Tversky and Kahneman, and examines future research directions.

Appendix A contains the questions in the probability interview.

Chapter I - Introduction

There is an endemic dissonance between the discourse about mathematics led by philosophers and logicians and picked up by educators and mathematicians and the actual practices of the creative mathematician.

Recently, the discourse about mathematics has begun to change bringing the two views into greater harmony. Instead of viewing mathematics as part of the rationalist tradition in which truth and validity are primary, a new paradigm is emerging, a view of mathematics through an interpretive framework in which meaning making is primary.

Indeed, once we adopt an interpretive framework, the history of mathematics, conspicuously absent both in our schools and in our mathematical talk, takes an important role. This history is replete with examples of mathematicians of great distinction arguing about the fundamental meanings of mathematical objects, operations and ideas.

Examination of this history reveals the path to our current mathematical conceptions was filled with argument, negotiations, multiple and competing representations. Mathematical objects are seen to be highly ambiguous, capable of multiple constructions and interpretations. Choosing among these many constructions is a major activity of the mathematician and these choices are governed by how the mathematical objects are connected to other objects both mathematical and non-mathematical. Mathematics is messy and not the clean picture we see in textbooks and proofs.

But our educational practice remains locked within the old paradigm. With the notable exceptions of educators such as Ball (1990b), Confrey (1990), Harel (1988), Hoyles & Noss (1992), Lampert (1990) and Schoenfeld (1991), the mathematics classroom of today is not recognizably different from the classroom of one hundred years ago. This is especially so at the higher educational levels. Many of us are all too familiar

with the “litany” of definition/theorem/proof chanted day in and day out in the mathematics classrooms of our universities. This image of mathematical practice portrays mathematics as a dead subject – inquiry is unnecessary because our concepts have been formally defined in the “right” way and our theorems demonstrated by linear and formal means. This approach is intrinsically formal – using linear deductive style, stipulative definitions, and non-contextual proofs. Very little effort is put into motivating the mathematical ideas, linking them to other ideas both mathematical and non-mathematical, or any discussion of the "meaning" of the mathematics being taught.

Those rare students who, despite this antiquated approach, do manage to succeed in understanding the mathematics, do so outside of the classroom. They are then pointed to as success stories – validating the myth of the gifted few – only those elite individuals who are natively gifted can really understand mathematics. The purpose, then, of the traditional mathematics class is to discriminate between those who are "called" and those who "really ought to be doing something else."

However, those students who do "get it", that is understand the mathematics they are taught, employ a variety of techniques and approaches in order to do so. These techniques are not taught in the classroom, nor is the kind of approach to mathematics embodied in these techniques even hinted at – most learners are unaware of its existence. And when they are engaged in an activity that involves building mathematical intuitions and connecting different mathematical ideas, when they are acting most like mathematicians, they often report that they are doing something other than mathematics. But, as we shall see, even those students who are successful at school mathematics, nevertheless encounter unexpected perplexity – confusions about the meanings of basic concepts and operations they thought they understood. What is needed is an approach to mathematics learning that addresses fundamental epistemological issues, encourages construction and negotiation of meaning – a new pedagogy for a new paradigm.

With the notable exceptions mentioned above, researchers in the mathematics education community are also embedded in the traditional mathematics paradigm. When analyzing the difficulties learners have in mathematics exercises, researchers often catalogue syntactic errors, rules that learners fail to follow such as: Johnny adds fractions by adding their numerators and denominators instead of making a common denominator. No attention is paid to what Johnny thinks a fraction is. Instead, these educators prescribe more practice in applying these rules, or perhaps computer aided instruction programs which will help Johnny drill.

More recently, in response to the changing paradigm, some researchers have begun to describe learners' difficulties as false theories or misconceptions, such as: Maggie thinks you can't subtract a bigger number from a smaller, or divide a smaller number by a bigger. The prescription offered here might be creating a simplified computer environment in which Maggie can play around with numbers, but is constrained to operations that are mathematically valid. In this way she will construct the true conception of, say, division instead of a misconception.

While the second view is an improvement on the first, both of these viewpoints share the assumption that there is a "right" way to do or understand the mathematics and Johnny and Maggie are not there. Their goal is the transmission of this "right way". But in throwing out the "bathwater" of error, they lose the "baby" if the learner never enters the messy process of negotiating meaning, constructing different representations and critiques of these representations. If we deprive learners of this opportunity, we strip mathematics of its essential character and deprive them of real mathematical experience.

We also deprive them of respect. As we have mentioned, mathematicians throughout history have constructed many different meanings for mathematical concepts and argued their relative merits. If mathematicians of distinction needed to go through this process in order to make sense of the mathematics, why do we expect that the learner

will take our conceptions on faith? We respect the learner by viewing her as a mathematician in a community that is still negotiating the meaning of the new concept.

Since so much of mathematics class is occupied with learning formal definitions and techniques, rarely are we encouraged in mathematics classes to think qualitatively about mathematics or to develop our mathematical intuition. Some researchers and educators seem to believe that our mathematical intuitions are fixed and consequently the only method by which we can progress mathematically is to learn mathematics in a formal way.

But in this thesis we shall make an analogy between the mathematical development of adults and children's development as observed in the classical conservation experiments of Piaget. Children develop from "intuitively" seeing liquid volume as changing when its appearance changes, to seeing it as equally intuitive that liquid volume is not affected by the shape of the container. Similarly, adult mathematical learners can find new mathematical objects counterintuitive at first but, through development of their mathematical intuitions, they can come to see them as intuitive.

We are accustomed to thinking that this kind of development happens only in children I shall argue in this thesis that this kind of development is characteristic of gaining deeper understanding and is a preferable way to make mathematical progress. But just like conservation in children, building mathematical intuitions or concrete relationships with mathematical objects takes time. Simple instruction does not dissuade the child nor the learner from their current intuitions. They must be in an environment where they can construct new meanings and structures for themselves.

The broad goals of this research are to: understand the obstacles to learners mathematical understanding, understand the failures of school mathematics to overcome these obstacles and to create environments which facilitate deeper mathematical

understanding. Earlier in this introduction, I introduced important themes related to the first two goals. I will now describe the approach I developed to meet the third goal and go on to motivate the choice of probability as the major domain of study.

As has been mentioned, some researchers in the mathematics education community (e.g. Papert, 1972, Lampert, 1990, Ball, 1990b, Steffe & Wood, 1990; Streefland, 1991; Von Glasserfeld, 1991) have begun describing and constructing mathematics learning environments consistent with the new paradigm. But, up to now, they have focused primarily on younger children doing elementary mathematics. Very little has been done with adult learners and advanced mathematics where the methods are, if anything, more rigidly *instructionist* (see Papert, 1990), provide less opportunity for self-directed exploration and conceptual understanding than do classes in the lower grades.¹ In this thesis, I develop an approach to learning mathematics that emphasizes learners' negotiation of meaning. I employ this approach with relatively advanced adult learners of the mathematics of probability. I call this approach "connected mathematics".

Among its key features are:

- Explicit discussion of the epistemology of mathematics, as in: What does the notion of probability mean? What validates the mathematical knowledge we are gaining? How does this proof prove anything?
- Mathematical ideas multiply represented. Never just one definition of a mathematical notion.

¹ Some educators have conceded that perhaps children and mathematical novices might "need" informal approaches to mathematics, but once someone has arrived at advanced mathematics, these informal "crutches" are no longer needed.

- A collection of especially interesting problems, puzzles², or paradoxes which highlight the key ideas and distinctions in a mathematical domain. Often these take the form of opposing arguments, each by itself compelling, which appear to contradict each other.
- Mathematical ideas linked to other mathematical ideas and related cultural ideas.
- An environment where conjectures can be made and tested resulting in engaging feedback to the learner. The environment can be social -- involving a community of learners -- or computer-based. In either case, it must be a learning environment that promotes user designed constructions and conjectures in an open and non-constrained way.

In this thesis, my major mathematical content area of focus will be probability. Probability (which in this thesis I will use broadly as an abbreviation for the mathematical areas of probability and statistics) is an important domain to study for several reasons. It is traditionally regarded as difficult for learners. Learners often say it is alienating, yet rely almost exclusively on formal and quantitative techniques while expressing confusion as to what the calculations they are performing or evaluating mean. However the development of qualitative and intuitive understandings of probability would be very useful for everyday life. These considerations also make it a good testbed for a connected mathematics approach.

Probability is an area of mathematics that is particularly vulnerable to the neglect of the intuitions. In their now classic work Tversky & Kahnemann (1982) document the persistent errors and “misconceptions” that people make when making probabilistic

² In speaking of logic as it enters into pure mathematics, Bertrand Russell once remarked: "A logical theory may be tested by its capacity for dealing with puzzles, and it is a wholesome plan, in thinking about logic, to stock the mind with as many examples as possible".

“judgments under uncertainty”. Among these errors are systematic misconceptions about probabilities of conjunctions, inverse probabilities, updating probabilities after getting evidence and perceiving pattern in random data.

Tversky & Kahneman speculate as to the origin of these systematic biases in people's assessment of likelihoods. They theorize that people's mental resources are too limited to be able to generate probabilistically accurate judgments. Consequently, people are forced to fall back on computationally simpler “heuristics”.

Many interpreters of Tversky & Kahnemann have taken their results to mean that human beings are simply not capable of thinking intuitively about probability. Instead they claim, in order to do probability we should distrust our intuitions and rely on formalisms and formulae. This claim has become very influential. Given the prevalence of this "quasi-innatist" view, it would be of great interest if it could be shown that these systematic probabilistic biases could be transformed into good probabilistic intuitions by a suitable learning environment.

A conjecture of this research is that people's difficulties with probability are of the same kind as their difficulties with other mathematical ideas. That is, these difficulties are not just errors in calculation or procedure but reflect a failure in understanding. This failure is a result of learning environments, both in school and out, inadequate for the acquisition of good mathematical intuitions. In many cases they are the result of brittle, formal, rule-driven instructional techniques that keep knowledge fragmented -- disconnected from its sources of meaning.

In support of my research goals, I have conducted interviews with adult mathematical learners. Interviews were designed to both elicit the structure of the learners' current knowledge and to provide them with experiences of connected mathematics. Interviews touched on three domains of study -- the primary study of

probability as well as two pilot projects in fractions and in recursion. These interviews were structured so as to flow naturally, like conversation, but were directed to various topics and issues of concern. Some of these conversations took place in the context of a project chosen by the interviewee using a computer-based learning environment -- the language *Logo³ (pronounced Starlogo), extended and adapted for exploring some of the mathematical ideas in the interviews. In these projects learners collaborated with me to design and implement *Logo programs for exploring some mathematical idea that they wanted to understand better. I interviewed them both during the process of design as software designers and, upon completion of the programs, as users of their own software. Through these interviews, I have elucidated the structure of the learners' present mathematical knowledge and shown how that knowledge develops and understanding deepens when a connected mathematics approach is used.

I have not conducted pre- and post- tests in order to evaluate the learning of the interviewees. Rather, in large measure, I have left the interviews to “speak for themselves”. It is my hope that you, the reader, will put yourself into the interview questions and conduct your own inquiry. In so doing, I trust that the reader will recognize the “good mathematics” that is done in these interviews. In this way, I am in a position not unlike that of the mathematician G.H. Hardy (1967) who attempts a defense of mathematical activity - “why is it really worthwhile to make a serious study of mathematics?” As Hardy concludes, this defense cannot rest on the utility of the mathematics to some future enterprise, but instead depends on the recognition of the beauty and seriousness inherent in the mathematics.

This thesis makes several contributions:

³ Originally designed by Resnick & Wilensky. Coded by Resnick (1992) for his doctoral dissertation. - extended and enhanced by Wilensky as part of this doctoral work.

I show how certain historical developments in the philosophy and practice of mathematics have been inappropriately imported into the practice of mathematics education and have resulted in brittle fragmented understanding. A critique of the traditional notions of abstract and concrete knowledge is used to propose new theoretical characterizations of the abstract and concrete. The account of this concretization process provides a more detailed and finely textured picture of the "messy" nature of mathematical understanding. Understanding how to concretize our mathematical knowledge shows us how to build and develop mathematical intuitions instead of relying on formal procedures. Several examples of connected mathematics learning environments are presented and analyzed in detail. Through the interviews of learners in these environments, we see how discomfort about the epistemological status of mathematical entities presents significant obstacles to deeper mathematical understanding. In particular, we see how discomfort with the concepts of randomness, probability and distribution prevents learners from a deeper understanding of and connection to probability. We see how these obstacles can be overcome through learners actively constructing and negotiating meanings for fundamental concepts of probability. Through a "connected mathematics" approach, which explicitly addresses these epistemological issues, learners develop more sophisticated probabilistic intuitions, build concrete relationships with the probabilistic concepts and gain a deeper mathematical understanding.

Chapter II - A New Paradigm for Thinking about Mathematics

Mathematics has long been viewed as the pinnacle of the rationalist tradition. In this chapter I will argue that a new paradigm for discourse about mathematics has begun to emerge. I begin by situating this new paradigm in the context of other post-modern attacks on rationalism. The historical and epistemological roots of the problems with rationalism in mathematics are then explored. The increasing role of the computer in mathematics will be seen as contributing to the downfall of earlier concepts of proof but, as in so many other disciplines into which it has entered, the effect of the computer has been contradictory -- pushing mathematical practice both in more formal and more intuitive directions.

Rationalism has come under attack in our post-modern era. It has been criticized by hermeneutic critics such as Packer & Addison (1989). Among the differences, they cite in their critique, between the rationalist perspective and the hermeneutic (or interpretive) perspective, four dimensions of difference are salient:

	Rationalism	Hermeneutics
Ground of Knowledge	Foundation provided by axioms and principles	Starting place provided by practical understanding: articulated and corrected.
Character of Explanation	Formal, syntactic reconstruction of competence	Narrative accounts - a reading of the text
Relationship to Researched	Detachment: abstraction from context	Familiarity with practices: participation in shared culture
Justification of Explanation	Assess correspondence with knowledge of competent person	Consider whether interpretation uncovers an answer to its motivating concern.

Historians and Sociologists such as Latour (1987) have critiqued the traditional scientific method and emphasized that science can only be understood through its practice. Feminist critics such as Gilligan (1986) and Belenky et al (1986) have criticized rationalism from a psychological perspective. They show how the formal knowledge of rationalist procedures has created a “separate knowing” which has alienated many and women in particular. They propose a revaluing of personal knowledge and what they call “connected knowing”⁴ - a knowing in which the knower is an intimate part of the known. The construction of the known through the interpretation of the knower is fundamental and unavoidable for non-alienated understanding. The concept of “connected knowing” will be explored in more detail in chapter IV, in the discussion of “concretion.”

While these critiques of the rationalist (and empiricist) traditions have made serious inroads into the hegemony of the dominant epistemology, the calls for interpretive frameworks have largely focused on the social sciences and to a lesser degree on the natural sciences. To a great extent, mathematics has still escaped the full glare of this critique. However, this is beginning to change -- we are witnessing the emergence of a new paradigm for thinking about the character of the mathematical enterprise.

In order to understand the nature of the paradigm shift, it will be helpful to situate this discussion in the history of mathematical epistemology. In the next section, I will sketch the development of mathematicians’ thinking about their own practice. We shall see that mathematics, which for so long was seen to be about the search for truth, or the discovery of the properties of Platonistic entities, derived its meaning from the “reality” of the Platonic “world”. Crises in mathematics arising in the 19th and 20th centuries created problems for this view. Two classes of responses to these crises have emerged. The first, and earliest, a formalist agenda whose aim was to eliminate ambiguity, interpretation and meaning from mathematical symbols. This response, which arose from the desire to rescue mathematics from contingency and secure

⁴ Or as Gilligan refers to it - the “web of connectedness”.

the indubitability of its foundations, greatly influenced the development of a pedagogy for mathematical instruction (see e.g. Kleiner, 1990). However, a second response, motivated by the preservation of meaning in mathematical statements, has begun to take hold. This latter view sacrifices the idea of a realm of mathematical entities, and places the emphasis on the construction of these entities by a community of mathematical meaning makers. In so doing, it replaces the centrality of truth and validity in mathematics by the interpretation and negotiation of meaning. Seen from this perspective, we shall see that mathematical proof, which was seen to be the foundation of mathematical validity, becomes instead a method for connecting new knowledge to our personal knowledge web and thus imbuing it with meaning.

Epistemology of Mathematics - a Short History

While mathematical activity has been going on as far back as is recorded, the idea of mathematical demonstration apparently only emerged in about 600 B.C.E. Thales of Miletus is credited with bringing geometry from Egypt to Greece. Although the Egyptians had an implicit geometry in their surveying practice, as far as we know they did not prove any geometrical theorem. Thales is credited with having proved a number of elementary theorems including that the base angles of an isosceles triangle are congruent. By about three hundred years later, Euclid had axiomatized plane geometry so that all theorems could be deduced from five "self-evident" postulates.

The Ontological Status of Mathematical Objects

Since the beginning of this proof era of mathematics, mathematicians have also concerned themselves with the epistemology of mathematics. What kinds of things are mathematical objects? What justifies mathematical arguments? Are mathematical truths *a priori* (prior to experience) or are they derived from our everyday experience?

Plato (see Hackforth, 1955) gave one of the first answers to these questions and to this day most mathematicians regard themselves as Platonists with respect to the reality of mathematical objects (see Davis and Hersh, 1981). Plato's theory was that the world of everyday objects is just a shadow world. Beyond the shadows lies a real world of transcendent objects in which live "ideal forms". Ideal forms are unique, immutable, and embody timeless truths. So, for example, while in the passing shadow world we may see many instances of stones, in the transcendent real world, there is one stone form of which all everyday stones are shadows. In this real world, mathematical objects are first class citizens. In the everyday world, one can find two apples, but the concept-object "two" does not exist at all except in the world beyond the senses. Central to this view is the notion that ideal forms are apprehensible, but unlike everyday objects which are apprehended through the senses, concept-objects must be apprehended by a process of pure ratiocination, a direct intellectual connection to the transcendent world.

This view has been elaborated on through the rationalist tradition and falls under the heading of what is now called Mathematical realism. One consequence of this view is that our intuitive notions of mathematical objects must be coherent since they reflect the real objects in the transcendent world. Mathematical knowledge is certain according to this view and the foundation of its certainty lies in the real relationships that obtain between real mathematical entities.

One implicit consequence of this view is that mathematics is true in the everyday world since that world is derived from the ideal world. Thus Euclidean Geometry was thought to be necessarily true of the world.

Throughout the Middle Ages, mathematicians tried to derive Euclid's fifth postulate (the so-called "parallel postulate") from the other four. The fifth postulate was not as self-evident as the other four and this was bothersome to the sense of necessity that geometry was supposed to have. The fifth postulate however resisted all attempts to derive it from the other four or from yet another more "self-evident" postulate. Finally, in

the nineteenth century, Lobachevsky, Bolyai, and Riemann were able to construct models of geometry in which Euclid's first four postulates were true, but the fifth was not. In Riemannian geometry, through a point not on a given line, one cannot draw a line parallel to the given line. The sum of the angles of a triangle is always more than two right angles, and the ratio of a circle's circumference to its diameter is always less than π .

If however, Euclidean geometry expressed truths about the world, then these alternative geometries would have to be proved false. Naive Platonism was dealt a blow when relative consistency theorems were proved showing that if Euclidean geometry is consistent so are the alternative geometries.

The meaning of the statement "Euclidean geometry is true," epistemologists had said, is that the postulates are true when the geometrical terms are interpreted in normal fashion. But what does it mean to interpret the terms in normal fashion? For example what does the term "straight" mean in "straight line"? One way of telling whether a line segment is straight is to see if you can find any measuring rod which is shorter and connects its end-points. Or one could sight along it - this would mean that straight is the path light takes, or one can define it as the path on which a stretched rope lies when it is under high tension. Another way to define it is to say that a straight line is the shortest distance between two points. Which of these definitions captures the essence of straightness? At the time, all criteria seemed to agree, so this question wasn't carefully addressed and Euclidean geometry was assumed to be true under standard usage of the word "straight".

When Einstein demonstrated the theory of general relativity, a further question arose as to the truth of Euclidean geometry. The truth of Euclidean geometry then appeared to depend on empirical observations - how does light travel, what does the rod measure? When Einstein showed that light does not travel in a Euclidean line, did he show that Euclidean geometry is false?

Some mathematicians do indeed take this view. According to them, our notion of straightness is empirically based and comes from our experience with light and rods. Hence relativity (by showing that light triangles have more than 180 degrees) proves that space is Riemannian not Euclidean.

But there is another view that says, no, straightness is defined in terms of shortest distance where distance is determined by an Euclidean metric. Under this view, space is not Riemannian, it is still Euclidean, it is only that gravity (like temperature, pressure, refractive medium) bends rods and light so that they are no longer straight.

Einstein himself espoused the former view as expressed in his famous quote: "As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality".

The Formalist Response to the Crisis

To escape such questions about the truth of geometry, some mathematicians began to tighten up Euclid's postulates so that they could be put on a more secure footing -- turned into a formal system. Thus was born the *logician* program to reduce mathematics to logic. According to this view, Euclidean and Riemannian geometry are equally valid -- they state nothing more than: if their postulates are true then their theorems follow logically. Some such as Poincaré (1908) carried this further into so-called "conventionalism" and said that the truth of Euclidean geometry vs. Riemannian geometry was a matter of convention -- how we stipulate the definitions of our language.

The logicist program sought to formalize mathematics so that there would be a purely syntactic procedure for determining the truth of a mathematical statement. To do so would require that all referent terms in a mathematical expression remain uninterpreted, without meaning apart from their formal stipulations. A term in such a system is a "dead" object defined only in terms of its relations to other dead objects. A

proof, in this view, contains all the information necessary for believing and asserting a theorem.

Parallel to the controversy about alternative geometry, another group of mathematicians was engaged in trying to formalize arithmetic. Peano and Dedekind had provided axiomatizations of the natural numbers and in 1879 Frege was the first to articulate the aim of reducing all of mathematics to logic (Frege, 1953). However, Frege's formulation relied on the intuitive notion of class. Russell showed that this formulation was inconsistent. In the intuitive notion of class was the idea that any describable collection could be a class. In particular, the class consisting of all classes is just another class. Here already the intuitive notion becomes muddy since at once we have this class being as big as imaginable and at the same time it is just one member of the large collection of classes which are its members. Russell formalized this paradox by considering the class of all classes that are not members of themselves. Call this class *C*. Is *C* a member of itself? If it is, then by the definition of *C*, it's not. If it's not, then by the definition of *C*, it must be⁵. Russell resolved this paradox by assigning types to classes and classes that contain other classes are of higher type than their members. But the cost of this move was high. The intuitive notion of class which seemed perfectly innocent had to be replaced by the unintuitive notion of class of type "n". This was another blow to Platonism. What happened to the ideal form of class? Are we to suppose that there are infinitely many ideal forms corresponding to a class of each type?

Logical positivists (e.g. Carnap, 1939; Ayer, 1946) responded to this crisis of meaning by adopting an extreme form of conventionalism which said that mathematical statements do not need any justification because they are true by fiat, by virtue of the conventions according to which we stipulate the meanings of the words we choose in

⁵ Referring to the significance of this paradox, Frege said: "Arithmetic trembles".

mathematics. In so doing, they removed mathematics both from doubt and from any possibility of making personal connections to, and meaning from mathematical objects.

Others, (e.g. Kitcher, 1983; Putnam, 1975) who saw that the procedure for choosing axioms for set theory and arithmetic was one of making hypotheses, deducing consequences and then evaluating these consequences, moved in the opposite direction and began to see mathematical process as much more akin to scientific hypothetico-deductive process rather than in a removed realm all its own.

The Failure of Formalism

The logicist dream of reducing mathematics to logic and Hilbert's derivative program to prove mathematics consistent were dealt a final blow by the results of Gödel (1931). Gödel's theorems show that any formal system large enough to incorporate basic arithmetic cannot be proven consistent within that formal system. Furthermore, if the system is consistent then there are statements in the system which are true but unprovable.

Since Gödel, further unclarity in the basic notion of the integers⁶ emerged when specific hypotheses such as the continuum hypothesis⁷ were proven independent of the standard axiomatizations of arithmetic. Mathematicians were confronted with the

⁶ In the last century, old arguments about the reality of the continuum, which date back to Zeno in the early 5th century BCE, were revived. Why is it permissible to divide space up into infinitely small pieces? What legitimates making infinitely many choices in finite time? Even if we can make infinitely many choices, can we make an uncountably infinite number of choices? These questions have a crucial bearing on what the character of the so-called "real" numbers is. Mathematical Constructivists such as Kronecker asserted that the real numbers were not real at all but a convenient fiction useful for some human purposes. (As in his famous quote: "God created the integers, all the rest were created by man.") Recently, some mathematicians, notably Brian Rotman (1993), have done Kronecker one better and questioned the legitimacy of the integers themselves. By what criterion are we allowed to assert that we can count forever? The intuitions that allow us to say that addition is, for example, commutative are formed from experience with small numbers -- what legitimates the extrapolation to numbers infinitely far away? Here we see clearly that making mathematics requires acts of imagination. Which imaginings or, if you will, "mathematical dreams" are we entitled to and which are we to say are irrational? (or ir-rational?)

⁷ Which asserts that there are no infinities "larger" than the integers yet "smaller" than the reals.

situation of having to explicitly choose whether to adopt an axiom by criteria such as plausibility, simplicity, and productivity in generating good mathematics.

These developments in formal mathematics permanently halted the logicist agenda. It would seem that mathematics is beyond logic and its truths are not formalizable.

Post-modern View of Mathematics and Science

In the latter half of the twentieth century, developments in the philosophy and history of science have tended to push mathematicians into seeing mathematics as more continuous with natural science. Works by Popper (1959), and more decisively by Kuhn (1962) have shown that the progress of science is not linear and hypothetico-deductive (see Hempel, 1963) as the logical positivists have claimed, but rather science proceeds by revolution, i.e. by changing the meaning of its basic terms. New theories are not incremental modifications of old theories, they are incommensurate in that what they posit as the basic entities of the world are fundamentally incompatible. Furthermore, says Kuhn, by presenting science as a deductive system devoid of history, positivists have robbed us of the fundamental character of science – the negotiation and construction of its basic entities. Mathematicians such as Polya (1962) and Lakatos (1976) have shown that mathematical development has been similarly mischaracterized. By placing such a strong emphasis on mathematical verification and the justification of mathematical theorems after their referent terms have been fixed, mathematics literature has robbed our mathematics of its basic life. Mathematics, according to Lakatos, is a human enterprise. Advances in mathematics happen through the negotiation of a community of practitioners. Moreover, the development of mathematical proofs is not linear, but rather follows the “zig-zag” path of example, conjecture, counter-example, revised conjecture or revised definition of the terms referred to in the conjecture. In this view, mathematical meaning is not given in advance by a transcendent world, nor is it stipulated in an

arbitrary way by conventions of language; rather, mathematics is constructed by a community of practitioners and given meaning by the practices, needs, uses and applications of that community.

Proof

The notion of proof has also undergone change. If there is one idea which traditionally separated mathematics from the empirical sciences, it was the notion of proof. Mathematical proof was seen to be an indubitable method of ascertaining knowledge in contrast to natural induction, which is required by empirical pursuits. However, a number of factors, among them the use of computers in mathematical proof, have occasioned a revision of both the concept and the significance of proof.

In a stimulating reflection on the four color map problem, Tymoczko (1979) suggests that the “proof” of the four color map “theorem” (henceforward 4CT) by Appel, Haken and Koch (1977) (which required a computer to supply the proof of a crucial lemma which had too many cases to be checked by hand) necessitates a significant change in the mathematical notion of proof. Tymoczko gives three reasons why mathematical proof is thought to be a reliable method of gaining knowledge. These three characteristics of proof are :

- Proofs are convincing

“It is because proofs are convincing to an arbitrary mathematician that they can play their role as arbiter of judgment in the mathematical community.”

- Proofs are surveyable

“A proof is a construction that can be looked over, reviewed, verified by a rational agent. We often say that a proof must be perspicuous, or capable of being

checked by hand. It is an exhibition, a derivation of the conclusion, and it needs nothing outside of itself to be convincing. The mathematician *surveys* the proof in its entirety and thereby comes to *know* the conclusion.”

- Proofs are formalizable

“A proof, as defined in logic, is a finite sequence of formulas of a formal theory satisfying certain conditions. It is a deduction of the conclusion from the axioms of the theory by means of the axioms and rules of logic..... Formal proofs carry with them a certain objectivity. That a proof is formalizable, that the formal proofs have the structural properties that they do, explains in part why proofs are convincing to mathematicians.”

As Tymoczko says, the first feature of proofs is centered in the anthropology of mathematics, the second in its epistemology and the third in the logic of mathematics.

The question he then raises is: are surveyability and formalizability two sides of the same coin?

Some intuitionists such as Brouwer (1981) and Heyting (1966) deny that all proofs can be captured by formal systems. Gödel’s result shows that regardless of your belief in the formalizability of every proof in some formal system or other, not all proofs can be formalized within the *same* formal system. In every formal system, one can find a surveyable proof which is not formalizable in the system. As René Thom (1971) has described it: “Formalizability is a local characteristic of proofs, not a global one.”

But what about the other side of the coin -- are all formalizable proofs surveyable? Clearly the answer to this is no, for in any sufficiently rich formal system there must be proofs too long for any mathematician to survey them⁸. But Tymoczko

⁸ A group theory proof I tried to survey while an undergraduate comes close to the limit of some mathematicians’ surveyability threshold -- see (Feit & Thompson, 1963).

argues that, up until the advent of the computer, all formalizable proofs came to be known by mathematicians either by being surveyable or through informal surveyable arguments that establish the existence of a formal proof.

Most mathematicians would say that the 4CT does have a formal proof. However, the proof of the 4CT is not surveyable by mathematicians -- it depends on the reliability of computers. Thus empirical considerations have been introduced explicitly into mathematical arguments. Our belief in the proof of the 4CT depends on our belief in the reliability of computers. Note that effect of this change is not solely in the direction of supporting the new paradigm. The change works in two different directions: it weakens the rationalist view that mathematical truth is to be found solely in the head of the individual mathematician, but it also leads some mathematicians to accept proofs which are, in a sense, even more formal – that is, beyond the intuitive apprehension of the mathematician. This shake up in the “sacred” notion of proof has exposed divisions in the mathematical community. Some mathematicians, more comfortable with the introduction of empirical concerns into the concept of proof have pushed further, challenging the status of mathematical concepts and offering new probabilistic mathematical concepts. Rabin (1976), for example, uses a number-theoretic algorithm to find large primes. “Primes” found by this algorithm aren’t deductively shown to be prime. Rabin defines the concept “prime with probability p ” and declares a number to be prime if it is prime with probability “ p ”, where “ p ” is sufficiently close to 1. These new notions of mathematical concepts and validity lead other mathematicians to give up on proof as a secure foundation for mathematical truth and take them back to the skeptical Wittgensteinian (1956) view that proofs are just those arguments that are convincing to mathematicians.

A New View of Proof - Connected Proof

The relations revealed by the proof should be such as connect many different mathematical ideas. - G. H. Hardy

As we have just seen, the mathematical notion of proof thought to be the sine qua non of mathematical culture has evolved over time and is again undergoing transformation in response to the computer's entry into mathematics. I shall argue that in contrast to claims that proofs are just those arguments which are convincing to mathematicians, the function of proof is actually both much more and much less than a convincing argument.

Proofs are much less than convincing arguments in that quite often and perhaps most of the time, it is not proofs which convince us of the truth of a mathematical claim. In this age of computer tools this has become increasingly obvious. In using tools such as Geometric Supposer (Schwartz & Yerushalmy, 1987), Cabri Géometre (Baulac et al, 1992) Geometer's Sketchpad (Jackiw, 1990) or Geometry Inventor (Arbel et al, 1992), one becomes convinced of the truth of many geometric relations (e.g. the medians, angle bisectors, and heights of a triangle meet at a point) by simply varying enough parameters and watching the result.⁹ If the prior probability attached to a particular claim is low then just watching it work out in a few "random" cases is enough to convince us that the claim is true¹⁰.

But even before the advent of computational technology, mathematical claims that were regarded as self-evident were subject to proof. Are we to suppose that Euclid got convinced of the non-intersection of concentric circles because of his proof of this theorem? Clearly not.

But without proof, empirical evidence which may convince us leaves the truth of a mathematical claim as a "brute fact". This is a very unsatisfactory way to know anything. Brute facts isolated as they are from other parts of our knowledge have scarcely

⁹ A similarly useful and powerful tool for probability is Resampling Stats (Simon, 1991).

¹⁰ See Goldenberg, Cuoco & Mark (1993) for a nice example of this. They trisect each side of a triangle, and join each trisection point to the opposite vertex. The resultant hexagon has area 1/10 that of the triangle. As they move a triangle vertex around dynamically, one can see the area of the triangle and the area of the hexagon update on the screen, yet remain in a ratio of 10:1.

any meaning for us. From this perspective then, *the important function of proof is to connect our new knowledge to other pieces of knowledge that we have.* Proofs and especially good proofs, ones that give us insight, relate new ideas to old familiar ones, situating our new knowledge in a network of “friends” where it can interact and make a contribution to what we will refer to in chapter IV as the “society of knowledge.”

While I have presented this connected view of proof as new, in fact, thoughtful mathematicians have articulated a similar view. The quote from Hardy (1967) in the epigraph to this section is one such articulation. But despite this recognition by the community of creative mathematicians, the proof as presented in the classroom and in the mathematical journal is tidied up and made so “elegant” that the messy process of discovery, the connections made and unmade are hidden from view.

Proof and Pedagogy

“There is a real conflict between the logician’s goal and the educator’s. The logician wants to minimize the variety of ideas, and doesn’t mind a long thin path. The educator (rightly) wants to make the paths short and doesn’t mind -- in fact prefers -- connections to many other ideas. And he cares not at all about the direction of the links.” - Minsky (1987).

It is interesting to note that proofs arose in mathematics to a great degree because of the need to teach. The Greeks, who first opened academies for teaching mathematics, were also the inventors of the concept of proof. Later, the greater rigor introduced into the calculus by Lagrange, Cauchy, Weirstrass and Dedekind, was also motivated by a need for organization for pedagogical purposes. More recently, Bourbaki has continued this trend of formalizing for the sake of more “efficient” teaching.

It is a common saying among educators that “if you want to understand something, teach it”. Projects like “Instructional Software Design” (Harel, 1990) in which children write instructional software to teach younger students about fractions renew this

lesson. The older children who write the software greatly improve their understanding of fractions. But, the product itself, the completed software is not likely to bestow on its users much understanding of fractions. In a similar way, the process of connecting different pieces of our mathematical knowledge which leads to formalization and proof can greatly increase our understanding, but the product, the proof itself may be inert -- hiding from us all the connections that were made, the pieces included and those deliberately excluded in order for it to come into being.

The Computer Comes to Mathematics

“The intruder has changed the ecosystem of mathematics, profoundly and permanently.” (Steen, 1986)

The advent of the computer has impacted mathematics in several dimensions. It has further weakened the hold of some firmly entrenched mathematical notions including proof. Computers have also dramatically changed the work practice of mathematicians. Finally, they have given us new metaphors and objects to think with.

Computers have unleashed a torrent of new ways of thinking about mathematics. Discrete simulations of “continuous” processes have led many physicists to make models of the world which do not postulate a continuum of infinitely dense real numbers (see e.g. Toffoli, 1990 and Fredkin, 1990). The availability of computing technology has shown us how much the mathematics of the past has been constrained by the limited calculating power of antiquated technology. Curves of non-integral dimension, the existence of which Cantor worked so hard to prove, are now ubiquitous fractals, programmed even by young children. Complex mathematical situations which were of necessity idealized as the “sum” of linear pieces, can now be simulated and reveal their non-linear complexity. Rates of change can now be seen to be a much more general notion than the linear approximation of a derivative handed down to us by pre-computational calculus. And, yes, even, the sacred notion of the infinite can be more easily challenged when we know any computer has a maximum number it can store.

Computers have become ubiquitous confederates in our daily lives. Mathematicians now use computers to aid in and relieve the burden of computation, generate data from which to make conjectures, and simulate the behavior of complex systems. The acceptance of the partnership between computers and humans in making

mathematical proofs breaks the “Fregean barrier”. Frege said that what matters in mathematics is only the context of justification not the context of discovery. But if computers are admissible in justification, how much more so in discovery. The partnership between the mathematical learner/discoverer and the computer has already transformed the culture of practicing mathematicians and will soon radically alter the mathematical learning culture.

Computers have given us new objects to think with. The computer’s capability to transform processes or relations into manipulable objects is particularly useful for mathematics learning (see Noss & Hoyles, 1992). The importance of this feature will be further amplified in chapter IV. Computer languages which make procedures into “first class” objects foster thinking of complex relations as entities, entities which just like numbers can be inputs and outputs to other procedures. As we shall see, by allowing us multiple modes of interaction with these relations, we make them concrete and bring them into closer relationship.¹¹

Computers are also becoming the dominant metaphor for explaining the human mind. Early information processing models of human cognition were inspired by the linear computations of contemporary computers. Marvin Minsky has brought the new technology of parallel computing to the Society of Mind model of thinking (Minsky, 1987), which he developed with Seymour Papert.

As we saw in an earlier section, Tymoczko claimed that the partnership of computers in proofs has altered our notion of proof and introduced into it empirical considerations. The view of the brain as a computer suggests that Tymoczko’s argument applies to all proofs. We can no better survey the workings of our brain in the act of

¹¹The distinction between relation and entity is often thought of as a rigid barrier, but we forget that familiar “entities” such as the number “2” started out as complex relations between such disparate collections as 2 apples, 2 forks, 2 children. We will explore this more in chapter IV.

proof than we can the program of a computer. Hence, mathematical proofs were always empirical -- relying as they do on the reliability of the “hardware” of the human brain.

New Paradigm, New Pedagogy

In this chapter, we have seen how the rationalist paradigm in mathematics has come under attack and is yielding to an interpretive view - a view of mathematics as a meaning-making activity of a community of mathematical practitioners. In this new view, mathematical objects are not given by the world nor are they just conventions, but instead are constructed and legitimated by the meaning negotiation of the community. We have also seen how the notion of proof has evolved from a linear form of argument whose function is to secure a new truth, to a web of connected ideas whose function is to situate the new knowledge in a society of knowledge. And, lastly, we have seen how the introduction of the computer into the mathematical community weakened the hold of the rationalist paradigm of mathematics and by so doing has accelerated the emergence of the new paradigm.

The old paradigm has inundated the school mathematics class and in turn is supported by the practices of teachers, textbooks and educators immersed in it. For those of us who have embraced the new paradigm, the challenge is to create a new pedagogy, one that is molded from the new characterization of the mathematical enterprise and which in turn will foster its acceptance.

Chapter III - Related Work

This research is situated within a constructionist paradigm for learning and in response to the methods of traditional mathematics education.

School Math

In his book *Mathematical Problem Solving*, Schoenfeld (1985) shows that most high school math students solve algebra word problems by the "keyword" method. If, for example, they see the problem:

John has 5 apples and gives 3 to Mary. How many apples does John have left?

Students scan the text from the right, see the keyword "left" and interpret the context to be subtraction. They then search out the two quantities to subtract, find 5 and 3, perform the subtraction, and obtain "2" as a result.

But if we try to fool them and give:

John has 5 apples. Mary has 3 apples. Mary left the room. How many apples does John have?

they use the same keyword method to produce the incorrect answer, 2. Surely, one would think that in most mathematics problems, this strategy would not be too effective.

Surprisingly, in a major textbook series analyzed by Schoenfeld, 97%¹² of all word problems yielded the right answer when worked by the keyword method. 97%! From a purely instrumental point of view, it is rational to use the keyword method under such circumstances. It is a faster method for getting correct answers than understanding the problem. Note here that 97% means that if you are a student working accurately by the keyword method, you're not only passing your math course, you're acing it!

¹² In a more recent article, Schoenfeld (1991), estimates that 95% of average textbook exercises can be solved by the keyword method.

It is much easier now to see how students successful in school mathematics nevertheless have trouble understanding the mathematics they can “do”. An amusing parody of students’ responses to the routine of mathematical exercises is given by Ihor Charischak (1993) in his “A Student’s Guide to Problem Solving”:¹³

Rule 1: If at all possible, avoid reading the problem. Reading the problem only consumes time and causes confusion.

Rule 2: Extract the numbers from the problem in the order in which they appear. Be on the watch for numbers written in words.

Rule 3: If rule 2 yields three or more numbers, the best bet for getting the answer is adding them together.

Rule 4: If there are only two numbers which are approximately the same size, then subtraction should give the best results.

Rule 5: If there are only two numbers in the problem and one is much smaller than the other, then divide the smaller into the larger if it goes in evenly. Otherwise multiply.

Rule 6: If the problem seems like it calls for a formula, pick a formula that has enough letters to use all the numbers given in the problem.

¹³ What makes the parody especially effective is that the rules the students give are of exactly the same form as the rules they are given in school mathematics. They are just applied in a “meta-context” — that of the school mathematics culture.

Rule 7: If rules 1 - 6 don't seem to work, make one last desperate attempt. Take the set of numbers found by rule 2 and perform about two pages of random operations using these numbers. You should circle about five or six answers on each page just in case one of them happens to be the answer. You might get some partial credit for trying hard.

Rule 8: Never, never spend too much time solving problems. Remember, *the best student in mathematics is the one who gets to the bottom of the page first!* This set of rules will get you through even the longest assignments in no more than ten minutes with very little thinking.

Some math educators have attempted to get around this problem by giving more sophisticated rules for when to apply each operation. But, as long as these rules are purely syntactic, they are bound to fail in most real-life contexts. They fail as the parody above suggest for affective reasons - no relevance to the student's life or connection to her personal "community of ideas". But even on their own terms, rule-based approaches must fail in real-world contexts. Consider the following two examples (Hersh and Davis, 1981)

*One can of tuna fish costs \$1.05. How much do two cans cost?
A billion barrels of oil cost x dollars. How much does a trillion barrels of oil cost?*

Syntactically, these look like simple addition or multiplication contexts. But, if these are to be interpreted as problems in the real world, than no such simple operations will do. In the first case, supermarkets often discount 2 cans of the same product. In the second, the law of supply and demand dictates that a trillion barrels of oil (being a fair fraction of the world supply) will cost much more than 1000 times the cost of a billion. At first glance these may seem like trivial objections, but if you stop to consider that real situations in the world are mostly of this kind, then you start to see the problem in training students to

apply syntactic rules to problems that cannot be solved by these rules in virtually any context outside the textbook.

The above leads us to see that:

“When students are trained to solve problems without understanding, they will come to accept the fact that they don't understand; they will implement procedures without trying to understand.” (Schoenfeld, 1985)

Constructionism

Based on the constructivist psychology of Piaget (e.g., 1952) and the progressive educational tradition exemplified in Dewey (1902), Constructionism (Papert, 1987) is the term coined by Papert to emphasize that the best way for the learner to construct knowledge in the head is to build something tangible - a meaningful product. But Constructionism as a philosophy in practice has gone beyond its original definition. Constructivist and constructionist teachers and researchers have fleshed out this definition into a body of beliefs and practices, many of which are inspired by Papert's seminal work, *Mindstorms* (Papert, 1980).

Among its central tenets are¹⁴:

- People learn by *doing* and then reflecting on what they did¹⁵.

In particular this means that any transmission theory of education is hopelessly wrongheaded. Metaphors that describe knowledge as being poured into an empty vessel head or being deposited and withdrawn from a mental bank misrepresent the nature of learning. Verbal explanations prior to relevant experience don't work very well. The

¹⁴ Inspiration for the following came from Uri Leron and his "Reflections on the Constructivist Approach".

¹⁵ I do not mean by this formulation to imply a discrete seriality. Reflecting while doing and the circle of doing/reflecting/doing are the prototypical modes of learning.

useful function of verbal explanations is to summarize knowledge gleaned from activities and to name the knowledge for future reference.

- Learners build up mental structures ("concepts").

In its simplest form this tenet is a rejection of a behaviorist psychology and an acceptance of a cognitivist model of the mental life. But constructionism takes this much farther than a simple-minded cognitivism. From Piaget we have learned that people assimilate new information to their current mental structures, and in turn their past structures must accommodate themselves in order to fit the new information (Piaget, 1970). So new ideas are getting linked in to a vast mental network that will allow some ideas to readily assimilate while yet others will be transformed radically by the assimilating structure. If you try to teach a concept that is well assimilated to your own internal structure but the structures of the learners are sufficiently different from yours, than what you teach will be radically transformed.

Furthermore, there is no way through the network filter. Once discovering a conceptual "mistake" a student has made, correcting it will do no good -- the correcting information will pass through the same assimilating structures and again be suitably transformed (see e.g. Carey, 1985; diSessa, 1983; McCloskey, 1988).

- Good activities are open-ended, and allow for many different approaches and solutions.

Open-ended activities encourage learners to pose their own problems and make conjectures as to what mathematical relationships obtain. These skills in themselves may be at the core of what mathematical practice is about (Brown, 1983; Lakatos, 1976; Polya, 1954). Obviously, if learners have gotten engaged enough in doing mathematics to have invented their own problems and made their own conjectures, their sense of ownership of the mathematical activity is much greater than when problems are handed

them by an external authority. If you've invented your own problem, you know why it is interesting to you, what motivated the interest and what kind of thing could make sense as a "solution". If you've made your own conjecture, you've put yourself at risk for being wrong -- you have a stake in defending your conjecture.

At first, not having a clear-cut externally validated goal may be anxiety provoking for the "school-habituated" learner. However, once she is involved and has set her own goals, she is encouraged in her creativity, individuality, and self-expression. Many different goals can be set from any one activity, many different conjectures advanced. Some of these will prove more fruitful than others. This gives learners the opportunity to develop the highly complex judgments by which mathematicians evaluate the quality of as yet unanswered questions and as yet unproved conjectures. (see Davis and Hersh, 1981)

- Group activities stimulate reflection.

Working in a small group affords you the opportunity to let your own ideas be shared with others in a not-too-risky atmosphere where others may even appreciate them. Making contributions to a mathematical community initiates a positive feedback loop of contributions in which others react to, elaborate on, or try to invalidate the contribution. This feedback loop results in the "bootstrapping" of a mathematical culture (Papert, 1980). If the activity is open-ended, then almost everyone in a group will have something worthwhile to contribute. Learning to see the many different features that are salient for others in the group and the myriad different styles of thinking that are possible expands the space of possible inquiry. When one style is not privileged by an authority, the group can make its own meaningful evaluation of which approaches are fruitful under which circumstances.

In theory, large class discussions can also have the same effect. In practice however, it is a rare classroom which has been made safe enough for learners to try out uncertain ideas without personal risk.

Constructivist Alternatives to School Math

Some educators such as Magdalene Lampert (1990) and Deborah Ball (1991b), while at the same time trying to increase classroom safety, have exploited the riskiness of open conjecture in motivating students to defend their ideas in classroom arguments.

Students who have a stake in maintaining their reputation as mathematical conjecturers are creative in finding clever arguments, examples, and counterexamples to support their own ideas and refute others'. In this way, the notion of mathematical proof is constructed as a limiting form of everyday classroom argument. A proof is a very convincing argument, and what counts as convincing depends very much on the sophistication of the community one is attempting to convince.

Another illuminating example of large group mathematical activity is to be found in the doctoral work of Idit Harel (1988). In Harel's project, a fourth grade class embarked on a project to author software for instructing third graders about fractions. Students were free to work alone or in groups, but in practice new ideas of how to present material circulated in the classroom in a "collaboration in the air" (Harel & Kafai, 1991). Even children who were most comfortable working alone ended up participating in the creation of a "fractions culture" and children who had been outside the social circle of the classroom enjoyed increased self-esteem as a result of the recognition of their contributions to that culture.

In the past twenty years there has been some (though still not much) movement in the elementary and even high schools toward teaching children to be mathematicians instead of teaching them mathematics¹⁶ (see Papert, 1971). A considerable literature has grown up describing experiments in constructivist teaching in the schools (e.g. Hoyles & Noss, 1992; Steffe & Wood, 1990; Streefland, 1991; Von Glasserfeld, 1991). A plethora of creative approaches have begun to proliferate:

Hoyles and Noss (1992) integrate the computer and the language LOGO into the mathematics classroom, providing new means of expressing mathematics, and encouraging a diversity of styles and approaches to learning/doing mathematics.

Increasing availability of computers in classrooms has facilitated the development of computer microworlds for mathematical explorations (see e.g. Edwards, 1992; Schwartz & Yerushalmy, 1987; Leron & Zazkis, 1992) and emphasized dynamic mathematical processes as opposed to static structures. (see e.g. Vitale, 1992).

Strohecker (1991) explores ideas of mathematical topology through classroom discussions of and experiments with tying knots.

Cuoco & Goldenberg (1992) focus on geometry as providing a “natural” bridge between the visual intuitions of the high schooler and more symbolic areas of mathematics and, by so doing, “expands students’ conception of mathematics itself.”

Confrey (1991) points out how easily we fall into the misconceptions trap and fail to listen to the sense, the novel meaning making of the mathematical learner.

Schoenfeld (1991) moves from a college problem solving class using pre-decided problems to a focus on learners’ creation of related problems and, ultimately, to the development of classrooms that are “microcosms of mathematical sense making”.

¹⁶ Although some of this movement may be due to a mistaken lesson learned from Piaget -- that children, because they think differently from adults, need a "simpler more concrete" approach until they are "mature" enough to be lectured at.

Konold (1991) works with beginning college students on probability and focuses on how students interpret probability questions. He concludes that effective probability instruction needs to encourage learners to pay attention to the fit between their intuitive beliefs and 1) the beliefs of others 2) their other related beliefs and 3) their empirical observations.

Rubin & Goodman (1991) have developed video as a tool for analyzing statistical data and supporting statistical investigations. Nemirovsky (1993) & Kaput (in press) have developed curricular materials for exploring a qualitative calculus. They view themselves as part of a calculus reform movement that seeks to understand calculus as the mathematics of change and thus make calculus into a “fundamental longitudinal strand” for the learning of mathematics.

Most of the above alternatives took place in the elementary school setting, and even those that took place in high school or college tended to focus on relatively elementary mathematics and relatively novice learners. As of yet, there has been very little written about the methods used at higher educational levels (late college through graduate school through life beyond school) which if anything are more rigidly instructionist (see Papert, 1990), provide less opportunity for self-directed exploration and conceptual understanding than do classes in the lower grades.

Chapter IV - Concrete Learning

“No ideas but in things” - William Carlos Williams

I will now advance a theoretical position which is the underlying force behind this study. This will involve a critical investigation of the concept of “concrete” and what it means to make mathematics concrete.

Concepts that are said to be concrete are also said to be intuitive - we can reason about them informally. It follows that if we hold the standard view that concreteness is a property of objects which some objects have and some do not, we are likely to be led into the view that our intuitions are relatively fixed and unchangeable. This last conclusion, that of the relative unchangeability of intuitions, is the force behind the prescriptions to avoid intuitive probability judgments. When we replace the standard view of concrete with a new relational view, we also see the error in the static view of intuitions. In understanding the process of developing concrete relationships we learn how to support the development of mathematical intuitions.

In the present chapter, I will expose the weakness of the standard view of concrete, and probe the very notion of “concreteness”. I propose a new definition of “concrete”. In the new model, concreteness is a property not an object itself but of our relationship to that object. A concept is “concrete” when

- it is well-connected to many other concepts
- we have multiple representations of it
- we know how to interact with it in many modalities.

As such, the focus is placed on the process of developing a concrete relationship with new objects, a process I call “concretion”.

Implications for pedagogy of this new view of concrete will be explored, and a multi-faceted analogy will be constructed with conservation experiments of Piaget.

Many of these ideas are expounded in greater detail in (Wilensky, 1991a).

Concretion

Seymour Papert has recently called for a “revaluation of the concrete”: a revolution in education and cognitive science that will overthrow logic from “on top and put it on tap.” Turkle and Papert (1991) situate the concrete thinking paradigm in a new “epistemological pluralism”-- an acceptance and valuation of multiple thinking styles, as opposed to their stratification into hierarchically valued stages. As evidence of an emerging trend towards the concrete, they cite feminist critics such as Gilligan's (1982) work on the contextual or relational mode of moral reasoning favored by most women, and Fox Keller's (1983) analysis of the Nobel-prize-winning biologist Barbara McClintock's proximal relationship with her maize plants, her “feeling for the organism.” They cite hermeneutic critics such as Lave (1988), whose studies of situated cognition suggest that all learning is highly specific and should be studied in real world contexts.

For generations now we have viewed children's intellectual growth as proceeding from the concrete to the abstract, from Piaget's concrete operations stage to the more advanced stage of formal operations (e.g., Piaget, 1952). What is meant then by this call for revaluation of the concrete?

And what are the implications of this revaluation for education? Are we being asked to restrict children's intellectual horizons, to limit the domain of inquiry in which we encourage the child to engage? Are we to give up on teaching general strategies and limit ourselves to very context specific practices? And what about mathematics education? Even if we were prepared to answer in the affirmative to all the above questions for education in general, surely in mathematics education we would want to make an exception? If there is any area of human knowledge that is abstract and formal, surely mathematics is. Are we to banish objects in the head from the study of mathematics? Should we confine ourselves to manipulatives such as Lego blocks and

Cuisinaire Rods? Still more provocatively, shall we all go back to counting on our fingers?

The phrases "concrete thinking", "concrete-example", "make it concrete" are often used when thinking about our own thinking as well as in our educational practice.

To begin our investigation we will need to take a philosophical detour and examine the meaning of the word *concrete*. What do we mean when we say that something--a concept, idea, piece of knowledge (henceforward an *object*)--is concrete?

Standard Definitions of Concrete

Our first associations with the word concrete often suggest something tangible, solid; you can touch it, smell it, kick it¹⁷; it is real. A closer look reveals some confusion in this intuitive notion. Among those objects we refer to as concrete there are words, ideas, feelings, stories, descriptions. None of those can actually be "kicked." So what are these putative tangible objects we are referring to?

One reply to the above objection is to say: No no, you misunderstand us, what we mean is that the object referred to by a concrete description has these tangible properties, not the description itself. The more the description allows us to visualize (or, if you will, sensorize) an object, to pick out, say, a particular scene or situation, the more concrete it is. The more specific the more concrete, the more general the less concrete. In line with this, Random House says concrete is "particular, relating to an instance of an object" not its class.

Let us call the notion of concrete specified by the above the standard view. According to the standard view, concrete objects are specific instances of abstractions, which are generalities. Concrete objects are rich in detail, they are instances of larger classes; abstract objects are scant in detail, but very general.

¹⁷ As in Samuel Johnson's famous refutation of idealism by kicking a stone

Given this view, it is natural for us to want our children to move away from the confining world of the concrete, where they can only learn things about relatively few objects, to the more expansive world of the abstract, where what they learn will apply widely and generally. Yet somehow our attempts at teaching abstractly leave our expectations unfulfilled. The more abstract our teaching in the school, the more alienated and bored are our students, and far from being able to apply their knowledge generally across domains, their knowledge displays a “brittle” character, usable only in the exact contexts in which it was learned.¹⁸ Numerous studies have shown that students are unable to solve standard math and physics problems when these problems are given without the textbook chapter context. Yet they are easily able to solve them when they are assigned as homework for a particular chapter of the textbook (e.g., DiSessa, 1983; Schoenfeld, 1985).

Critiques of the Standard View

Upon closer examination there are serious problems with the standard view. For one thing, the concept of concrete is culturally relative. For us, the word “snow” connotes a specific, concrete idea. Yet for an Eskimo, it is a vast generalization, combining together twenty-two different substances. As was noted by Quine (1960), there are a multitude of ways to slice up our world, depending on what kind and how many distinctions you make. There is no one universal ontology in an objective world.¹⁹

An even more radical critique of the notion of a specific object or individual entity comes out of recent research in artificial intelligence (AI). In a branch of AI called

¹⁸ A fruitful analogy can be made here between this kind of brittleness and the brittle representation of knowledge that so called expert systems in AI exhibit. In effect this kind of abstract teaching is akin to programming our children to be rule-driven computer programs. A similar kind of brittleness can be found in simple animals such as the sphex wasp. For a discussion of "sphexish" behavior and how it differs from human behavior, see Hofstadter (1982) Dennett (1984).

¹⁹ A tenet of the empiricist world view. For an illuminating comparison of the empiricist, rationalist, and hermeneutic stances, see Packer and Addison (1989).

emergent AI, objects that are typically perceived as wholes are explained as emergent effects of large numbers of interacting smaller elements. Research in brain physiology as well as in machine vision indicate that the translation of light patterns on the retina into a “parsing” of the world into objects in a scene is an extremely complex task. It is also underdetermined; by no means is there just one unique parsing of the inputs. Objects that seem like single entities to us could just as easily be multiple and perceived as complex scenes, while apparently complex scenes could be grouped into single entities. In effect the brain constructs a theory of the world from the hints it receives from the information in the retina. It is because children share a common set of sensing apparatus (or a common way of obtaining feedback from the world, see Brandes & Wilensky, 1991) and a common set of experiences such as touching, grasping, banging, ingesting²⁰, that children come as close as they do to “concretizing” the same objects in the world.²¹ The critique of the notion of object connects also to the work of Piaget. In Piaget's view of development, the child actively constructs his/her world. Each object constructed is added to the personal ontology of the child. Thus we can no longer maintain a simple sensory criterion for concreteness, since virtually all objects, all concepts which we understand, are constructed, by an individual, assembled in that particular individual's way, from more primitive elements²². Objects are not simply given to the senses; they are actively constructed.²³

²⁰ I ignore here the social experiences, which play a large role in determining which objects are useful to construct.

²¹ Alternatively, we can say that children construct a *model* of the world through feedback they receive from their active engagement with the world (see Wilensky, 1991a).

²² In other words, whether something is an object or not is not an observer-independent fact; there is no universal objective (object-ive) way to define a given composition as an object. It thus follows that when we call an object concrete, we are not referring to an object "out there" but rather to an object "in here", to our personal constructions of the object.

²³ Even the recognition of the most "concrete" particular object as an object requires the construction of the notion of object permanence

We have seen that when we talk about objects we can't leave out the person who constructs the object.²⁴ It thus follows that it is futile to search for concreteness in the object -- we must look at a person's construction of the object, at the relationship between the person and the object.

Towards a New Definition of Concrete

I now offer a new perspective from which to expand our understanding of the concrete. The more connections we make between an object and other objects, the more concrete it becomes for us. The richer the set of representations of the object, the more ways we have of interacting with it, the more concrete it is for us. Concreteness, then, is that property which measures the degree of our relatedness to the object, (the richness of our representations, interactions, connections with the object), how close we are to it, or, if you will, the *quality of our relationship* with the object.

Once we see this, it is not difficult to go further and see that any object/concept can become concrete for someone. The pivotal point on which the determination of concreteness turns is not some intensive examination of the object, but rather an examination of the modes of interaction and the models which the person uses to understand the object. This view will lead us to allow objects not mediated by the senses, objects which are usually considered abstract—such as mathematical objects—to be concrete; provided that we have multiple modes of engagement with them and a sufficiently rich collection of models to represent them.

When our relationship with an object is poor, our representations of it limited in number, and our modes of interacting with it few, the object becomes inaccessible to us. So, metaphorically, the abstract object is high above, as opposed to the concrete objects,

²⁴ Seymour Papert has said: “You can’t think about thinking without thinking about thinking about something”. We might paraphrase him here and say: You can't think about something without thinking about *someone* thinking about something.

which are down and hence reachable, “graspable.” We can dimly see it, touch it only with removed instruments, we have remote access, as opposed to the object in our hands that we can operate on in so many different modalities. Objects of thought which are given solely by definition, and operations given only by simple rules, are abstract in this sense. Like the word learned only by dictionary definition, it is accessible through the narrowest of channels and tenuously apprehended. It is only through use and acquaintance in multiple contexts, through coming into relationship with other words/concepts/experiences, that the word has meaning for the learner and in our sense becomes concrete for him or her. As Minsky says in his *Society of Mind*:

The secret of what anything means to us depends on how we've connected it to all the other things we know. That's why it's almost always wrong to seek the “real meaning” of anything. A thing with just one meaning has scarcely any meaning at all (Minsky, 1987 p. 64).

This new definition of concrete as a relational property turns the old definition on its head. Now, thinking concretely is seen not to be a narrowing of the domain of intellectual discourse, but rather as opening it up to the whole world of relationship. What we strive for is a new kind of knowledge, not brittle and susceptible to breakage like the old, but in the words of Mary Belenky, “connected knowing“ (Belenky, Clinchy, Goldberger, & Tarule, 1986).

Consequences of the New View

Below is a pithy summary of what I see as the consequences, both theoretical and practical, of the above reformulation of the concrete/abstract dichotomy:

- It's a mistake to ask what is the "real meaning" of a concept.

This "thing with only one meaning" is a characterization of a formal mathematical definition. An object of thought which is given solely by definition, and an operation given only by simple rules, is abstract in this sense. Like the word learned only by

dictionary definition, it is accessible through the narrowest of channels and tenuously apprehended. It is only through use and acquaintance in multiple contexts, through coming into relationship with other words/concepts/experiences, that the word has meaning for the learner and in our sense becomes concrete for him or her.

- When concepts are very robust they become concrete.

From Piaget's conservation experiments, we see that what appears to us to be a very concrete property of an object, like its very identity, is constructed by the child over a long developmental period. After a brief period of transition, when the child acquires the conserved quantity, she sees it as a concrete property of the object.

Similarly, a machine vision program must build up notions of objects from more primitive notions in its vision ontology. Which things it adds to its ontology, making them concrete "wholes", depends on what primitive elements it started with and what objects it has already built up. As Quine has shown in his famous "rabbits vs. undetached rabbit-parts" example (Quine, 1960), there is no way unique way to slice up the world into objects.

- There's no such thing as abstract concepts - there are only concepts that haven't yet been concretized by the learner.

We mistakenly think that we should strive for abstract knowledge since this seems to be the most advanced kind of knowledge that we have. The fallacy here is that those objects which we have abstract knowledge of are those objects that we haven't yet concretized due to their newness or difficulty to apprehend. When we strive to gain knowledge of these, it seems like we strive for the abstract. But in fact, when we have

finally come to know and understand these objects, they will become concrete for us.
Thus Piaget is stood on his head: development proceeds from the abstract to the concrete.

- Concepts Become Concrete

- Through connection to other concepts
- When they are multiply represented²⁵
- Through many modes of interaction
- By engaging in activities followed by reflection

Mathematical Concepts are Societies of Agents

In his book “The Society of Mind”, Marvin Minsky (1989) describes the workings of the human mind in a radically decentered way. We are, says Minsky, a collection of thousands of distributed “agents”, each by itself stupid, specialized, able to communicate only a very small part of what it knows to any other agent.

One example which Minsky and Papert analyze through the society of mind theory is the water volume experiment of Piaget (1954). In this classic experiment, a child is shown two full water glasses, one taller and one wider glass. He/she is then asked which glass has more water. Children before a certain stage (usually around age 7) say that the tall glass has more. If the tall glass is then poured into another wider glass and fills it, then if the children are asked which of the two full glasses has more, they say the

²⁵ The issue of representation is controversial when applied to distributed concepts. A deep treatment of this issue is beyond the scope of this thesis. For our purposes, a concept is a cluster of connections and a representation is one of these connections which we take to refer to the cluster.

two glasses have the same amount of water. But then if the water is poured back into the tall glass, they again assert that the tall glass contains more water.

This phenomenon is robust and has resisted all attempts to explain it away as merely a semantic artifact. Piaget explains the children's behavior as a construction of the concept of liquid volume. Before roughly age 7, children do not conserve volume, liquids can become more or less if their appearance is altered.

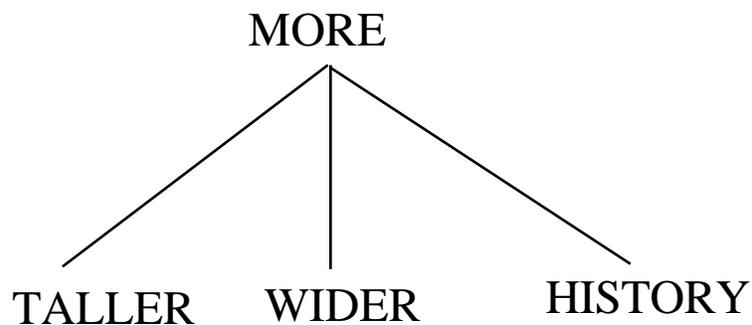
Minsky and Papert's analysis of this situation posits the existence of an *MORE agency* (a system or society of agents) and three sub-agents:

TALLER: This agent asserts which of two glasses is taller.

WIDER: This agent asserts which of two glasses is wider.

HISTORY: This agent asserts that two glasses have the same amount of liquid if nothing has been added or removed.

In the pre-conservational state, a child's *MORE agency* can be thought of as connected as follows:

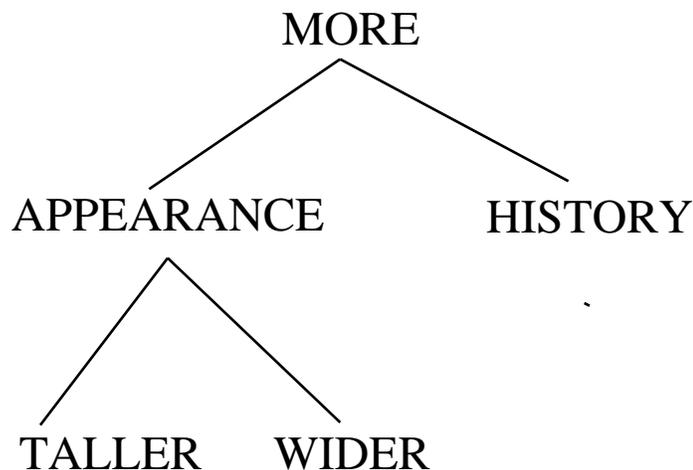


If **TALLER** asserts itself, then the *MORE agency* listens to it. If **TALLER** is quiet, then *MORE* listens to **WIDER**. If **WIDER** is also silent, then *MORE* listens to

HISTORY. Thus, the child's agents are arranged in a linear order. If one glass is taller than the other, MORE will assert that that glass has more liquid.

But post-conservation, say Minsky and Papert, the child's agents are connected differently. An APPEARANCE agent is created that subsumes TALLER and WIDER. The APPEARANCE agent functions as a middle manager between MORE and TALLER/WIDER.

In this new organization, MORE listens to APPEARANCE first. But APPEARANCE listens to both TALLER and WIDER. If only one of these asserts itself or if they agree, then APPEARANCE speaks up about which glass has more. However, if the two agents TALLER and WIDER disagree, then the "principle of non-compromise" is invoked and APPEARANCE is silent. MORE, then, listens to HISTORY in making its decision.



Minsky summarizes the lessons of this example as Papert's principle:

“Some of the most crucial steps in mental growth are based not simply on acquiring new skills, but on acquiring new administrative ways to use what one already knows.”

An important corollary of Papert’s principle is that, when new skills are developed, it doesn’t follow that old habits, intuitions, agents are dispelled. They may still resurface in other contexts. But their connections in a particular agency will have been greatly altered.²⁶

While it happened here that the few agents in this example are organized hierarchically, this can be expected to be true in only the simplest cases. Nor is it generally true that all agencies have a clear “execution model”. In the general case, agencies interpenetrate, call each others' sub-agents and super-agents and form a tangle of connections. As Minsky says:

“In its evolutionary course of making available so many potential connections, the human brain has actually gone so far that the major portion of its substance is no longer in its agencies but constitutes the enormous bundle of nerve fibers that potentially connect those agencies. The brain of Homo Sapiens is mainly composed of cabling.”

One way then to measure the strength of an agent or agency is to see how well connected it is. We can then recharacterize our definition of concrete in agent language: *an agent is more or less concrete depending on the number of agents to which it connects.*

The water volume experiment also illustrates the value of agent conflict as an impetus to learning. It was the conflict between the TALLER and the WIDER agents that led to the silence of the APPEARANCE agent²⁷ and the subsequent attention to

²⁶ Old agents never die, they just get reconnected?

²⁷ How the two agents TALLER and WIDER get grouped, or what causes the APPEARANCE agent to arise is another interesting question but is beyond the scope of this thesis.

HISTORY. In order for this conflict to be detected, TALLER and WIDER had to both be relatively strong. If TALLER continued to dominate WIDER as it did in the first figure then the appearance of APPEARANCE would do no good. Thus, it was the conflict between two strong trusted agents that motivated the reordering and enrichment of the child's MORE agency.

In the course of this thesis, we shall draw a strong analogy between intellectual development in adults and conservation experiments of Piaget. The child moves from agent dominance to agent conflict to a new agent network. This process takes a long time, but results in a new set of intuitions about volume. We take this as a model for adult development as well.

The usefulness of agent conflict in this example suggests that paradox is generally a powerful motivator toward new understanding. In a paradox, two trusted and reliable arguments (or agents) that have done their jobs faithfully before with good results are seen to be in conflict.

Paradox

Throughout history, it has often been the case that paradoxes have led to major reconstructions of mathematical and scientific systems. This should no longer surprise us since much of what is paradoxical about paradoxes is their counterintuitiveness. But it is precisely through encountering paradox that intuitions are built. When two trusted agents conflict, we are ripe for new agent links and middle managers which embody new intuitions. Paradox has played this role in the history of science, as Anatol Rapoport (1967) notes:

Paradoxes have played a dramatic role in intellectual history, often foreshadowing revolutionary developments in science, mathematics and logic. Whenever, in any discipline, we discover a problem that cannot be solved within the conceptual framework that supposedly should apply, we experience shock. The shock may compel us to discard the old framework and adopt a new one. It is

to this process of intellectual molting that we owe the birth of many of the ideas in mathematics and science....Zeno's paradox of Achilles and the tortoise gave birth to the idea of convergent infinite series. Antinomies (internal contradictions in mathematical logic) eventually blossomed into Gödel's theorem. The paradoxical result of the Michelson-Morley experiment on the speed of light set the stage for the theory of relativity. The discovery of the wave-particle duality of light forced a reexamination of deterministic causality, the very foundation of scientific philosophy, and led to quantum mechanics. The paradox of Maxwell's demon which Leo Szilard first found a way to resolve in 1929, gave impetus more recently to the profound insight that the seemingly disparate concepts of information and entropy are intimately linked to each other.

Rapoport's examples are among the most salient, but many more mundane examples are to be found in the history of mathematics and science. Questions such as the ones below were hotly debated paradoxes and perplexed mathematicians in their times.

- What is the result of subtracting a larger number from a smaller?
- What is the result of dividing a positive number by a negative number?²⁸
- What is the square root of a negative number?
- Which are there more of: integers or even integers?
- How can a finite area have infinite perimeter?
- How can an quantity be infinitely small, yet larger than zero?

But parallel to their phylogenetic importance, the analogy of paradox to agent conflict lets us see their important role in ontogenetic development. It is through the resolution of these conflicts that new meanings and epistemologies are negotiated. But to be of value in development, both agents in a paradox must be strong and trustable. If one

²⁸ If this doesn't seem paradoxical, consider the following argument: negative numbers are modeled as debt. Now we all know that if we have \$12 and our friend has \$6, then we have twice as much as she does. If she has only \$4, then we have 3 times as much. As she spends more of her money, we have a larger and larger relative amount. When she spends all of her money, we have infinitely more than she does. Now, suppose she goes into debt \$10. Don't we now have even infinitely more than she does than we did when she had nothing? It was considerations like this that led the tenth century mathematician Wallis (famous for his infinite product formula for $\pi/4$) to conclude that there were many different kinds of infinities and the one generated by division by zero was smaller than those generated by dividing by a negative number. This difficulty was resolved when a directional model of negative numbers was proposed. I am indebted to David Dennis (1993) for this piece of history.

is much more trustable than the other then, like TALLER, it will dominate and no paradox will even be perceived²⁹. If both agents are weak, then the paradox is not compelling and only serves to undermine further the credibility of each agent.³⁰ In order to foster development then, we need to strengthen each of the conflicting agents. If, as is so often done in mathematics classes, we resolve the conflict prematurely by declaring one agent a victor, as being the “right answer”, the “right way to do it”, or the “right definition”, then we undermine the processes that work to concretize these agents, linking them in to the network of agents. Unless we experience the conflict, see which way we might have gone if we hadn’t gone the “right” way, then we will get lost on the next trip. The result of only getting the “right answers” is brittle formal understanding. We shall see in the probability interviews in Chapter VIII examples of all of the above responses to paradox.

Concepts are Messy³¹

Mathematicians have argued that proofs are the essence of their enterprise. What distinguishes Mathematics from other disciplines is the certainty that is obtained through the rigor of proofs³². But in fact proofs are not the source of mathematical certainty. They are a technique used by mathematicians to create a uniform procedure for verification of mathematical knowledge. The technique consists of “linearizing” the complex structure that constitutes the ideas in a mathematical argument. By means of this linearization,

²⁹ We will see, in Chapter VIII, a strong analogy between this situation and the situation of most learners encountering the Monty Hall problem.

³⁰ Magicians often remark that many of their tricks do not impress children. This is not surprising when we consider that a child without a strong object-permanence agency will not be terribly impressed by a rabbit disappearing in a hat.

³¹ To better understand the structure of this messiness, it is useful to attempt a model of particular concepts in terms of a society of knowledge agents (Minsky, 1987) or knowledge pieces (diSessa, 1985).

³² The history of mathematics however has brought us many examples of "mistaken proofs" (e.g. Kummer's famous proof of Fermat's last theorem). As we have seen, recently mathematicians such as Rabin (1976) have tried to salvage the notion of proof through probabilistic methods.

mathematical proofs can be checked line by line, each line either an axiom or derived from previous lines by accepted rules of inference.

But the hegemony of the standard style of communicating mathematics (definition/theorem/proof) constitutes a failure to come to terms with the mind of the learner. We attempt to teach formal logical structures in isolation from the experiences that can connect those structures to familiar ideas. The result is that the idea too often remains “abstract” in the mind of the student, disconnected, alien, and separate, a pariah in the society of agents.

Even if the motivation is there to communicate our mathematical ideas to those who don't already understand them, we may have lost the organization of mental agents that existed before the idea was acquired, and forgotten the mechanisms by which we ourselves acquired the ideas. This is clearly illustrated by the conservation experiments of Piaget (1952). In a typical such experiment, a child is shown a tall thin glass filled with water which is then poured into a shorter wider glass. When asked which glass contains more water, so-called “pre-conservation” children say that the tall glass has more. A year or so later, these same children now “conservational” are shown video tapes of their earlier interviews on the subject. These children do not believe that they could ever have made such a “ridiculous” claim. “Of course the glasses contain the same amount of water - the videotapes must have been faked.” (Papert, 1990). Piaget's experiments are dramatic examples of what is a quite common phenomenon -- people have a hard time remembering what it was like to not understand a concept which they now understand. What is it like to not accept the identity of objects through time? What is it like to not understand what addition is? Because of this phenomenon, a well-meaning mathematical educator often cannot reconstruct the process by which he/she was convinced of the efficacy of a definition or the validity of a proof and mistakenly believes that the linear proof contains all the information and structure necessary for the conceptual understanding.

If asked to justify the formal style of their expostulation, an author or teacher of mathematics may respond that the linear structures (definition/theorem/proof) capture most economically the essence of the material. But it may be that this reason for hiding the messy structure of mathematical ideas is not the whole story. Revealing that the structure of mathematics in your head does not mirror the clean elegant lines of the mathematical text can be quite embarrassing. Yet to reveal that process may be perceived as an admission of vulnerability and weakness. The theorem and proof is the logical mental construct of which we are proud; the web of connections, intuitions, partial understandings, and mistakes from which that logical construct arose may be a source of shame. The traditional form of mathematical expostulation is a shield that mathematical culture has developed to protect the mathematical teacher or author from embarrassment. It allows the mathematician to present a controlled and logically inexorable understanding without exposing him to the risk of revealing his own messy internal thought processes.

Covering up is Damaging

This covering up of the hidden messy structure of mathematical ideas can be damaging to the mathematical learner. If learners believe that the mathematics as presented is a true picture of the way the mathematics is actually discovered and understood, they can be quite discouraged. Why is their thinking so messy when others' is so clean and elegant? They conclude that clearly mathematics must be made only for those born to it and not available to mere mortals. Mathematical discourse is not a form of persuasion continuous with daily discourse, but is instead in some special province all its own, a purely formal phenomenon. These mathematical learners are deprived of the experience of struggling for a good definition, and the knowledge that mathematical truths are arrived at by a process of successive refinement not in a linear and logically inexorable fashion. (see Lakatos, 1976)

Unfortunately, this kind of mathematical culture is self-reinforcing. Those who survive the stark, formal modes of presentation and manage to concretize mathematical structures sufficiently to pursue a career in mathematics have learned along the way that to reveal their internal thought processes is to violate a powerful social norm. In parallel to the standard mathematical curriculum, the student has learned the following “lessons” of mathematical culture:

- No one is interested in your personal constructions of mathematical ideas.
- There is one canonical way of understanding mathematical objects.
- The correct procedure for understanding mathematical ideas is to replace your former unrigorous intuitions with the received rigorous definitions.
- Mathematics is to be grasped instantaneously - struggle with an idea is a sign of dullness and lack of ability.

The student in a mathematics classroom who braves a revelation of his/her tentative understanding of an idea is too often confronted with a demeaning response, “How can you not see the answer to that? -- it’s trivial.”³³ Not only does this culture impede learning, it is almost certainly an inhibition to new mathematical discovery. In fact, embarrassment at expressing new, half-formulated ideas is a powerful force for conservatism in mathematics³⁴. It is difficult to challenge old ideas, or to formulate new

³³ The word “trivial” has become the greatest fear of many a student of mathematics. Among graduate students in mathematics, it is often joked that “any theorem that is not profound is trivial”. Indeed after a while this becomes a sort of badge of honor and mathematical textbooks are written so as to allow “trivial” exercises to the reader. There are various humorous and probably apocryphal stories which communicate this point. One such story is told of Norbert Weiner. Lecturing one day to a scholarly audience, he asserted that a particular line of his proof was “obvious”. A member of the audience asked Weiner how it followed. Weiner shook his head, paced back and forth, stared at the blackboard, walked out of the room for fifteen minutes, then returned saying: “As I was saying, it is obvious that....”

³⁴ An important contribution would be made by the person who does a psychological study of the use and power of shame in forming and maintaining a mathematical culture.

ones, in the absence of a culture that supports the floundering, messy process of mathematical exploration.

Chapter V - Fractional Division

The study of fractional division will be used to demonstrate that the knowledge of mathematically sophisticated adults in what is thought to be an elementary mathematical subject area is in fact brittle and fragmented. Careful examination shows that confusions about meaning abound, and non-trivial questions about the epistemology of mathematics arise which when explored can enrich the mathematical understanding.

Division of fractions is an area of mathematics that isn't often encountered in everyday life. As such, the adults in this study, though shown to have only a formal and rule-based understanding of fractional division, do seem to get away with it - in the sense that for most situations they encounter in life, their formal understanding suffices. However, the scope of this early research was limited and it remains an open question if, had we probed deeper, we would have discovered that the limited understanding and lack of connectedness of the ideas of fractions and division did limit and constrain these adults in their mathematical growth. What is clear, is that in the process of the interviews, they were surprised to discover this lack of understanding. Furthermore, they experienced a sense of satisfaction when they succeeded in bringing fractions into the fold of their understanding of division. This satisfaction is one payoff secured by a connected mathematics understanding.

Fractions Setting

The fractions research was carried out over three years in a variety of settings. Eleven personal interviews about fractions and division of fractions were conducted. Of the interviewees, three were in fourth grade, one in high school, one undergraduate, three graduate students, and three post-graduate adults. In addition, I observed group discussions about division of fractions in an epistemology and learning research

seminar³⁵. I also posed the question: "What does it mean to divide two fractions?" as a beginning stimulus to discussion in a number of talks that I gave to university audiences.

Discussion

Most of us draw upon a very limited set of representations when asked to describe or show a fraction (see Streefland, 1991). Typically, in the US, a fraction is represented as a shaded-in piece of a pie. Harel (1990) has documented that, when asked what a fraction is, many children draw this prototypical shaded piece of pie. When asked if the unshaded piece is also a fraction, they deny it. Furthermore, when they are prompted to think about time, money, calculators, or clay, as connected to their representations of fractions, they are unable to do it. This is a clear case in which it can be seen that even though the typical representation of fraction (as a piece of pie) seems concrete in the old sense of the term, (it is related to physical objects which children have lots of experience with) the fact that it is the sole representation makes it abstract in the sense developed in Chapter IV.

Research by Ball (1990) and Wilensky (1989) has shown that when faced with the question "What does it mean to divide two fractions?" or "What does $3/4 \div 1/3$ mean?", most adults do not have satisfactory models. They often say that it means "inverting and multiplying," but are uneasy with this explanation. They sense that this purely formal description of what it means to divide two fractions is not adequate but they recognize that they have no other. When pressed to give a meaning to fractional division other than the invert and multiply rule, most interviewees floundered. Often they go on to give flatly incorrect models of the division. The incorrect models do not give the same result as the invert and multiply rule (which they know and apply effortlessly), yet the conflict between the models they present and the results they can calculate goes undetected.

³⁵ A weekly meeting of our epistemology and learning section of the MIT Media Laboratory in which members presented their research and discussed broad educational and learning issues.

Consider this fragment of an interview with Katya, an MIT undergraduate:

U: What does $3/4 \div 1/3$ mean?

K: Mean?

U: Earlier you were talking of fractions in terms of parts of pizzas, what does $3/4 \div 1/3$ mean in terms of pizza?

K: I guess if you have three quarters of a pizza and divide it up among three people, that's $3/4 \div 1/3$.

U: So how much pizza does each person get?

K: Each one gets a quarter of a pizza.

U: But earlier you said that $3/4 \div 1/3$ is equal to $9/4$.

K: Yeah, three quarters divided by a third is three quarters times three is nine fourths.

At this point in the interview, Katya did not seem to notice that she had calculated two different results for $3/4 \div 1/3$. Later on in the interview when Katya detected this conflict, she realized she “was using two different heads” for the two different “problems”. Katya seems to be speaking as if she were accessing two different agents to express her understanding. One agent deals with fractional division problems, an altogether different and unconnected agent deals with the dividing up pizza context.

Other interviewees simply fail to give a model of fractional division or assert flatly that there is no meaning to dividing fractions other than the rule they have been taught.

For example in this interview fragment, David, a graduate student in psychology says:

D: "I never asked what it means. You just take the second fraction, flip it, and then multiply it by the first fraction, so $3/4 \div 1/3$ is $3/4 \times 3/1 = 9/4$."

U: But is that what $3/4 \div 1/3$ means?

D:Well, I guess you could say it means $3/4$ of 3 things ... since that's what the multiplication means.

U: Is there a direct interpretation of the division in the same way that you used the word "of" to interpret the multiplication?

D: I don't know. I guess not. Division of fractions is different from normal division. I'm not sure if it has a meaning.

In this fragment, David clearly does not see fractional division as an extension of integer division (which he gave a perfectly coherent account of using the word "into"). Instead his knowledge of fractional division is split off (or balkanized, see Papert 1980) from his knowledge of division in general and instead is special "abstract" (in the Wilensky (1991) sense) knowledge about fractions.

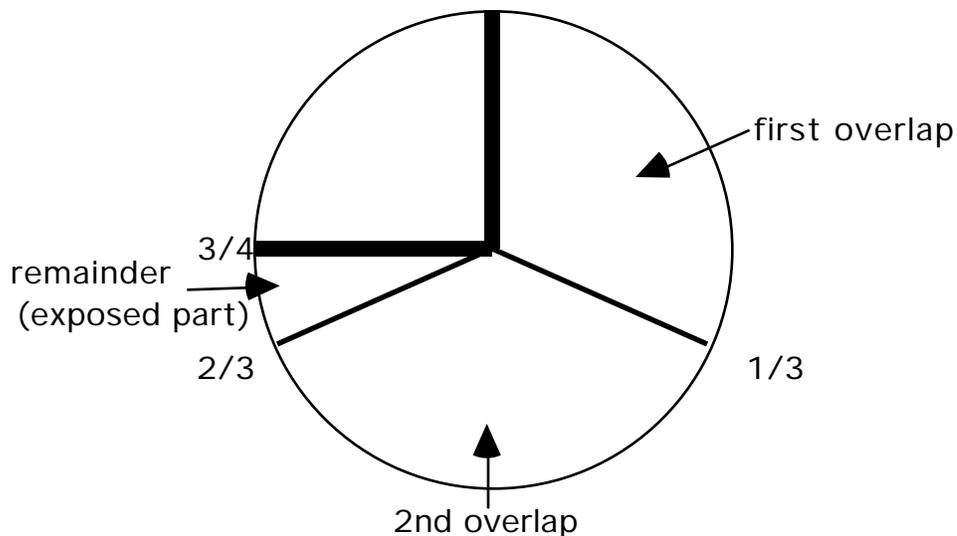
Even in the case where a person gives a correct model of the division problem, they will often choose an unwieldy representation.

Consider José, a graduate student in media arts and sciences:

J: You take $\frac{3}{4}$, [draws the bold line in the figure below] then lay 2 pieces of size " $\frac{1}{3}$ " on it. You see that the two pieces almost cover the $\frac{3}{4}$ size piece, but not quite. So the answer is a little more than 2, in fact we know by calculating that the answer is 2 and $\frac{1}{4}$.

U: But how do you know from the figure that it fits exactly 2 and $\frac{1}{4}$ times?

J: You can't tell exactly from the figure, but it gives you a good estimate, so if your answer is very different from that, you can recalculate and correct your mistake.



José has a more advanced representation of fractional division as a generalization of integer division. But he does not know how to get exact answers using his representation. Also, he continues to use the "pie" representation of fraction that is not that well suited to his diagram. A linear representation of fraction for example would yield a much clearer picture of what's going on. Note also that José does not explicitly map his "story" onto a plausible context or application.

José's explanation is a generalization of what's known as the measurement model of division. In the measurement model of division, to explain the meaning of dividing 6 by 2, you take 6 things, say pizza slices, and assign a measurement unit, say 2 slices = 1 unit (or 1 portion). Then the meaning of $6 \div 2$ is: How many portions can you get from 6 slices?

Measurement Model of Division

$$6 : 2$$



The measurement model of division generalizes pretty well to division of fractions, and José's explanation of that generalization is a reasonable one. However, there is another common model of division -- the partitive model of division. In the partitive model of $6 \div 2$, we take 6 pizza slices and divide them into 2 portions. The meaning of $6 \div 2$ is then: How many pizza slices are there in a portion?

Partitive Model of Division

$$6 : 2$$



These two models of division are quite different and when put in English, it is not at all obvious why the same operation, division, should be used in both cases.

Most English speakers favor the partitive model. This is probably due to the fact that it is more natural in English to think of dividing a pizza for two people to share than it is to think of dividing a pizza into portions of size two slices each. However the partitive model doesn't generalize as easily to fractional division. What does it mean to divide $\frac{3}{4}$ of a pizza among $\frac{1}{3}$ people? We understand dividing $\frac{3}{4}$ of a pizza among integer numbers of people, but $\frac{1}{3}$ of a person is not a very natural notion.

This last observation brings into focus a key point about mathematical representation. Having too limited a collection of representations (or models) of a mathematical concept can result in a trap where none of your representations are sufficient to deal with a new situation. We have seen that both the partitive and the measurement models of division work pretty well when applied to the division of integers. However, if one has only the partitive model, one is much less likely to generalize it to the division of fractions, thus leaving fractional division as a disconnected, "balkanized" piece of knowledge removed from previous concepts of division. Indeed it is quite likely that you may have no model of fractional division and after learning the invert and multiply rule may forget that fractional division is about

division at all. Now if you were "lucky" and happened to hit upon the measurement model of division, then you probably could generalize integer division to fractional division.

Let us be sure to draw the right conclusion from the above paragraph. It would be wrong to conclude that the answer to this problem is simply teach the measurement model of division and exclude the partitive model, since it doesn't generalize to fractional division. Fractional division is only one small example of the different ways in which division can be generalized, linked, and applied. We have seen that in being more closely allied to English usage, the partitive model is more useful for understanding ordinary English conversation. Moreover, understanding that the models are conceptually different yet arithmetically equivalent is itself a deep fact about mathematics. The moral of this story is that having a richer collection of models leads to deeper mathematical understanding. It is a mistake to search for the right model or the best model.

Chapter VI - Recursion³⁶

“To iterate is human , to recurse divine.”

The recursion study will describe the kinds of learning achieved by adults who are put in an environment where qualitative understandings and broad multiple connections are encouraged. This is in contrast to the fractions study where the interviewees exhibited a formal understanding with very few connections. In a typical mathematics or computer science class, recursion is introduced formally, and the immediate educational goal is usually to teach the students how to write mathematical induction proofs or recursive programs. In this class, the immediate goal was to create many and broad connections to the idea of recursion.³⁷

Recursion Setting

The recursion research took place initially in a special interest group (SIG) I conducted together with Kevin McGee on Recursion, Chaos, and Fractals. This SIG was part of a larger workshop called Science and Whole Learning (SWL) that was conducted in the summer of 1989 by the Epistemology and Learning (E&L) group. Approximately seventy teachers, primarily elementary school teachers from the Boston area, participated in the three week long workshop. The workshop was organized around several key activities: a morning lecture, workshops in Logo and Lego/Logo followed by free time for playing with and exploring these environments, reflection and discussion (R&D) groups at the end of the day, and SIGs on various topics such as Lego/Logo and art, experiments in motion, computer simulated creatures, knot exploration (Strohecker, 1991), Lego/Logo creatures (known as "weird" creatures (Grannott, 1991), etc. The

³⁶Recursion³⁶

³⁷ A similar but more intensive approach to recursion is developed (for children) by McDougal (1988) in her doctoral dissertation.

recursion SIG was conducted in six three hour sessions – two sessions in each week of the workshop. Approximately fifty teachers attended at least two sessions of the SIG and thirty teachers attended all six. During the workshop, I wrote down my observations and reflections on the SIG in a journal. About three months after the workshop ended, I conducted interviews with five of the participants in an effort to get at what they had learned from the recursion SIG.

The Workshop

In the planning stages of the SWL workshop, a colleague and I proposed to lead a SIG on recursion, chaos, and fractals. This proposal met with an unfavorable initial reaction from our other colleagues. We were told that the subject matter was too "abstract" for the intended audience, that recursion was a difficult topic for experienced programmers and would doubtless be too hard for these teachers who were by and large novice programmers.

My colleague and I had an intuition that this workshop would succeed so we carried on despite the initial discouragement. We designed the workshop to have a very open-ended format, to be focused on a few activities and to leave plenty of time for large group discussion as well as collaboration on projects among SIG members. We also decided that each of us would take greater responsibility for one content area, and that during the time allocated to that content area that one of us would be foreground and the other person would watch the goings on and help "debug" the process. I took on the responsibility for the recursion part of the workshop, and for designing stimulating activities for exploiting recursion.

I started off the first day of the SIG by writing down on the blackboard the following definition of recursion: Definition: Recursion - see Recursion.³⁸ Then I asked the participants to speak about what thoughts came up for them when they heard the word recursion. After the initial ice was broken, a prolonged discussion arose about instances of recursion that people knew in their lives. Though some of the participants had written recursive programs, no one said that recursion was a concept that applied only to computer programs. Some typical examples they came up with were: nested Russian dolls, Logo programs to draw "trees", musical rounds, Escher drawings, etc. Most often, when an example was put forward by someone as being a case of recursion, many others objected and said that this was not a case of recursion.

After this discussion, I gave a brief talk on the history of mathematical ideas connected to recursion. I talked about mathematical induction, self-reference paradoxes, Russell's theory of types, and Gödel.

The rest of our time was divided between a few structured activities and group discussion. One day I asked the class to wander around the campus and come back with as many examples of recursion as they could find. They returned with many examples and were surprised to find that many of their examples were found in nature: snails, leaves, tree branches, etc. On another day I brought in a set of drawings that could be thought of as recursive³⁹. I asked the class to divide into small groups and to think about each picture, give it a name⁴⁰, and then try to come up with some kind of description of it. Many different kinds of descriptions were raised in this activity. Afterwards the class discussed and argued about which descriptions were better, more useful, more elegant. This led to a long discussion about whether shorter descriptions were better and whether

³⁸ For this and other stimulating quotes on recursion see Eric Roberts's (1986) gem of a book: "Thinking Recursively".

³⁹ I first saw the use of such pictures as an aid to thinking about recursion presented by Shafi Giv'on at the 4th international Logo and Mathematics conference (LME4) in Jerusalem.

⁴⁰ Because in the words of Gerald Sussman (one of the creators of the legendary 6.001 - the introductory programming course at MIT): "If you name a spirit, you have power over it."

nature may use recursion because it has limited resources and therefore covets compact descriptions.

Toward the end of the SIG, I asked the class to choose a project that would produce a tangible product in the time left, and to work on it either by themselves or collaborating with others. Most of the class members chose programming projects. Among these were: designing a new and pleasing fractal, reproducing a favored recursive design, writing a program to solve the Towers of Hanoi puzzle, coming up with criteria for when a design was more easily programmed "bottom-up" or "top-down", and many others.

Discussion

The SIG did indeed succeed. Yet our colleagues initial objections were also valid -- most of the participants did find it very hard to learn to program recursively, and if we used this standard alone to judge the success of the workshop it would probably not have met the test.⁴¹ What then do I mean when I say that the workshop was a success? Well for one thing: the teachers were enthusiastic participants. They came regularly, participated actively, formulated their own questions, devised experiments to answer these questions, wrote extensive journal entries about the unfolding of their learning about recursion⁴², and collaborated with each other to form a miniature "recursion culture". In addition, they actively sought supplementary materials: they read books on recursion and chaos, clipped articles from the paper, found videos of the Mandelbrot set, brought in poems, stories, medieval art that they saw as related to recursion. Lastly, and most importantly, there was evidence of a new perspective emerging -- they began to see

⁴¹ I may be being a little harsh here. What I mean by "not meeting the test" is that if I had given them a test consisting of a standard set of exercises that can be solved by recursive programs, I suspect that most of the teachers would not have passed. However, considering that most of the participants had very little programming experience before beginning the SWL workshop, they actually achieved a remarkable level of recursive programming proficiency. I did not expect however, that this level of proficiency would be maintained by the time the interviews were conducted, and this expectation was by and large confirmed.

⁴² One teacher wrote a recursive poem about the SIG and its leaders which she inscribed on a large poster that was hung in the lobby. Another teacher drew a picture of the "recursive leap of faith".

the world in a new way -- they had started concretizing a new concept and using it to interact with the world.

In an interview a few months after the workshop, I asked Rhonda, a music teacher in a suburban elementary school what she had learned from the workshop.

R: Well, I can't remember the details. I'm not sure that I could still solve the Towers of Hanoi puzzle.

U: So what do you remember?

R: Well, it's like I got new glasses ... a new way of looking at the world. At first it was driving me crazy because I saw recursion everywhere. Everywhere I went, everything I saw, seemed to involve recursion.

U: Like what?

R: Oh. Well, like I went to the Mosque exhibit at the museum and there were all these designs in the mosque tiles and when I looked at them they were recursive.

.....

R: And I thought now why would these eighth century folk draw designs of this kind? Obviously these designs were pleasing, but why were they pleasing? And then I thought about how recursive programs are short, they compact a lot of information into a small space.

U: Like we were talking about with the plants?

R: Yeah. And I thought well maybe the human mind evolved to be able to detect high information patterns and that's part of what makes us think things are beautiful. That kind of blew my mind you know? And I'm a musician so I was thinking in what way does this apply to music. And I thought of rounds and canons and I just began to wonder if the brain uses the same mechanisms to appreciate music, art, as it [the brain itself] is built from. Being a biological thing, it's probably got many levels some of which are simpler versions of themselves. So that's a kind of recursion too, the brain is built out of recursion so it appreciates recursion in the world.

In answer to a similar question, Joseph, an inner city fourth grade teacher said:

G: My daughter asked me the other day: if God created the world, who created God? I was just about to answer: "no one did, God just exists" when I thought well maybe God created himself. But is that a kind of recursion? The problem is that like those self-reference paradoxes we talked about there is no way of whittling God down till there is a smallest "Godlet" or "Godling."

And Dave, a middle school science teacher:

D: My kid brother was taking a philosophy of science class in college. After the workshop [SWL], I read some Kuhn and Gleick so I was curious what they were teaching in college. When Al came home for Thanksgiving, I asked him what he was studying. He told me he was learning about explanations, what they were and why they worked. I laughed out loud. I said to him: "that's recursive -- you're trying to explain explanations!"

As can be seen from the above transcripts, teachers in this SIG got engaged with the ideas connected to recursion. They made personal connections between ideas linked to recursion and their own lives.

The following anecdote shows that the approach of this workshop was important, that this was not just a providential group of teachers --a group that would have been motivated in any learning environment⁴³. One day, a visiting mathematician stopped by the workshop. When he heard that there was a SIG on recursion, his interest was piqued, so I invited him to sit in. When he came in, he saw two pictures displayed in the room. The first was a drawing that a teacher, Winn, had made of "the recursive leap of faith". This idea had come up in the context of a discussion on reductionism vs. holism which another teacher, Monica, had read about in "Godel, Escher, Bach" (Hofstadter, 1979). Monica noticed a similarity between the holistic stance and the way one programmed recursively, in that in contrast to a reductionist working out of each case, in recursive programming one took a view of the whole and related the whole program to a slightly similar one. This faith that one didn't have to unpack all the details of the recursive calculation for it to work, became known as the "recursive leap of faith". This phrase struck Winn as particularly meaningful and she was therefore inspired to draw the picture below:

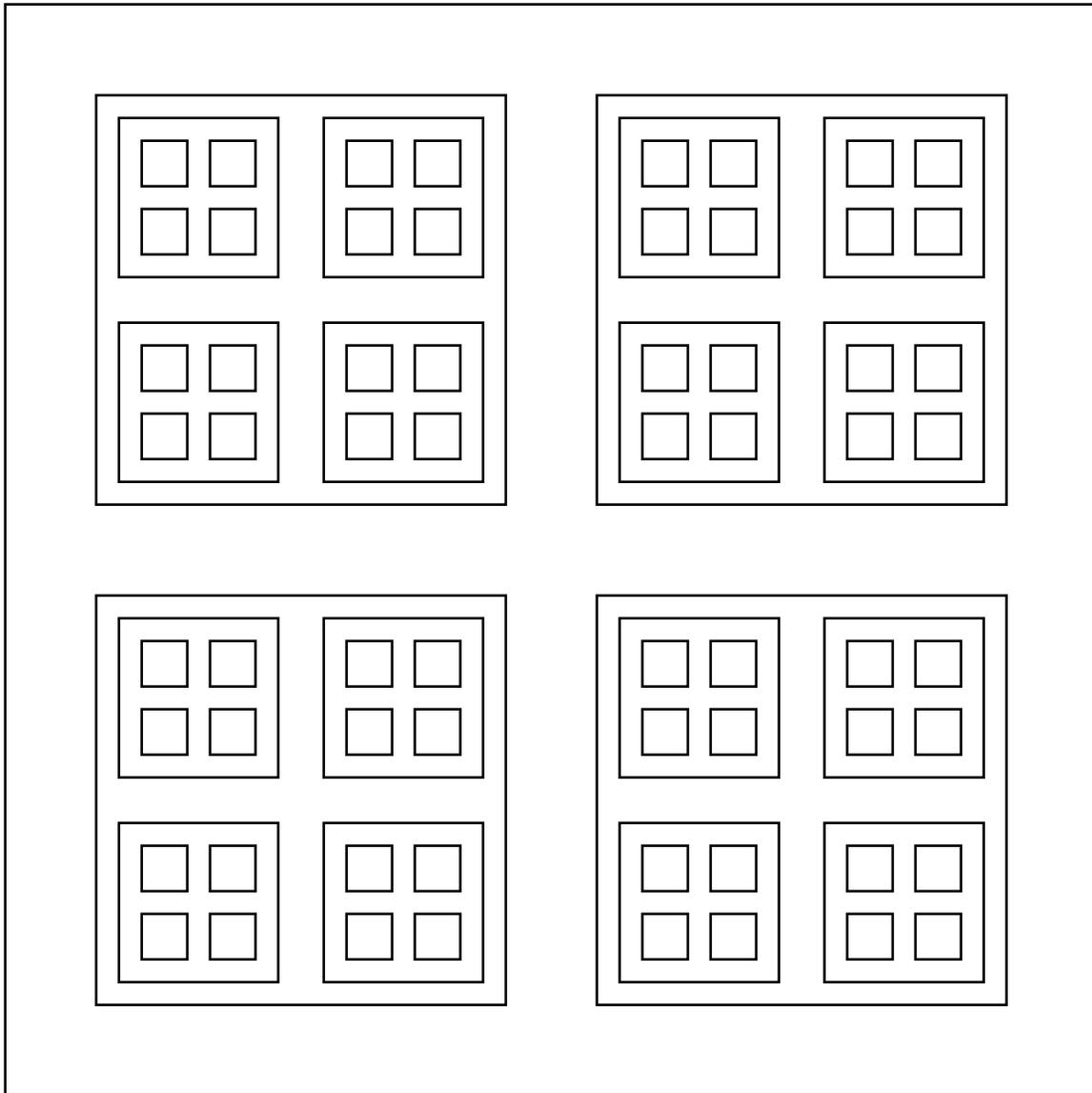
⁴³ Quotations in this anecdote are drawn from memory. Videotaping was planned for these sessions, but it failed to materialize.



Marmon did not like this picture. He called out that this was a picture of a loop, possibly infinite, but not recursive. He then proceeded to go to the blackboard and write down some code fragments to distinguish a linear recursive process from an iterative process.

The second picture that was displayed was one of the "recursive drawings" that I had brought in. One of the teachers had named it a "squish" and above it had written the words: "This is a squish. A squish is a square with 4 squishes in it."

A squish is a square with 4 squishes in it.



Marmon did not like this definition either. Once again he came to the blackboard and talked for 10 minutes or so, on the necessity for a terminal step in a recursive definition. He wrote down more code, along with some equations and logical formulae to try to drive home the point that this was an incorrect definition and that no finite picture could accurately be called a recursive picture.

The class reaction to Marmon's lesson was loud and vocal. "What is he talking about?", "Why is he writing down all these formulae?", "This is exactly what turned me off to math in school".

Marmon decided he had enough of this "class" and departed the room. Later he told me that he was dismayed that I was teaching "incorrect mathematics". "You must be very careful to get the details completely correct", he told me. "Otherwise you introduce damaging misconceptions that will be very difficult to root out". "Also, don't mix ideas from art, philosophy, and other sciences into a lesson on recursion. It will just confuse matters more than they already are".

What can we learn from this anecdote? It is clear that Marmon's objections did have technical merit. Many of the teachers did not have a clear grasp of the distinction between iterative and recursive processes. They also hadn't yet appreciated how important a size parameter is in any finite recursive procedure. So, perhaps Marmon was right and I was fostering sloppy mathematical thinking?

Obviously, I don't think so. Marmon's objections are just another manifestation of the traditional dogma about how to teach mathematics: Mathematics is to be taught by specifying precisely and formally the definitions, theorems, and procedures that most parsimoniously express the underlying ideas. Mathematical concepts should be specified abstractly and uniquely and not conflated with their look-alikes and applications in the observable world. Furthermore, steps in a proof should follow linearly from previous steps as in a logical derivation.

As can be seen again from the teachers' reaction, this approach to teaching mathematics was not effective. It removes the connections between mathematical ideas and their real-world counterparts. It deletes all the analogies and problems that provide motivation for the generation of these ideas. Furthermore, it removes the ideas completely from the historical or developmental context of their discovery, and by doing so hides from the learner all the Lakatosian zig-zags (Lakatos, 1976), the examples and counterexamples, the long history of negotiated meaning whereby these concepts come alive. In a very real sense, I claim, this educational strategy creates a dead mathematics – one disconnected from all its earthly ties and incapable of provoking any interest in the learner.

The intent of the dogmatic math educator is good. He wants to remove any extraneous distractions from the matter at hand. He wants to present the cleanest, the most elegant, the minimal definition, one that will lead to all the right examples working out and banish the counterexamples to the land of never-considered. But his effect is disastrous. By segregating mathematical concepts from their cultural counterparts, he does not remove distractions, he removes any relevance to the learner. Paradoxically, the more distractions of this kind, the more linking between concepts, the better as far as learning is concerned. Even if the linked concept is not familiar to the learner, learning both concepts at once can be easier than learning either separately if they synergize -- help provide a context and justification for each other's relevance. Similarly, if the candidate examples and counterexamples that naturally arise when trying to understand recursion are omitted entirely, it is very difficult to see why any formal definition of recursion is formalized in exactly this way rather than an alternative, and why theorems that employ this formal definition are formulated with exactly these conditions and not others.

This educational process transforms appreciation for the creativity of making a new distinction that results in a new definition or condition into dumbfoundedness at the sheer unnaturalness of the definition or theorem. "Where did that come from?", I've heard in numerous mathematics classes. In some cases rather than turning the learner

completely off from doing mathematics, this unmotivatedness can inspire an awe of the mathematician. "Only a genius could have come up with this amazing idea." Some mathematicians, out of vanity, are not above encouraging such notions so that they can bask in the reflected glory and augment their own status as "genius mathematicians".

Chapter VII - Probability: Settings and Rationales

Why Probability?

At least four motivations have guided my decision to make the probability study the major focus of this investigation. First, although concepts and intuitions from probability theory are ubiquitous in many aspects of our "non-mathematical" lives, most people mistrust the subject matter, their probabilistic intuitions, and all those who cite statistics in order to persuade. Second, the theory of probability presents subtle and interesting issues at its foundation which, I believe, are closely related to learners' difficulties in the area. Third, I wish to challenge the response of educators to the research of Tversky and Kahneman (1982) which has been used to argue for a teaching method in probability that is the diametrical opposite of what I espouse here. Fourth, I wish to take advantage of the new computational field of emergent phenomena, which offers a rich array of (essentially probabilistic) unexpected results that offer challenge to our deterministic intuitions. These reasons are set out in detail in the five sections below.

The Subject Everyone Loves to Hate

Concepts and intuitions from probability theory are ubiquitous in many aspects of our "non-mathematical" lives. When we decide whether to pay more for a car that is reported to be "more reliable"; when we choose to safeguard our children by living in a "low-crime" neighborhood; when we change our diet and exercise habits to lower the risk of heart disease; when we accept or refuse a job with other options pending; all the above actions involve implicit probability judgments.

Arguably, no other area of mathematics can be said to apply so widely, to be so potentially useful in making decisions, both intimate and professional, and affords such opportunities to make sense of phenomena, both natural and social, and empower us through the knowledge we gain. Yet, despite this potential for empowerment, people's

attitude towards probability can be summed up in the well worn adage: “There are three kinds of lies”, said Disraeli, “lies, damn lies, and statistics.”

Probability courses are anathema to most graduate students. Students in the social sciences are required to take a statistics course, and frequently report being able to work the textbook problems but having no idea “what they mean,” or how to apply them to novel situations (see, e.g., Phillips, 1988). They also say that they can make many different arguments to solve a problem and they all can sound plausible, but they give different answers -- how to choose among the plausible alternatives?

Unresolved Philosophical Underpinnings

A second source of reasons to look more carefully at probability comes out of the philosophy of science. As we shall see below, the meanings of the basic notions of probability theory, the ideas of "randomness", "distribution", and "probability" itself are still quite controversial. We are living in a time when these meanings are not yet fixed but are being negotiated by mathematicians and philosophers⁴⁴. Due to the lack of cultural construction of basic probabilistic concepts, we are, as probabilistic learners, in a similar position to the pre-conservational child. As teachers of children, having the experience of conflicting probabilistic arguments that we cannot resolve, serves as an empathy pump, reminding us of how difficult the acquisition of new mathematical ideas can be, of the slow pace of conceptual change, and that a post-conservational perspective is inadequate for helping learners move beyond their pre-conservational conceptions.

Epistemology of Probability

⁴⁴ Many historians of science would say that the real period of "probabilistic revolution" occurred in the 1830's. (see, e.g., Hacking, 1990; Cohen, 1990)

Among the dominant interpretations of the meaning of probability we can identify four schools:

propensitists: believe that probabilities are essential properties of objects. Just as a coin has mass, volume, and color, so it also has another property which is a propensity when thrown to land on heads 50% of the time and tails 50% of the time. Discovering these properties can be done in many different ways but is analogous to discovering other objective properties.

frequentists: probabilities are limiting ratios of frequencies. When tossing a coin many times, we record the ratio of numbers of heads to number of tosses. As the number of tosses increases without bound, this ratio approaches the probability of throwing a head.

subjectivists: probabilities are degrees of belief. When tossing a coin, the probability of its coming up heads is relative to the beliefs of the coin tosser. We can measure subjective degrees of belief by giving people bets at various odds and seeing which they judge to be fair bets. So, in assessing the degree of their belief in the proposition "the next toss of the coin will be heads", we can offer them the choice of taking \$5 and "running" or taking \$10 if the next toss is heads and zero if it's tails. If they judge this choice to be a toss-up (pun intended) then we say their degree of belief in the proposition is $1/2$. If they'd rather have the \$5, we say their degree of belief is less than $1/2$.

credibilists or Bayesians this view identifies probability with rational degree of belief given the evidence. Accordingly, probabilities are not just subjective degrees of belief, nor are they objective properties of objects, rather they are those degrees of belief which a rational person would hold given the evidence she has available. Since two people might have different evidence, they might have different probabilities for the same

event⁴⁵. In judging the probability of a coin toss, one person seeing a streak of heads may judge that the probability of the next toss being heads is still $1/2$, while another may judge that the probability of heads is greater than $1/2$. They could both have rational degrees of belief if one has evidence of the common circulation of biased coins and the other does not.

This controversy may seem rather theoretical, but consider what these differences mean for how we understand and apply probability to our daily lives.

If we are naive frequentists, then, for most situations in our life, probability is irrelevant. After all, probabilities are limiting ratios of frequencies. But, for an event to have a limiting frequency ratio, it must be repeatable. Usually, when we act in the world, we consider ourselves to be in a unique situation, one unlikely to exactly repeat itself. Under this view, probability is irrelevant to most life situations. Only in a few narrow contexts, such as gambling, can the ideas of probability apply to our life decisions. Thus, the naive frequentist, to secure the bedrock of well defined probability⁴⁶ and normative rules dictating “correct action”, sacrifices the connection of probability to most life contexts. Of course, a more sophisticated frequentist might argue that we can still apply probability to a unique situation if we can find the right level of description for that situation. If we find a description that identifies our unique event with other events and therefore forms a class of events, then we can treat the unique event as a repeatable event.

⁴⁵ In this view, a probability is not a property of an object, but rather a property of our knowledge of the object.

⁴⁶ Actually, the bedrock is not quite so solid. Even in "obviously" repeatable situations such as bets at roulette, the frequentist still must specify a criterion for identifying two situations. In order for your roulette bet to be repeatable, subsequent bets must be considered "the same". But in the next bet, someone new is standing next to you, the air movement is slightly different, ... The frequentist is forced to say that these differences don't make a difference and therefore the two bets are equivalent, but determining which variables can be relevant to considering two situations the same or different is a serious problem for the frequentist. The philosopher Nelson Goodman (1983) has shown that one cannot even be sure that a single predicate has the same value in two different situations. As we saw in Chapter IV, deciding when two situations or objects are the same is an act of construction by the individual involved, and different individuals will construct different "sameness" maps.

In contrast, the subjectivist allows all events to have probabilities. Each individual may assign any probability to any event. The price of the richness of application however, is the lack of guidance. Each of us may have a completely different probability for an event to occur. Which is right? How can we evaluate our probability, our beliefs?

At first glance, the Bayesians seem to have the best of both worlds above. On the one hand, they allow all events to have probabilities, and thus probability is relevant. On the other hand, the Bayesian procedure for updating our probabilities gives us rational guidance. Outlandish probabilities will not stand the test of the evidence. But there are problems with the Bayesian view as well. In order to "ground" the Bayesian update formula, each event must be assigned an initial or "a priori" probability. How is this probability to be chosen? One answer is known as the "principle of insufficient reason". It says, that if you don't know how to assign probabilities, if you are "ignorant", then you should assign equal probabilities to all possibilities. But this answer is fraught with paradox. Consider the following example: You want to know if there is life in orbit around a star, say Mithrandir, in a far away galaxy. Since you know nothing about Mithrandir, you decide to assign equal probability to the belief that there is life on Mithrandir as to its contrary. But now consider the question: are there planets surrounding Mithrandir? If, to this, we admit there are three possibilities:

P1: There is life orbiting Mithrandir.

P2: There are planets but no life.

P3: Mithrandir has neither planets nor life.

then we are completely ignorant of which possibility obtains. By the principle of insufficient reason, we should then say that the probability of each event is equal to $1/3$. Therefore, the combination of P2 and P3 will now have probability $2/3$. But the combination of P2 and P3 correspond to the assertion that there is no life orbiting Mithrandir which, by the principle of insufficient reason, was assigned probability $1/2$

before. In this manner we can create completely contradictory probability assignments and it thus appears that there is no rational way to choose our prior probabilities. Bayesians have found various routes around this difficulty⁴⁷ but none have been universally adopted. Furthermore, in order for Bayesians to be able to use their formula for updating probabilities, they must, like the frequentists, be able to determine when two events are to "count" as the same event. This too is fraught with difficulties.

I maintain that the confusion commonly experienced by first-year probability students is not unrelated to the confusion and controversy surrounding these core notions of probability. The difficulties Tversky and Kahneman report are not merely computational; they reflect a deep epistemological confusion as to what the quantities to be calculated are⁴⁸? If the latter conjecture is well-founded, then building a connected mathematics learning environment for probability might go a long way toward grounding students in the subject matter.

Tversky & Kahneman

Yet a third source of motivation for my focus on probability is the research uncovered by Tversky and Kahneman (1982) on the persistent errors that people make when making judgments under uncertainty. Some of their findings have been outlined earlier in this thesis. Now, we will look at their results in greater detail.

The psychologists Tversky and Kahneman have spawned a vast literature which documents people's systematic biases in reasoning when reasoning about likelihoods.

⁴⁷ Shafer (1976) finds an interesting way around this dilemma which results in a very different notion of probability.

⁴⁸ I have recently come across fascinating research by Cliff Konold (e.g. Konold, 1989; 1991) which begins to suggest an affirmative answer to this question. In asking people questions such as "The weatherman predicts there's a 70% chance of rain tomorrow and it doesn't rain, what can you say about his prediction?", many subjects reported that the weatherman was wrong. According to Konold this shows they were understanding probability as a way of predicting the next single outcome. In doing that, they anchored all probabilities to three basic quantities: 0 = no possibility, 1 = certain, and 1/2 = don't know. In the weatherman case, 70% was interpreted as anchored to 1.

Consider the following example:

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements by likelihood:

- 1) Linda is a bank teller.*
- 2) Linda is active in the feminist movement.*
- 3) Linda is a bank teller and is active in the feminist movement.*

Tversky & Kahneman analyzed people's responses to questions of this type. A persistent "misconception" that respondents exhibited was to rank Linda's being a bank teller and a feminist as more likely than just being a banker. This ranking is in violation of the rules of logic which require that any statement is more likely than its conjunction with another statement. Yet, even when the respondents had sophisticated knowledge in logic and probability they were somewhat seduced by the incorrect answer - seeing it as intuitively right and trying to find some way to justify it⁴⁹.

Tversky & Kahneman explained this misconception as stemming from people's use of a "representativeness" heuristic in making likelihood judgments. By a "representativeness heuristic" they mean that people look for an ideal type that represents their answer and then judge probability by closeness to this type. In the example of Linda, the text leads us to represent Linda as a someone who is likely to be a feminist and unlikely to be a banker. (So, when we judge the likelihood of the three statements, we see

⁴⁹ Some researchers have compared this phenomenon to visual illusions such as the face/vase illusion. They conceptualize Tversky & Kahnemann kinds of examples as "cognitive illusions".

that it is unlikely that she is a banker, likely that she is a feminist, and yes perhaps she could be a banker because maybe it was the only job she could get, if she's still true to the feminist type.)

Tversky and Kahneman have collected many examples of this type. In one experiment, people are asked whether there are more words in the English language that begin with “r” or that have “r” as their third letter. Most people say there are more words that begin with “r” whereas in fact there are many more of the latter kind. In this case, Tversky and Kahneman argue that the error is attributable to the heuristic of "availability". People can much more easily "retrieve" or recall words that begin with “r” than words with “r” in the third position. Since the words beginning with “r” are more available to them, they judge them more likely. Another way to describe the availability heuristic is to think of people who are recalling words with “r” in it as conducting sampling experiments. They sample a few words with “r” in them and then compute the relative frequency of “r” in the first or third position. Under this interpretation, the error that people make is in wrongly attributing randomness to their procedures for generating samples of words with “r” in them.

Yet a third example of the errors Tversky & Kahneman report:

A group of “subjects” was given the following story:

A panel of psychologists have interviewed and administered personality tests to a group of 30 engineers and 70 lawyers, all successful in their respective fields. On the basis of this information, thumbnail descriptions of the 30 engineers and 70 lawyers have been written. You will find on the form five descriptions, chosen at random from the 100 available descriptions. For each description, please indicate your probability that each person is an engineer, on a scale from 0 to 100.

Subjects in another large group were given the same exact story except that there were 30 lawyers and 70 engineers. Both groups were then presented with 5 descriptions. For example:

Jack is a 45 year old man. He is married and has 4 children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues, and spends much of his free time on his many hobbies, which include home carpentry, sailing and mathematical puzzles.

Subjects in both groups judged Jack to be much more likely to be an engineer. The data indicate that the prior probability or "base rate" of lawyers or engineers did not make an appreciable difference. But when the same subjects were asked to make the same judgments in the absence of a personality description, they did use the base rates. Tversky & Kahneman concluded that in the presence of specific descriptions, prior probabilities are ignored.

Because these systematic errors are repeatable and don't seem to go away even when people have had significant training in probability, there is a widespread belief that humans are incapable of thinking intuitively about probability. As many tell the story, the human brain evolved at a time when probabilistically accurate judgments were not required and, consequently, resorted to heuristic shortcuts that were not so taxing on mental resources. As a result, we have been "hard-wired" not be able to think about probability and must circumvent our natural thinking processes in order to overcome this liability.

Tversky & Kahneman speculate as to the origin of the systematic biases they uncovered in people's assessment of likelihoods. They theorize that people's mental resources are too limited to be able to generate probabilistically accurate judgments.

People are forced to fall back on computationally simpler heuristics such as the representativeness heuristic they describe.

This view has become very influential and has spawned a large literature. Interpreters of Tversky and Kahneman seem to come in two varieties: those who make the strong claim that our brains are simply not wired for doing probability, that evolution did not spend our mental resources so profligately, and those who simply hold the weaker claim that as far as probability is concerned, our intuitions are suspect. However, the effect of these claims on probability education has been the same -- a reliance on formal methods and a distrust of intuitions.

Recently, I took an introductory graduate course in probability and statistics. When we came to the section on inverse probabilities, the professor wrote down Bayes theorem for calculating inverse probabilities and then baldly announced:

"Don't even try to do inverse probabilities in your head. Always use Bayes formula. As Tversky and Kahneman have shown, it is impossible for humans to get an intuitive feel for inverse probabilities"⁵⁰.

Given the prevalence of this formalist view, it would be of great interest if it could be shown that these systematic probabilistic biases could be transformed into good probabilistic intuitions by a suitable learning environment.

And after all, we do have some reasons to doubt that the lack of robust intuitions about the meanings and applications of probabilistic concepts is due to some inherent deficiency in the "wiring" of the human brain. It may instead stem from a lack of concrete experiences from which these intuitions can develop. Rarely, in our everyday lives, do we have direct and controlled access to large numbers of experimental trials, measurements

⁵⁰ Quotation from memory.

of large populations, or repeated assessments of likelihood with feedback. We do regularly assess the probability of specific events occurring. However, when the event either occurs or not, we don't know how to feed this result back into our original assessment. Suppose we assess the probability of some event occurring as say 30%, and the event occurs, we have not gotten much information about the adequacy of our original judgment. Only by repeated trials can we get the feedback we need to evaluate our judgments.

It is a plausible conjecture that the computer (with its large computational resources, capacity to repeat and vary large numbers of trials, ability to show the results of these trials in compressed time and often in visual form) may be an important aid in construction of a learning environment which gives learners the kinds of concrete experiences they need to build solid probabilistic intuitions.

In the following chapter, I will argue that, even though many of the empirical results of Tversky & Kahneman do hold for many people today, concluding from this fact that people are innately unable to think about probability is unwarranted. This argument will rest on two claims, one theoretical and one empirical. The theoretical claim which we discussed in Chapter IV is that people's mathematical intuitions are constructed, not innately given. Both the lack of good learning environments for probability and the cultural and epistemological confusion surrounding probability do not support the construction of good probabilistic intuitions. Personal and cultural development can lead to more sophisticated probabilistic intuitions and greater mathematical understanding. In the interviews presented in the next chapter, we see learners beginning this concretizing process and starting down the road toward development of strong, reliable probabilistic intuitions.

Emergent phenomena

The fourth source directing the inquiry into probability comes from the new fields of systems theory, emergent dynamics, and artificial life. There has been a rash of publications over the past few years about the difficulties people (and scientists) have with understanding emergent phenomena - phenomena that are the result of the interaction of numerous distributed but locally interacting agents. Mitchel Resnick (1991) has written eloquently about these difficulties and postulated the existence of a "centralized mindset" - a globalized tendency to think in terms of top-down, centrally organized, single agent control structures. To help people go beyond this mindset, he designed the language *Logo, which allows the programmer to control thousands of graphic turtles and thousands of "patches" (i.e., small pieces of the environment on which the turtles move. The patches alone can be thought of as cells in a cellular automaton). *Logo provides primitives for basic turtle and patch calculations as well as communication among turtles and patches and interactions between them.⁵¹

One of the difficulties encountered when trying to understand emergent phenomena is that though the resultant pattern is often stably determined, the sequence of steps by which it is reached is not at all deterministic. In one of the examples from Resnick's (1992) doctoral dissertation, a *Logo program simulates the interactions of termites and wood chips. Before the program is run, the patches are seeded randomly with wood chips and termites are scattered randomly on the screen. After running the program, you see the termites picking up chips and putting down chips and after a few minutes piles of chips begin to take clear shape on the screen. How does the program work? If one is operating from a centralized mindset, one might explain the behavior in terms of a planner termite who organizes all the termites into teams and tells them each what to do to achieve this effect. But in fact the program works by two simple rules: at each time step, move randomly. If you come to a wood-chip and are carrying one, then

⁵¹ I have added primitives to *Logo to facilitate working with probability and statistics.

drop it and turn around, if you come to a wood chip and aren't carrying one, then pick one up. At first glance this procedure doesn't seem to be right. You might object: "but the termites will start to make piles and then destroy them - this is no way to build up a set of piles."

A key insight into seeing why this works is to note that the number of piles can never increase. Since termites are always putting down chips on top of already made piles, they can never start a new pile. Since in the long run, the number of piles will decrease, the average pile size must increase and eventually a few large piles appear on the screen⁵².

This example shows some of the characteristic features of emergent phenomena. The overall eventual pattern of the piles can be said to be determined, but on each run of the program, the path that the termites take and the details of the pile formation are quite different.

Emergent phenomena are essentially probabilistic and statistical. Some interesting questions to ask are: Are the difficulties in understanding emergent phenomena due to their probabilistic character? Is there such a thing as a deterministic mindset by analogy with a centralized mindset? Is the change that transpires when we say someone understands probability a matter of incremental knowledge -- a mere mastering of subject matter in a new mathematical area? Or is the change more fundamental -- a global change in the entire way of looking at the world? And if there is such a thing as a probabilistic mindset, does getting there require a quantitative understanding of probability or are qualitative arguments akin to the ones we gave in the termite example sufficient, or even preferable? The beginnings of answers to these questions will emerge from the interviews discussed in the next chapter.

⁵² If the program is run long enough, the number of piles should reduce to 1. It is an interesting piece of mathematics to try to calculate how many iterations this should take. Indeed, one lesson we take from Minsky and Papert's (1969) "Perceptrons", is that, while it is highly desirable to have a qualitative understanding of algorithms like "termites", it is also important to understand their complexity.

Probability Setting

The Probability research has been conducted in a variety of research settings. Seventeen in-depth interviews (typically lasting at least two hours each, and some lasting as long as eighteen hours face to face!) were conducted. The interviewees consisted of seven women and ten men, one high school student, four undergraduates, five graduate students, and seven post-graduates. The interview format was designed to elicit discussion on four main topics:

- 1) How to make sense of basic probabilistic notions such as randomness, distribution, probability?
- 2) How to go about solving particular probability problems? The problems were chosen by me for their potential for raising fundamental issues and likelihood to link to new and related problems. Some were chosen due to their counter-intuitive and paradoxical nature.⁵³
- 3) How to interpret statistics from newspapers, articles or books. "How would you design a study to collect this statistic?"
and
- 4) What is your attitude towards thinking probabilistically?

Interviews were not time limited: they were allowed to continue until they reached a natural conclusion. Protocols for the interview were flexible: a total of 13 separate topics had to be covered before the interview was done, but no rigid order was imposed by the interviewer. A goal of the interview design was to be broad, to get at the relationship of the interviewee to Probability in a variety of contexts and to explore psychological and social dimensions of this relationship as well as the mathematical

dimensions. The interview was experienced by most interviewees as a conversation. Most often, the conversation grew out of the responses of the interviewee to the beginning question: "What does the word probable mean in ordinary discourse, and is this meaning related to the mathematical disciplines of Probability & Statistics ?" Some interviewees started to give their personal accounts of experiences with courses in Probability before the initial question was asked. This was then followed up and became the germ from which the interview sprouted. Along the way the interviewer introduced a variety of prepared questions at points where they seemed natural. Among the topics explored were the relationship of the interviewee to uncertainty (what his/her general tolerance is for uncertainty in decision making), whether the interviewee participated in gambling activities, as well as specific Probability problems such as the "Monty Hall" problem recently publicized in the NY Times. The interviewer also presented statistics from newspaper articles (e.g., the divorce rate in 1987 is 50%, the reliability of a certain form of birth control is 95%), and asked the interviewee to interpret the meaning of the statistic, how it might have been collected, what kind of study would he/she design in order to obtain this statistic.

Great effort was made by the interviewer to combat the inhibitions of the interviewees to talk about their partial and imperfect understandings. The interviewer explained his belief that all mathematical understanding is partial and a tangle of messy concepts. He also modeled the pushing through inhibition by talking about his own mathematical development and the many confusions and subsequent connections made on the road to mathematical understanding.

After the interview topics were covered, all interviewees expressed a desire to talk more about the problems and to "find out what the right answers were." At this point, the interviewer discussed his understanding of the solutions, and in cases where discourse did not settle the matter, experiments were designed by both parties and conducted by the interviewee in order to deepen the understanding of the problem.

In a few cases, the interviews could not be conducted (or completed) face to face either because the interviewee was no longer in the area or because the amount of time needed was more than could be arranged. Some of these interviewees elected to continue the interviews over electronic mail. In particular interviews on the "envelope paradox" were often conducted in this way. In email interviews, I sent each interviewee an initial message describing what I hoped for in our email dialogue. Most interviewees appeared to have no problem adhering to my guidelines and the email medium proved to allow a rich set of exchanges.

In addition to the interviews, some students worked with me in computer learning environments. Five students elected to develop *Logo programs designed to explore some area of probability that they wanted to understand better. In order to facilitate its use as an environment of exploring probability, I made some modifications to *Logo and added some primitives especially useful for probabilistic explorations. Among the programs developed are a microworld for exploring distributions of gas particles in a closed container, an environment for exploring the meaning of bell-shaped and binomial distributions, a test environment for "Monty Hall" type problems, and various projects modeling newspaper statistics, and creating physics and ecology simulations. The nature of these projects, what was learned through both their design and their use will be explored in detail.

Chapter VIII - Probability Interviews

This chapter will describe and analyze the probability interviews that I conducted. These interviews were both designed and conducted so as to emphasize epistemological foundations and meanings of probabilistic concepts. The failure of even the best university mathematics instruction to address students' epistemological obstacles to learning probability may be gleaned from the following anecdote:

Recently I had a chance to test some new probability and statistics software that was developed through the National Science Foundation for curricular use at the university level. The software was indeed impressive, with beautiful graphics, an easy user interface and a large selection of features. A principal emphasis of the software package was on distributions. The package contained many different distributions, both continuous and discrete, each with its own name and associated text describing its characteristics. In addition, for each distribution, users were given a list of parameters which they could manipulate and see how the graph of the "random variable" changed.

This software was used extensively by students in a number of university level courses. After each course was completed, the students were given a probability and statistics test. One of the questions on the test was: "What is a random variable?" A researcher in this software project allowed me to look at the tests from three of these classes. In all three classes not a single student gave a meaningful answer to this question. Most students just left it blank. The most frequent non-blank answer was: "a variable that has equal probability of taking on any of its possible values". When I asked the researcher if this was typical, he reported (Cohen, 1993) that in all the exams he has seen, not a single student had "given the correct answer", nor a single one mentioned the concept of distribution in his/her answer. This despite the fact that they had spent hours manipulating distributions and had plotted and histogrammed their "random variables". Why is this so? I shall argue that one major obstacle to learners' understanding

probability and statistics is the failure to explicitly address the issue of randomness -- what is it, how is it tied to probability distributions? A number of the questions in the interview were designed to draw out these issues.

Some of the interviews described below involved interviewees writing *Logo (pronounced Starlogo) programs. *Logo is a computer language that runs on a massively parallel supercomputer (the CM2). It is an extension of the computer language Logo. Logo provides an environment for controlling a graphical turtle on a computer screen. Children issue commands (such as forward, back, right and left) to the turtle to make it move. *Logo is an extension of Logo which allows the user to control thousands of graphical turtles. As such it is an engaging and powerful environment for exploring and playing with mass phenomena as in probability and statistics. I will not describe Logo or *Logo here. The reader not familiar with these languages should see (Resnick, 1992).

In the next seven sections, I present and analyze the probability interviews that I conducted. To get the most out of these interviews, the reader will have to read them quite closely. The mathematics done in the interviews, while not requiring that much formal background, is quite sophisticated in its use of subtle argument and developing intuitions. As I remarked in the introduction to this thesis, I believe it is only by actually doing the mathematics in these interviews that the reader will be convinced of the quality of the mathematics the interviewees have done.

Normal Distributions - Why is Normal Normal?

Let me introduce Lynn, a psychologist in her mid-thirties who at the time of her interview had recently received her Ph.D. For her dissertation, Lynn had done a statistical comparison study of different treatment modalities for improving the reading comprehension of dyslexic children. While in graduate school, she had taken an introductory probability and statistics class as well as two classes in statistical methods. As such she would "naturally" be classified as

having a fairly sophisticated background. In this interview fragment we will see that despite her sophisticated background, basic ideas of randomness and distribution are alienated for her - not trusted or appropriated. Even though Lynn was capable of applying formal statistical methods in her coursework, she was fundamentally confused about the “madness behind the method”. In this interview, she starts to negotiate meaning for “normal distribution”. Although her attempts take her to a position vis a vis distributions which would be considered just plain wrong in most university courses and may indeed be incoherent, she is for the first time grappling with the meaning of the fundamental ideas of randomness and distribution and in so doing she takes a step towards developing intuitions about and appropriating these concepts.

The first item on the interview that Lynn and I discussed was the question: What would you estimate the probability of a woman in the US being at least 5'5" tall?

Here is the text of the ensuing conversation:

L: Well I guess it would be about 1/2.

U: Why do you say that?

L: Well height is normally distributed and I'd say the mean height of women is about 5'5" so half of the women would be taller than 5'5".

U: Why would half the women be taller than the mean?

L: Because the curve of the normal distribution is symmetric around the mean - so half would be below it and half above it.

U: What about the number that are exactly 5'5" tall?

L: Well I guess they could make a difference, but no - they shouldn't matter because they're just one data point and so can be ignored.

U: You can ignore any one data point?

L: Yes, because there are infinitely many data points so one doesn't make a difference.

U: But can you make infinitely many height discriminations? How do you measure height?

Here, I'm just trying to probe Lynn's thinking about discrete vs. continuous distributions.

L: Well.... I guess we measure it in inches -- so there probably aren't infinitely many data points. I'm somewhat confused because I know height is distributed normally and I know that for normal distributions the probability is 0.5 of being bigger than the mean, but how come you can ignore the bump in the middle? I guess 0.5 is just an estimate, it's approximately 0.5.

U: How do you know height is distributed normally?

L: I don't remember where I first learned it, but lots of the problems in the stats books tell you that.

My question was intended to elicit an explanation of why Lynn believed height was distributed normally in terms of her personal knowledge, but Lynn responded to it as about the context of gaining that knowledge and the authority it derives from.

U: Why do you think height is distributed normally?

L: Come again? (sarcastic)

U: Why is it that women's height can be graphed using a normal curve?

L: That's a strange question.

U: Strange?

*L: No one's ever asked me that before..... (thinking to herself for a while)
I guess there are 2 possible theories: Either it's just a fact about the world, some guy collected a lot of height data and noticed that it fell into a normal shape.....*

U: Or?

L: Or maybe it's just a mathematical trick.

U: A trick? How could it be a trick?

L: Well... Maybe some mathematician somewhere just concocted this crazy function, you know, and decided to say that height fit it.

U: You mean...

L: You know the height data could probably be graphed with lots of different functions and the normal curve was just applied to it by this one guy and now everybody has to use his function.

U: So you're saying that in the one case, it's a fact about the world that height is distributed in a certain way, and in the other case, it's a fact about our descriptions but not about height?

L: Yeah.

U: Well, if you had to commit to one of these theories, which would it be?

L: If I had to choose just one?

U: Yeah.

L: I don't know. That's really interesting. Which theory do I really believe? I guess I've always been uncertain which to believe and it's been there in the background you know, but I don't know. I guess if I had to choose, if I have to choose one, I believe it's a mathematical trick, a mathematician's game.What possible reason could there be for height, ...for nature, to follow some weird bizzaro function?

The above was a short section of the first probability interview I conducted⁵⁴. Until the last exchange transcribed, I had the feeling that the interview was not very interesting for Lynn. But after that last exchange she got very animated and involved in the interview. This question of the reality vs. “trickiness” of the normal and other distributions occupied her for much of the next few discussions. She came back to it again and again.

At the time, I was a bit surprised by this reaction. Earlier in the transcript, when I asked her about infinitely many height discriminations, I think I was trying to steer Lynn into making a distinction between a discrete and continuous distribution. This, I thought, might have resolved her dilemma about ignoring the “bump in the middle”. But I was not expecting to hear such basic doubts about the validity of statistics from someone with so much experience studying and using statistical methods.

Lynn’s question was, as she says, “in the background” all the time when she thought about statistics. How could it not be? At the same time that she was solving probability and statistics problems for homework, and later on when she actually implemented a study which used statistical tests to evaluate treatments, she was never really sure that the formulas she was using had any explanatory power, any real meaning other than mathematical conventions. Is it any wonder, then, that she was engaged by this question? Up until now, no one had engaged her

⁵⁴ Short at least in proportion to the total length of the interview which lasted eight hours and made me realize how much time this kind of research would require.

in discussion about how to interpret the mathematical formulae she was using. She was left feeling that her own work played by the “official rules” but in a way was not accessible to her -- and she harbored doubts about its validity. Her questions, while essentially epistemological, go to the heart of the mathematics of probability. Note that, in emphasizing the normal distribution, her courses had led her to see the process of describing height by a curve as a single process with a single agent (possibly the mathematician) controlling it. The idea of the height distribution as being emergent from many interacting probabilistic factors was not in awareness. In a later section, we will see how another interviewee, building a *Logo program to understand normal distributions, deals with this issue.

Divorce

The newspaper says that the divorce rate (1987) is 50%. What does this statistic mean?

How would you gather the data to compute this statistic?

In the following interview, two interviewees, Amy and Doug, struggle to understand the statistics of divorce rates. It is probably no accident that these two interviewees got so engaged with this question, as they were also both engaged to be married (not to each other).

Both Amy and Doug are undergraduates majoring in engineering. Both have taken university courses in probability and statistics. Earlier, in their interviews both said remarkably similar things when asked about their course - they did well but did not feel like they really “got it”.

The interview text itself is an ambiguous figure. Looked at from the perspective of the traditional mathematics curriculum, Amy and Doug are floundering. They are very confused about the meanings of the statistics they have and their progress is agonizingly slow and incomplete. Why not just tell them what the divorce statistic means? If we were just clear with them about the exact meaning of the statistic, then they would be able to make much faster progress and would not flounder so much.

But looked at from a connected mathematics perspective, Amy and Doug are succeeding in doing mathematics. They are engaged with a mathematical problem and in making meaning of the statistics in the problem. In the course of their interview, they spawn off new problems which are relevant to their question and pursue them. They interact with each other to form a community of shared goals and representations. Throughout the interview, again and again they come back to such questions as: What does a statistic mean? What is a good statistic to have and what makes it good? How can a statistic be collected? We see them as they construct their own statistics, compare and contrast different statistics, invent new representations of their knowledge, all the while using the computer as a natural extension of their thinking and a tool for verifying or invalidating their conjectures and doing experimental mathematics.

In choosing between these two perspectives of the interview, it is helpful to remind ourselves of Amy and Doug's previous experience. Both of them have had formal courses in probability and statistics. As such, they have solved numerous exercises in probability akin to the divorce rate problem formally posed. Yet this previous experience was not sufficient to guide them through this interview. The statistic they were trying to understand came directly from the daily newspaper. Unspecified statistics of this kind are reported regularly in the newspaper and often in scientific papers. All the instruction Amy and Doug received in their many years of schooling was not sufficient to help them interpret, connect, and understand the statistics they encounter in their daily lives.

Amy and Doug got absorbed by the divorce rate question and decided to try to work it out together. Here are some excerpts from their conversations. The transcripts have been "cleaned up" some (removing pauses, umms and many interjections) for clarity of the exposition.

Doug: Well I guess it's straightforward enough -- half of the people who got married in 1987 will divorce.

Amy: By when? Eventually?

Doug: Yeah.

Amy: But how can they know that? Many of the people haven't yet divorced?

Doug: Well, let's see. Suppose you've got the number of people who got married in 1987 and got divorced in 1987.

Amy: OK

Doug: And you have the number of those married in 1987 who got divorced in 1988, 1989, 1990 and 1991.

Amy anticipates Doug's argument.

Amy: So you're gonna extrapolate?

Doug: You have the beginnings of a curve. You can then complete the curve to see how many people will eventually get divorced?

Amy: But how do you complete the curve? You don't know what shape it is.

(5 minutes deleted. Amy and Doug discuss whether it is a bell shaped curve.)

Note that a bell curve is the anchor distribution - often selected whether it is appropriate or not. In another interview question, about the shape of the distribution of American cars on the road by model year, many interviewees drew a bell curve when in fact the curve is highly asymmetric and monotonically increasing. Amy and Doug had both worked through that question and were wary of “defaulting” to the normal distribution.

Doug: Well there must be some standard shape for divorce curves.

This is a crucial tension in their conversation. Doug seems quite convinced that whether they know the shape or not, divorce statistics must have some characteristic shape, there must be a lawlike regularity to divorce frequency, but Amy is unsure why this should be so and alternates between accepting Doug’s framework and questioning it. But both of them, like Lynn in the previous interview, have an unease about distributions that lies in the background. They feel this unease about divorce distributions, know it is there, but repress it. It is easy to follow a garden path, that there is something called a divorce curve and all we do have to do is fit the curve. But, left unexamined is the status of this curve - what kind of thing is it and what makes it takes its shape?

Amy: I don’t know how long they’ve been keeping statistics on divorces but if you could get the stats on divorces say from marriages in 1900, then you could see the shape of that curve, and then fit the beginning of this curve onto that curve and get the answer that way.

Doug: Yeah that would work. You’d get the shape from the old data but the slopes would be different and so you’d get a different total answer.

Amy: It’s a bit weird though.

Doug: Weird?

Amy: Well, we’re getting the beginning of the curve from current stats but we’re getting the shape from old stats. What justifies doing that? I mean what we’re saying is that the number of marriages dissolving in their beginning stages has changed from 1900 to 1988. But we’re also saying that that’s the only thing that has changed. The relative number of divorces from the n th year of marriage to the n -plus-first year of marriage stays constant from 1900 to 1988. Why should that be?

Doug: I guess because that's the shape of the divorce distribution. Just like if you take different height samples you always get a bell curve.

Amy: So what you're saying is: just like a bell curve is characteristic of height, this shape we're postulating is characteristic of divorce? And the first few years of data are sufficient for us to calculate the curve's parameters?

Doug: Yeah, like mean and standard deviation.

Amy: I'm not yet convinced. Don't you need something more to just say that divorce has its characteristic distribution? Why don't we look at some data samples and see if they all conform to a recognizable shape?

Here Amy is expressing skepticism that divorce, like height, has to have a characteristic distribution. What, she implies, would entitle them to make that assumption? Do they need a physical model of the mechanism creating a divorce distribution? At least, she suggests, they can try to look up some data and see if in fact, there is a characteristic shape.

Doug: I like that idea. I think Dewey's ⁵⁵still open.

(Two days later - Amy and Doug had gone to the library (each separately) and found some statistical data about divorces.)

Amy: Here it says that the divorce rate in 1987 is 4.8 per thousand population. What does that mean?

Doug: That 4.8 out of every thousand couples getting married in 1987 have gotten divorced so far? No, that's totally wrong.

Amy: 4.8 out of every thousand couples got divorced in the calendar year 1987. How do they count couples? Do you also add in the additional people who got married at any point during that year. And why do they say population? Do they mean out of every thousand people?

Doug: That would be weird - some of those people are unmarried. How can they count them in the divorce stats?

Amy: ... Maybe they need to compare these stats with other stats - you know like death rates or something so they want them in a uniform format?

⁵⁵ An MIT library

Doug: That could be. I guess they might want to compare divorces per population to ski accidents per population or something.

Amy: But it's still a bizarre statistic - it mixes up people who got married in very different years. So of 1000 marriages in 1987, I mean people, people in 1987, 4.8 got divorced that year. How are we going to figure from that any information on how many people who got married in a particular year, say, 1980 got divorced?

(10 minutes deleted)

Doug: Well, suppose the divorce rate is constant at 4.8.

Amy: But it's not. In 1986 it was 5.0 and ...

Doug: But let's just try to simplify that way.

Amy: OK.

Doug: So each year you're married you have a 0.048 chance of getting divorced.

Amy: 0.0048.

Doug: No, it's 0.048. Oh you're right, 4.8/1000. That's 0.0048.

Amy: Is that right? Isn't it half that? The "4.8" is divorces out of a thousand people but divorces involve two people.

Doug: Is it half or is it twice?

Amy: No, I see, each time a person gets divorced their partner has to get divorced as well.

Doug: Divorce is a symmetric relationship?

Amy: (Laughs) About the only symmetric thing in a marriage... But going back I think that means we can use the 0.0048 figure. If there's a thousand of you and each one of you has a 0.0048 chance of getting divorced then 4.8 of you or them or whatever will get divorced. So that works out.

Doug: OK then. We're home free. All we have to know is how long a typical marriage lasts and then we can figure out what the probability of getting divorced is.

Amy: So you're saying flip a coin. A coin that has 0.0048 chance of coming up heads.

Doug: Yeah, each year.

Amy: And if it comes out heads, you get divorced. If it's tails, you stay married. Do it for "n" years and you get the divorce probability.

Doug: Say typical life span is 70 years. Then we compute 0.0048^{70} and that's the divorce probability. Wait a minute, that can't be right - that's tiny?

Amy and Doug: one minus 0.0048!

Doug: To the seventieth power.

Amy: More like 40 or 50. We're assuming people are married, so they're at least 20.

Doug: We gotta know the average life span of a marriage. I mean a non-divorced marriage. If we say it's roughly 40, then (gets out calculator) $1 - 0.0048 = 0.9952$. And $0.9952^{40} = 0.8249$. Whoa, does that mean 82% get divorced? No, let's see. Wait a minute. 0.9952 is the chance of staying married and so 82% is the probability of still being married after 40 years. That's too high. It was supposed to be 50%.

Amy: Well 40 may have been too conservative. Do we have the statistics for the average age of marriage and the average age of death? Also, even if we had those then it might not be that good cause people who are married have a higher average life expectancy. I'm not sure how valid it is to keep sticking in averages here. I'm somewhat unsure about this whole approach. We keep replacing individuals with averages and that may not work out right.

Doug: Well how high could it be? Say it was 60 years. It can't be more than that. $0.9952^{60} = 0.7492$. Wow, even for 60 years only 25% get divorced. Let's try 80.

(Doug tries higher and higher values)

Doug: $0.9952^{140} = 0.5099$. Something's very wrong. It should take over 140 years to have a 50% chance of getting divorced!

Amy: Our model was too simplified. It's not a 0.0048 chance of getting divorced each year.

Doug: But look at the chart - the highest divorce rate was in 1981 and that was 0.0053. $1 - 0.0053 = .9947$. (calculates for a few minutes). At the highest rate it would still take over 130 years.

Amy: But we've assumed that the rate is constant each year. We know that the 4.8 figure is composed of many different marriage durations. Remember you were saying that the divorce curve would have a characteristic shape. We've assumed that it is flat but that doesn't make sense. I don't know how to think about this problem without the simplifications though.

Doug (facing Uri): We're stuck. I don't know if we can get any further with this. Any hints?

Uri: Want to write a *Logo program to think about this?

Amy: That sounds like a good idea.

Doug: Yeah. Let's give it a try.

Uri: Before going over to the CM, why don't you guys work out a plan? I'll then help you program it.

.....

Amy: OK. So we start with a 1000 turtles representing 1000 people.

Doug: And each turtle rolls a thousand sided dice and if it gets 4.8 or less it gets divorced.

Amy: And we give each turtle an age variable. And at a certain age they get married.

Doug: So for each age we're gonna have to know the probability of their getting married so they can throw a marriage dice as well and only if they're married will they then throw a divorce dice. Right, so we start out with 1000 turtles born in a given year, say 1900. Each year they throw the marriage dice and the death dice. If they're married they throw the divorce dice. We count up how many get divorced and how many die.

Amy: Is that stat a good one to calculate? I mean wouldn't it be kind of old by the time you used it since you have to wait for all these deaths? Strange again to be mixing deaths and divorces from marriages of different lengths.

Doug: I think my program won't work. I'm getting confused because do these turtles marry each other or what? If they do then I'll be counting divorces twice.

Amy: I have an idea. Let's make every turtle be a marriage. So we start with a 1000 turtles. Now each marriage can "die" in one of two ways - death of one of its members or divorce. That seems cleaner.

Doug: Hmm. That's kind of nice making the marriages the turtles.

Both of them like the new representation. They've created a new unit, a new object, a marriage. By letting a turtle which is a single entity and was previously seen to represent a single individual now represent a marriage, they no longer have to think about the individual people's characteristics, only the "birth" and "death" of the marriage.

Amy: So we can then just count how many marriages end in divorce and how many end in death and they should come out to be equal if the newspaper was right.

Doug: That way we get rid of the marriage dice. So now there are just two dice the divorce dice and the death dice. We have to know two things then: the probability of getting divorced at a certain age

Amy: The age of the marriage.

Now Amy seems to really be thinking of marriages as living aging beings.

Doug: Yeah, right, the age of the marriage. And the probability of dying at a certain marriage age. The first one seems pretty look-uptable but the second looks hard. We'd have to know the average age of each partner and each one's life expectancy. I don't know. Maybe since guys tend to be older and live less time, we can just use life-expectancy tables for the men and ignore the females since they'll rarely die first?

.....

Amy: Or maybe we can use your "average" idea. Just give each turtle an average life expectancy. So start each turtle at age 0 and then kill them all when they get to say 75, actually it would be more like 50, and then do the counting. But I'm still not sure that will work. What if we just give each turtle its own life expectancy? Say each turtle is born in 1980 which means its couple got married in 1980 but since the couple could be all different ages, we give some of the turtles life-expectancies of 10 years and others 70 years.

Doug: So we have to find a graph, a histogram which tells us what percentage of marriages die in years 1 to 70. I guess that might be hard to find. Cause some of the marriages end in divorce. What we need is the ones that don't end in divorce.

Amy: I need a way to understand the difference between giving the whole population the characteristics of the average and assigning them through some distribution function. Let's try writing a program.

With my help Amy and Doug then embarked on writing a *Logo program to answer Amy's question. They started by creating 1000 turtles and giving each of them an age of 0. To simplify matters they got rid of the "age" variable and set up each turtle with a life-expectancy of 40. At each step the turtle "rolled" a random number to see if it got divorced (Their *Logo procedure "divorced" was: to divorced? random 10000 < 48 end) and it decreased its life expectancy by 1. If its life expectancy got to 0 before divorce then the program incremented the "dead" counter. If it got divorced first then they incremented the "divorced" counter. In either case the turtle died and a new turtle (couple) came into being with the same life expectancy. They computed the probability of divorce to be $\frac{\text{divorced}}{\text{divorced} + \text{dead}}$. They ran this program and it "confirmed" that approximately 18% of the population got divorced using this model. They then tried to answer Amy's question about substituting averages for distributions. They ran the

program changing the life-expectancy from a constant to a variable specified by $20 +$ random 61⁵⁶. This gave an average life expectancy of 40 but each turtle had a different life expectancy. The result of this run was much the same as the run before. They then experimented with fixed life-expectancy and varied probability of divorce depending on marriage age. They worried that they were only taking into account marriage age and not the interacting factor of the age of the couple, but they proceeded on anyway. In their almanac, they were able to look up some data from which they concluded that the probability of divorce was much higher in the early stages of marriage then later on. They then reimplemented their “divorced?” predicate as the procedure “variable-divorced?”:

to variable-divorced?

```
if age < 3 [make "prob 41]
```

```
if between age 3 10 [make "prob 120]
```

```
if between age 10 20 [make "prob 75]
```

```
if age >= 20 [make "prob 10]
```

```
random 10000 < prob
```

```
end57
```

To their surprise, this did not substantially change their results. They had hoped that this was the missing factor that would account for the discrepancy and increase the divorce probability. They ran many experiments to try to understand why the result didn't change. In one experiment they assumed marriages lasted a potential 2 years. In each year they had a 50% chance of divorce. This led to a 75% divorce probability. But, if they

⁵⁶ Giving a flat distribution of life expectancies. Later I added a primitive to *Logo to make it easier to create bell-shaped distributions.

⁵⁷ This procedure took some time to write. In order for “variable-divorced?” to give the same average probability as “divorced?” the sum of the products of its constituent probabilities with their age ranges must equal $39 * 48 = 1872$. This was a natural motivation for thinking about expected values.

assumed that 90% got divorced the first year and 10% the second year, even though the average was the same over the two years, the probability of divorce over the two years increased substantially. This seemed like cause for hope in getting the divorce probability up if they could find a “bug” in the variable-divorced code. But they could not find a bug and eventually concluded that for very small probabilities, the average was a reasonable estimate of any distribution⁵⁸.

After some more interesting and productive side investigations, Amy and Doug were still unable to find a way to square the results of their model with the initial statistic. They decided something had to be wrong with the data rather than with their reasoning. Looking back over the data they discovered a “bug”. They had taken 0.0048 to be the probability of a couple getting divorced in a given year. But actually their own interpretation of the statistic was that 4.8 out of every 1000 people got divorced. As Doug had said, this includes people who are not married, so to get the probability for a married couple, one must know the proportion of married people in the general population. This they were able to look up and determined was equal to .454. Therefore the probability they should have used in their simulation was equal to $\frac{.0048}{.454} = .01057$. Substituting this value in their simulation led to a 35% divorce probability which they decided was sufficiently close to believable and resolved their dilemma.

⁵⁸ One way to frame the issue about averages that so concerned Amy is to ask when a product of “n” numbers with a given average, S, is maximal. It is not hard to show that this occurs when all the numbers are equal to $\frac{S}{n}$. Since marriage probabilities are multiplied in this way, Amy and Doug correctly intuited that if they replaced the turtles’ uniform probability of 0.0048 with a distribution that had mean 0.0048, then the product of the corresponding marriage probabilities (having mean 0.9952) would decrease. But a decrease in the marriage probabilities would have the “desired effect” (only in terms of their own goals - let me not be accused of being anti-family) of increasing the divorce probability. Amy started to reason in this way when she said “but x^2 is always bigger than $(x-a)(x+a)$ since that’s $x^2 - a^2$. However, they got confused when thinking about more than two terms and weren’t entirely sure that this held in the general case. They might have gotten an insight into their eventual conclusion that small deviations don’t make much of a difference by noting that even in their binomial case, the effect on the product of creating two deviations is “second order” i.e. the difference between the two products is proportional to the square of the deviation and so small deviations will not greatly affect the result. I am indebted to Aaron Brandes for stimulating discussion concerning this point.

To me, Amy and Doug's last solution still seemed to have a "bug" in it. I chose, however, not to intervene at this point.

What stands out for me in this interview are the many obstacles that are in the way of Amy and Doug's understanding of divorce rates. Among these obstacles is their lack of information as to how the data was collected, the ambiguity of the newspaper and almanac statistics and their ignorance of certain mathematical theorems. But even more salient are the epistemological obstacles. Amy and Doug are unsure what ontological and epistemological status to give to the concept of a divorce distribution. What kind of thing is it? How do we know that it is stable? What mechanism causes divorces to happen at a characteristic rate?

The mathematics of divorce rates is messy. There is no one best way to represent divorce rates. The choice of representation is guided by other mathematical representations to which we wish to connect, which assumptions we are willing to make about the data, the cost we are willing to incur for errors in these assumptions, and how we will collect and use the data. These kinds of questions cannot be meaningfully answered by learning formulas. These questions must be responded to by making representations, connecting and contrasting these representations, subjecting them to criticism, criticizing the critiques, and building a collection of tools, of probabilistic intuitions.

Monty Hall

Monty Hall on the TV show "Let's make a Deal" gives you a choice of 3 boxes. One of the boxes contains a valuable prize (a Ferrari), the other two contain undesirable prizes (goats) . You choose box #1. Monty, who knows what's in each box, opens box

#2 intentionally revealing a goat. He then makes you an offer: you can keep box #1 or you can trade it for box #3. Which should you do?

This problem has many variants and has recently gotten much attention in the mainstream press. A distinctive feature of the problem is that almost all people, upon initially hearing the problem, assert that it makes no difference whether you switch or not - “the odds are 50 50”. In its sharp form, the Monty Hall problem is a formal problem and has a correct answer, namely that switching is better with 2 to 1 odds. But even when people accept the formal terms and are told the “correct answer” and arguments for why it is correct, they do not accept it. I have both witnessed and participated in many arguments where both sides are convinced that their answer is the correct one. Typically, the contenders argue till they are blue in the face with neither side convincing the other.

In this, it bears a strong resemblance to Piaget’s conservation experiments such as the water volume experiment where pre-conservation children can not be argued into accepting that the two glasses have the same amount of water. For the purposes of this exposition, I’m assuming the reader is “post-Monty-conservational”.

From a “post-Monty-conservation” perspective some of the answers to Monty Hall interview questions seem quite strange. Consider Chris:

U: So what’s better, sticking or switching?

C: It doesn’t matter. It’s 50-50.

U: Why is it 50-50?

C: (a bit exasperated) There are two boxes so it’s 50-50.

U: OK, How about if I change the choice. Choose between box #1 and both boxes #2 and #3.

C: Well, then I’d take both boxes cause it’s a $\frac{2}{3}$ chance.

U: OK, [sensing that this argument will be convincing] Suppose I take box #2 and box #3 and put a big box around them. Now there are only 2 boxes, box #1 and box #4. Which one do you prefer now?

C: It doesn't matter. It's 50 -50 again.

This last answer surprised me. I had expected Chris to say that the "big" box was better. Her answer though reminded me of children who say that the tall glass has more, then when its contents are poured into another wide glass they say both glasses have the same. Similarly, when Chris saw the contents of the two boxes "poured" into one big box, she switched from seeing the two boxes as better than the single box to seeing the two boxes, big and little, as being the same. It is not hard to propose an agent model of the Monty Hall problem quite analogous to the one proposed by Minsky and Papert for water volume conservation. In fact we could integrate the Monty Hall problem into the Society-of-More by adding a few more agents to that society. I present two such analyses in (Wilensky, in preparation).

Sometimes, to facilitate thinking in terms of dollar value, I presented the Monty Hall problem as: One box contains \$99 and the other two boxes contain nothing.

It then became apparent that the idea of "probabilistic box worth" was critical in the conservation. Most people do not think of these boxes as having a worth. The new probabilistic idea of box worth must be constructed by the learner. Once constructed, it is connected to the non-probabilistic notions of worth and thus inherits a lot of power. If we start to evaluate the worth of the boxes, in accordance with expected value, we see that the box we chose was initially worth \$33 ($1/3 * \99). The other two boxes combined are worth \$66 ($2/3 * 99$). But at least one of these boxes definitely has \$0 in it. If we deliberately open that box, then we have not changed the worth of the two boxes (box worth is conserved!), hence its partner box must still be worth \$66. As was said above, most pre-conservational interviewees do not think of the box they have chosen as having a worth. Thus, they respond like Chris in the following fragment:

U: So, how much is the box you chose worth?

C: It's not worth anything.

U: It's worth nothing?

C: Well, you don't know. It could be worth \$100 or it could be worth nothing.

For contrast, here is a fragment from a post-conservational Gary:

U: How much is your box worth?

G: I guess it's worth \$33. I don't know if you could really sell it for that though, might only get about \$25.

And when I say to Chris:

U: Some people say that it is worth \$33.

C: Why would they say that?

U: Well, one time out of three it will be worth \$99, so it could be worth a third of \$99.

C: They [the people who say that] must be salespeople. I took a sales class and they try to get you to think that way. It's a trick you've gotta do if you're gonna call all those people.

I found this last remark quite interesting. In Chris's analysis, if the salespeople thought that most of their calls, which presumably did not result in sales, were worthless, they would get discouraged. If, however, they thought that, for example, 99 calls out of 100 wouldn't be sales, but 1 in 100 would result in a \$10,000 sale, then they could think of each call as being worth \$1000. By using this "trick", they've transformed themselves from being 99% failures to 100% successes!

There is one more important insight buried in Chris's fragment. What is it that makes Chris unwilling to adopt this trick of giving the box a "worth"? I think that Chris, like many of the other interviewees, is uneasy with taking a "long run" view of single instances. Even post-conservational interviewees found this point troubling. Consider Brian:

B: I definitely see that switching is the right strategy in the long run, but one thing still bothers me.

U: Yeah?

B: We're in this situation once, right?

U: I guess.

B: Well, if we're in it once, then what do probabilities have to do with it? I mean really, we're only doing this once, we bought our tickets and paid for our LA hotel, went to Disney world or land or whatever and now we're on the show. ...

Brian is explicitly asking the question that I believe lay underneath Chris's fragment and that many interviewees articulated or half-articulated - are probabilities relevant to our life situations? To events which are not obviously repeatable? In what situations do we say we are in a unique situation and when do we see our situation as being one of a class of similar situations?

As we saw in the previous chapter, this is at the heart of the argument between the subjectivists and the frequentists about the meaning of probabilities. But for Brian, this argument has real bearing on what to do in the Monty Hall situation. And for many interviewees, it is in the background, subtly discouraging them from using probability in life situations. We are all aware of numerous examples where people think they can "beat the odds". Smokers are notorious for saying that the statistics linking smoking to disease do not apply to them. When, then, does it make sense to live our life and make our decisions as if we were members of a statistical class, and when should we view ourselves and our situations as unique? Without confronting these basic issues of the meaning of randomness and probability, we are likely to have difficulties with applying probability.

When pre-conservation interviewees became switchers, like many converts they became zealous advocates for their cause. Taking advantage of this trend, I asked them to come up with the most compelling argument they could for switching and try to convince someone pre-conservational.

While there were many variants, two arguments were by far the most popular.

One of these went about its work by trying to generalize the Monty Hall problem to many boxes. So, for example, we are asked to suppose that there were 100 boxes in the original problem, we chose one, and Monty Hall opened 98 (all but one) of the rest of the boxes. Now we can choose between our box and the remaining box. It is now obvious, so say these “convincers”, that we should switch. Many of the convincers were quite taken with this argument and felt sure that it would convince anyone to switch. However, this argument rarely convinced anyone. Two sorts of objections were raised to it: Why did you generalize to opening 98 boxes, Monty opened one box and you should do the same whether you have 3 boxes or 100 boxes? If this hurdle got jumped, then most often the “convincee” restated their argument for why, even in this case, the odds were 50-50 so it didn’t matter. We might predict the failure of this argument as it does not address either the question of history and agency in the situation nor, directly, the notion of box worth.

A more successful argument commonly advanced directly addressed the question of agency in the situation. Typically, this “no-choice” argument started by drawing the following diagram:

No Choice Argument

You select the box on the left

Situation 1	car	Monty can show this goat	Monty can show this goat	Switching loses
Situation 2	goat	car	Monty must show this goat	Switching wins
Situation 3	goat	Monty must show this goat	car	Switching wins

The convincer goes on to say: So in situation 1, you chose box #1, the money is there, so you win if you stick, lose if you switch. But in situation 2, you chose #1, and the money is in box #2. Now since Monty must show a goat, he has *no choice* but to open box #3. So, you lose if you stick and win if you switch. The same holds for situation 3. So in one of three situations you lose if you switch, and in two out of three you win if you switch.

Note this argument has a lot of experience and sophistication built into it. The convincer simplified the problem by showing only three situations (all with you choosing box #1) instead of showing nine situations. This simplification comes from having worked out that the position of the chosen box is not important, an argument not usually

shared by the convincee. Similarly, it carefully avoids the trap of breaking down situation 1 into two cases, one in which Monty chooses door #2 and one in which Monty chooses door #3. It lumps the two together recognizing that the combined case is equally likely as the other two cases. All of this cleverness is hidden from the convincee, and quite often the simplicity of the argument is enough to convince him/her.

But the “no-choice” argument can be too clever for its own good. Interviewees who were convinced by this argument, and even many of those convincers who gave it, were likely to say that the following variant of the Monty Hall problem has the same resolution and responds to the same argument.

Suppose you’re in the original Monty Hall situation but now, instead of choosing a box to deliberately reveal a goat, Monty flips a coin to decide which of the two unchosen boxes he will reveal. He flips his coin and opens a box with a goat in it. Now, which is better sticking or switching?

Many holders of the “no choice” argument refer back to the diagram they have drawn or redraw it and “see” that it obtains equally well to the variant problem. Despite the name I have given to this argument, now that they are in a situation where Monty does have a choice (or at least the coin does), “no choicers”, seemingly so seduced by the simple diagram, forget that Monty’s having no choice (or “forced” choice) was crucial to the validity of their argument.

In the preceding paragraphs, we have seen how interviewees grappled with “solving” the Monty Hall problem. However, the primary use interviewees made of the Monty Hall problem was after they had switched viewpoints (post-conservational). Then the Monty Hall problem was often cited and referred to as a point of departure and connection in thinking about other probability situations.

Consider this interview fragment of Arne’s:

U: So now that you have changed your mind and believe that switching is better, how do you explain your previous belief? And how would you convince someone else to change?

A: OK, why was it so obvious? [that sticking is the same as switching] OK, because you look at it in a very narrow frame. You really have to step back in order to see the other variables, the other information in the system, which you don't usually see. Also because you have such a prototype, such a hardwired prototype for, not a hardwired, but a very deeply burned in prototype for this kind of situation. Look, you've got two closed doors, one of them, then there's a 50-50 chance. There's no reason to switch, I mean you can switch if you want, but it's not going to help you one way or the other to switch or not.

Note, that Arne spontaneously comes up with (then retreats from) an innatist explanation of the error.

A: These are the types of assumptions, we have to establish certain assumptions about this problem. Is it a logic problem, is it an psychological kind of thing? And I think we're kind of dealing with it, not as a psychological, we're talking about it, this is a formal, either a probability or a logical... I don't know if this falls between the line of a probability question or a logic question. OK, there was a "ah-HAH" moment in this whole process, I'm trying to reconstruct it so it may not go exactly the way it happened, but, there is, you have to make a certain assumption about what's going on here, there's a certain set of rules, there's one rule that has to follow through this thing to make any sense, and the rule is that, Monty Hall isn't going to open up the door with the Ferrari. That's the only rule, that Carol Reul, [sic]whoever, Monty Hall isn't going to open up the door with the Ferrari, that's the rule. And that's the key thing, that's why this is so confusing, cuz he's only going to show you the donkey, he's always going to show you the donkey, so there's one fixed thing in the system.

U: Now, was that obvious to you, even in phase one when you said that was obvious?

A: No, like I said you don't think about that. You don't think about that.

U: So if someone asked you if it was possible for Monty Hall to open the door and you would have seen the Ferrari, what would you have said?

A: I would have said that doesn't make sense, because then there would be no problem, that's not how this scenario, that's not how this thing is scripted to happen, so there's something built into the script which is that no matter what door you choose, there's at least one donkey in a door that you didn't choose that's going to be opened. OK, so then you step back from this and what you realize is that.. you have a.. right, it works out, ends up being a 2/3, 1/3 thing, not a 50-50 thing, because the action.. because You have a 2/3 chance of choosing a donkey. OK? And a 1/3 chance of choosing the car. If you choose the car, I mean if you choose the donkey, then Monty doesn't have any choice what he's going to do, if you choose the donkey, either one, Monty Hall has no

choice what he's going to do, which means that 2/3 of the time the door that he, of the two doors you haven't chosen, the door that he decides not to open is the car. So 2/3 of the time, you can get the car, 1/3 of the time you will get the donkey. So if you do this 100 times, you're going to get 66 cars, versus 33 donkeys, and the point is that it's not a single choice, not a single event probabilistic model, it's a.. actually I don't know if it's a 2- or 3- event probabilistic model, it's actually.. it's a 2-event probabilistic model with a kind of determined action in the middle. And that's why it's a complicated kind of thing. You need a trick in order to understand it.

U: Let me see if I've understood what you mean by a trick. Is the trick that you've framed it in terms of Monty's choice? If you pick a donkey, which is going to be 2/3 of the time, Monty Hall has no choice, he has to show you the other donkey, so if you switch to the unopened box it has to be the Ferrari and you win 2/3 of the time?

A: Right. That's it, that's the key piece.

A: Well, that..... yeah, the other side of that is he's giving you information, He's telling you a lot. The other thing is that you don't know anything, and he knows something. He knows where everything is, and you don't know where anything is at the beginning, and then at a midway point you have information, so it's not like pure, it's not a pure probability thing, there's like a certain kind of, there's information in the system, and I don't know information theory and all this kind of stuff, but, there's probably, this is not a totally probabilistic problem, it's gotta be something else, or maybe probability has ways-- I'd like to hear a kind of formal description of this problem because I'm just kind of hacking, I mean it works, hacking works, and it allows you to explain it to other people, but I'd really like to, I never was able to kind of come up on my own or hear from anyone a formal, in a kind of formalistic logical, or probabilistic equation for, or sequence of how this works, but that's what I'd like from you. But the key thing, the common kind of thing for people is to see that..... that 2/3 of the time, Monty Hall has no choice, and therefore 2/3 if you switch, you're going to get it right. Maybe there's something better..

Arne has discovered two “tricks” or new ways of seeing the problem. Framing the problem in terms of Monty’s choice puts history into the evaluation and removes it from the first “burned-in” prototype where all that matters is what’s staring at you now. Secondly, Arne now sees Monty’s agency in the situation. Monty has information about which box has the Ferrari and he uses that information to make his choice, namely he deliberately avoids showing you the Ferrari. Through his choice, Monty conveys some of the information he has to you (or information “flows” from Monty to you).

U: So what about that other thing that said "Hey, there's two boxes in front of me?"

A: But there's a residue of past information, you have information. See there's two boxes, that's assuming it's a complete black box, but it's not a complete black box anymore.

U: It's very interesting to me one of the things that you said... that that wasn't probability, it was information theory or that it was logic..

A: That was, it was something, it was a hybrid, it wasn't pure probability, it was a hybridized probability, that.. .but I don't know the mathematics of ...

U: What do you think probability is if it's not that? Probability is when you have two black boxes in front of you?

Here I'm getting back to the first interview question which we never completed, but at the same time taking advantage of the opportunity presented by this attempt to classify the Monty Hall problem.

A: Good question. Well, I guess in probability, obviously there's going to be certain things that are determined. Definitely constraints, and a certain determinateness about it and a certain indeterminateness about it. So maybe this could fall into pro..., but somehow I think there's too much of a, like the Prisoner's dilemma, I don't think of that as probability, purely, if there's gotta be something more, my definition of probability doesn't, isn't broad enough to cover that. If there's some other piece which is, which is not wet like psychology .

Here Arne for the first time seems to be reconceptualizing his understanding of probability from being essentially about “randomness” to a new view in which there’s “a certain determinateness about it and a certain indeterminateness about it”. This is a crucial change which will pervade Arne’s thinking. It is very common in our culture for adults to see randomness as completely unordered and unlikelike, or as each possibility equally likely so it is impossible to know anything. This, of course, is in basic conflict with the idea of distribution and this conflict “did in” the NSF software users mentioned in this chapter’s opening anecdote.

U: Wet? like alive?

A: Yeah, like amorphous, ambiguous, nonformalizable, like psychology, but it's.. there's something like in information theory is the thing that comes to me, but I don't know enough about that to.. But there's a decision theory kind of thing, there are two players, one person who knows about the system, and one who.. there's like a knowing, what is known and what is not known. So it seems like somehow in my understanding of probability, it's... there's some pieces that come from outside probability that get injected into some probabilistic model, that have to be accounted for, and maybe, and maybe probability is broader and encompasses that, and that's cool.

Here Arne seems to be distinguishing between contexts in which there is no deliberate agent acting and calling those probability contexts and those in which agents act knowingly which he is not sure if to call probability.

U: OK, let's get one example of something that definitely is probability.

A: OK, flipping coins, rolling dice. OK, you want something better?

U: No, that's a perfect example, let's look at those. What's going on when you flip a coin?

A: All right, we're assuming a kind of a, certain kind of universe right? Where it really is an absolute 50-50 percent chance it's going to happen one way or the other, y'know, whatever. So what's happening is it's a random, purely random event with a ... purely random event that over many, many passes exhibits certain behaviors. That's... If I was to sit down and come up with a definition of probability, it probably would be better than that, but...

U: So probability takes place when something is purely random, is that what you're saying?

A: All right. That's kind of my strongest prototype. If I was pushed, I would kind of say.... it's ... I mean..... well, I'm kind of thinking that it's locally random, that there's spaces of randomness within a kind of architecture of.. nonrandomness, I don't know if that makes sense, but, that's I'm having this visual kind of thing. Y'know, maybe you kind of have this certain kind of relatively nonrandom kind of.. certain kind of blurs of randomness at the fingers of it.. And these are just kind of things, Maybe probability is more, is a mixture of randomness and nonrandomness.

U: Let's look at the coin example. You think that in reality flipping the coin. What does that mean to be random?

A: That's OK. Well, I mean, what it means for it to be random is, well. Like do you want me to deal with physics, or just however?

U: Anyway you want.

A: Random, when I think about probability and randomness and coin-flipping, that's kind of a pure mathematical kind of randomness, like we're assuming it's 50%, but in the real world maybe one side of the coin would be a tiny bit heavier, or there'd be certain very very obsc... y'know, kind of chaotic prototype effect weirdness stuff that would be leading things towards one way or the other, but. What it means is that, for all practical purposes, it would seem obscure to subscribe anything nonrandom to such an event. I don't know if that makes any sense, but..

U: You mentioned physics. Well if you're a physicist, then how could there be anything less random than the flip of a coin? With certain initial conditions, the weight of the coin, the force on the coin when you flip it, then gravity just takes over, and the next thing, the spin of the coin and the outcome is completely determined by the initial conditions..

A: Here's an interesting thought; what if you had a machine that was so precisely calibrated that it could flip a coin and have it exactly the same each time? I mean, you could imagine flipping a plastic Tupperware top and getting it to land y'know, the same side every time.

Here Arne transforms my suggestion to look at the possible knowledge one can have about the coin into a way of controlling the coin through building a machine. Arne asked me for a transcript of his interview. Later he wrote me that this point stood out for him, that I was talking about knowledge of the world and he was talking about control. He attributed to this conversation the beginnings of a shift in his perspective from looking at probability as “what is under my control and what simply random or beyond control” to looking at it as “what do I know”. It is interesting to note that in my summary of Arne’s Monty Hall “tricks”, I also recast his first trick to be less about Monty’s control and more about the chooser’s knowledge. However, Arne’s second “trick” was explicitly acknowledging the value of knowledge or information in determining probabilities. This dialogue of perspectives, the focus on control vs. the focus on knowledge, in which Arne is engaged, is reminiscent of the argument between propensitists and Bayesians: are probabilities real properties of objects or are they properties of our knowledge of objects?

U: In fact magicians are quite capable of flipping a coin and getting it to land the way they want.

A: In order to have a certain trick work to make it seem like it could happen any number of ways. OK, that makes a lot of sense.

Yeah, no, I like it, it's very provocative. So, it's like y'know, the pure probability action, I'm not sure anymore that it exists in any situations that I can think of. Yeah, that makes sense. So we can talk about it as a mathematical kind of abstraction, y'know, theoretical notion. Construct, mathematical construct. So, under certain practical conditions in your life its very, it's quite a heuristic then to assume that something is random, cuz it really, for all intensive purposes it behaves randomly in your life. And like flipping a coin before a soccer game, and assuming it's not a machine and, certainly nobody's a magician, then basically it's a random thing.

U: How would you respond to someone who said that coin tosses are a completely determined thing, but as far as your knowledge is concerned, as far you're able to figure out, you don't have enough access to the, to prior information to be able to deduce from that prior information

A: Not only deduce, but control...

I still haven't caught on to our "struggle" between deduction or knowledge and control.

U: What would you deduce...

A: But both, I agree. Hmmm. I see what you're getting at. So that's a good point, though, getting back to your original, it's a question of what we can know... So there's moments of ... there's moments of, there's points of certainty and points of uncertainty. In that model, I mean in that model. Certain things that are... Then how does it make it that much different from the Monty Hall problem; you also ignored the prior information there, but after some reflection you can deduce it.? It's not obvious, but you can..... Yeah, so it's kind of like, that's very provocative, that's like the classic, it's really nice because it's a coin problem is this classic random, if there's randomness, it's in the coin.. But, in a sense, it's utterly deducible or whatever, derivable... so it depends on kinda what you know. Yeah, it's very contextual, it's very, it's like how you frame it. So let's go back to the Monty Hall problem, there's two ways of framing, there's multiple ways of framing it, but we've pretty much talked about two.

Summary

The Monty Hall problem is a particularly evocative puzzle. As we have seen, a strong analogy can be made between understanding "Monty Hall" and the development of volume conservation in children. Because, most people's intuitions about Monty Hall are wrong, we could classify it as another example of a Tversky & Kahneman type systematic error. But Monty Hall does the test questions of Tversky & Kahneman one better. Most Tversky & Kahneman subjects' "misjudgments" were not robust. When Tversky & Kahneman's subjects were told why their assessment of a conjunctive

probability as greater than its constituents was wrong, most recognized their previous likelihood assessment as a mistake. In contrast, most pre-Monty-conservationists maintain their errors despite great efforts at persuasion. If Tversky & Kahneman like errors, which are recognizable by the subjects as errors, are to be regarded as ineducable intuitions, resolved only through the use of formal methods, the case should be even stronger for Monty Hall errors which aren't even recognizable to the interviewees as errors?

But as we have seen in the above interviews, learners do build up intuitions about the Monty Hall problem. Monty Hall can be resolved through building intuitions as opposed to just using formal methods. In fact, for all interviewees in this study, *it could only be resolved in that way*. But this resolution often takes substantial time. Simply strengthening the correct argument and repudiating the incorrect argument will not suffice. Conflicting intuitions need to be elaborated, explored and strengthened. New concepts such as the concept of probabilistic box worth (or expected value) must get constructed. Once these concepts are constructed, the learner has not just resolved the Monty Hall problem, but has built new tools and examples for thinking about probability which are applied and linked to new probability situations. In the process, the epistemological status of probabilities gets negotiated and elaborated, often emerging from its repressed background position into the light.

Many important epistemological issues were raised in the Monty Hall interviews. Interviewees such as Brian began to negotiate the projectibility of unique situations - to what extent can they be viewed as members of classes? And interviewees such as Arne grappled with the extent to which probabilities reflect our knowledge and can be characterized by flows of information. These ideas led Arne to a more robust sense of the meaning of probability distribution and the relation between a probability and a distribution.

In addition, we have seen interviewees noticing the importance of the selection process, the distribution, in determining probabilities. In the original Monty Hall problem, the selection process is tied to Monty Hall's agency, but if we change the process to a coin flip, as in the variant problem, the probabilities also change.

Finally, we have seen how the construction of the probabilistic notion of "box worth" was of fundamental importance in overcoming the strong "agent" which asserts that two possibilities "means" a 50-50 chance.

Circle Chords

From a given circle, choose a random chord. What's the probability that the chord is longer than a radius?

This question was, in one sense, the most formally presented question in the interview. On the surface, it most mirrored the kinds of questions that students get in school. But, because of its hidden ambiguity, and the relative mathematical sophistication of the interviewees to whom it was posed, many rich interviews arose from it. It was particularly rich in evoking epistemological questioning and investigations into the meaning of the word random and how "randomness" connects to both mathematical distribution functions and the physical world.

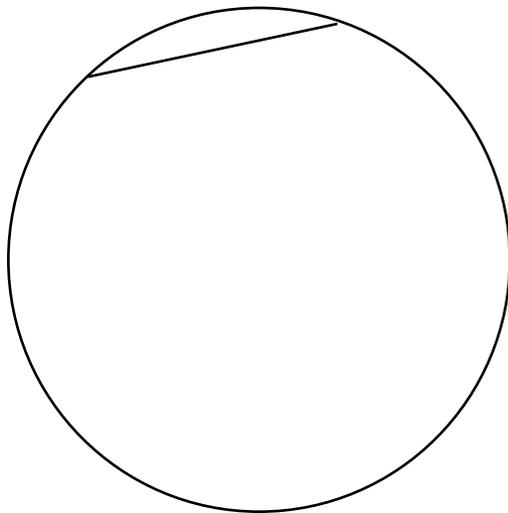
Like Amy and Doug in the divorce interview, Ellie gets into trouble understanding the meaning of random. Again, we could have resolved her difficulty by specifying a particular distribution of chords, or by describing a specific experiment to generate the chords. But if we had done that, Ellie would not have developed the insights she does about the meaning of randomness. It is difficult, sometimes, for us as teachers to watch learners struggle with foundational questions and not intervene. However, the

temptation to intervene is more easily resisted if we keep in mind that it is only by negotiating the meaning of the fundamental concepts, by following unproductive, semi-productive and multiple paths to this meaning that learners can make these concepts concrete.

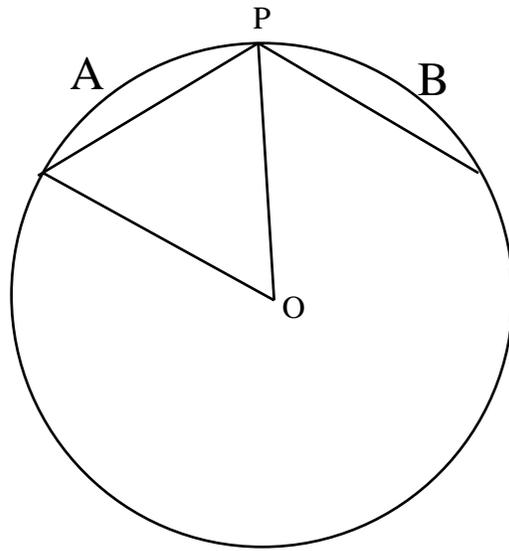
Many interviewees answered this question fairly quickly using the following argument. Chords range in size from 0 to $2r$. Since we're picking chords at random, they're just as likely to be shorter than " r " as they are to be longer than " r ". Hence the probability is equal to $\frac{1}{2}$.

Ellie, a computer professional who had a solid undergraduate math background got engaged with this question but approached it differently.

She began thinking about the problem by drawing a circle and a chord on it which she said had length equal to the circle's radius.



After contemplating this drawing for a while she then drew the figure below:



With the drawing of this picture came an insight and she pointed at the triangle in the figure and said:

*Ellie: It has to be equilateral because all the sides are equal to a radius. So that means six of them fit around a circle. That's right $6 * 60 = 360$ degrees. So, that means if you pick a point on a circle and label it P, then to get a chord that's smaller than a radius you have to pick the second point on either this section of the circle (labeled A in the figure below) or this one (labeled B in the figure below). So since each of those are a sixth of the circle, you get a one third chance of getting a chord smaller than a radius and a two thirds chance of a chord larger than a radius.*

Ellie was quite satisfied with this answer and I believe would not have pursued the question anymore if not for my prodding.

U: I have another way of looking at this problem that gives a different answer.

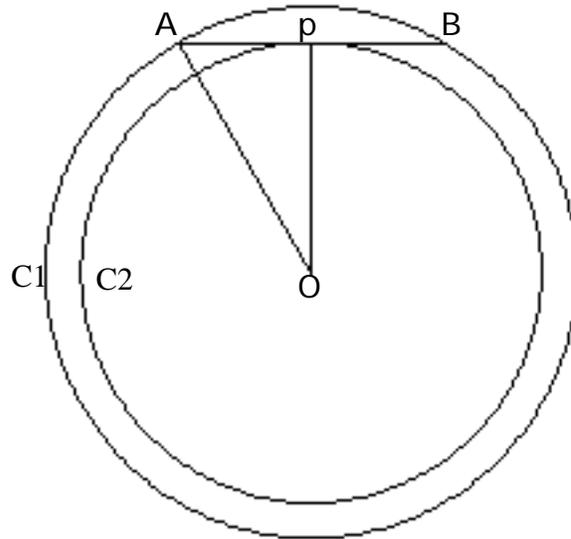
E: Really? I don't see how that could be.

U: Can I show you?

E: Sure. But I bet it's got a mistake in it and you're trying to trick me.

U: OK. Let me show you and you tell me.

I then drew the figure below:



U: Consider a circle of radius r . Draw a chord AB of length r . Then drop a perpendicular onto AB from the center of the circle, O , intersecting AB in a point, P . Then P is a mid-point of AB . Now we calculate the length of OP . We have $OA = r$ and

*$AP = r/2$. By Pythagoras we have $OP = \frac{\sqrt{3}}{2} * r$. Now draw a circle, $C2$, of radius OP centered at O . If we pick any point on $C2$ and draw a tangent to the circle at that point then the resultant chord has length r . If we pick a point, P' , inside the $C2$ and draw the chord which has P' as mid-point then that chord must be longer than r . Similarly, if we pick a point inside $C1$ but outside $C2$ and draw the chord which has that point as mid-point, then that chord must be shorter than r .*

Now pick any point, Q , inside $C1$. Draw a chord, EF , of the circle which has Q as mid-point. EF will be bigger than a radius if and only if Q is inside $C2$. It follows that the probability of choosing a chord larger than a radius is the ratio of the areas of $C1$ and $C2$.

$$\text{The area of } C1 = \pi r^2. \text{ The area of } C2 = \pi P^2 = \frac{3}{4} \pi r^2$$

So the ratio of their areas is $\frac{3}{4}$ and therefore the probability of a chord being larger than a radius is also $\frac{3}{4}$, not $\frac{2}{3}$ as you said.

This explanation had a disquieting effect on Ellie. She went over it many many times but was not able to find a “bug” in the argument. After many attempts to resolve the conflict she let out her frustration:

E: I don't get it. One of these arguments must be wrong! The probability of choosing a random chord bigger than a radius is either $\frac{2}{3}$ or $\frac{3}{4}$. It can't be both. I'm still pretty sure that it's really $\frac{2}{3}$ but I can't find a hole in the other argument.

U: Can both of the arguments be right?

E: No. of course not.

U: Why not?

E: It's obvious! Call the probability of choosing a chord larger than a radius p . Then argument #1 says $p = \frac{2}{3}$ and argument #2 says $p = \frac{3}{4}$. If both argument #1 and #2 are correct then $\frac{2}{3} = \frac{3}{4}$ which is absurd.

Here Ellie is quite sure that there is a definite and unique meaning to the concept “probability of choosing a random chord larger than a radius” even if she admits that she is not completely certain what that meaning is.

U: Would writing a computer program help to resolve this dilemma?

E: Good idea. I can program up a simulation of this experiment and compute which value for the probability is correct! I should have thought of that earlier.

Ellie then spent some time writing a *Logo program. As she worked to code this up, she already began to be uneasy with her formulation. A few times she protested: “But I have to generate the chords somehow. Which of the two methods shall I use to generate them?” Nevertheless, she continued writing her program. At various points, she was unsure how to model the situation. She experimented with a fixed radius as well as giving each turtle its own radius. She experimented with calculating the statistics over all trials of each turtle as opposed to calculating it over all the trials of all the turtles. Finally she decided both were interesting and printed out the “probability” over all trials as well as the minimum and maximum probability of any turtle.

Below are the main procedures of Ellie's program.

```
;;; this turtle procedure sets up the turtles
to setup
setxy 0 0   ;;; note that all turtles start at the origin
make "radius 10
make "p1x 0
make "p1y 0
make "p2x 0
make "p2y 0
make "chord-length 0
make "trials 0
make "big 0
make "prob 0
end
```

```
;;; This is a turtle procedure which generates a random chord.
to gen-random-chord
fd :radius
make "p1x xpos
make "p1y ypos
bk :radius
rt random 360
fd :radius
make "chord-length distance :p1x :p1y   ;;; the distance primitive calculates the distance
                                         ;;;between the turtle's location and another point

bk :radius
end
```

```
;;; this turtle-procedure gets executed at each tick of the clock
to turtle-demon
gen-random-chord
make "trials :trials + 1
if bigger? [make "big :big + 1]
make "prob :big / :trials
end
```

```
;;; is the turtle's chord bigger than a radius?
to bigger?
:chord-length > :radius
end
```

```
;;; the observer summarizes the results of all the turtles
to observer-demon
make "total-trials turtle-sum [:trials]   ;;; the total number of trials
make "total-big turtle-sum [:big]         ;;; the total number of chords bigger than a radius
make "total-prob :total-big / :total-trials
   ;;; print the summary results every 10 steps
   ;;; also print the results for the turtles with the highest and lowest probabilities
every 10 [type :total-big type :total-trials print :total-prob type turtle-min [prob]
print turtle-max [prob]]
end
```

Ellie ran her program and it indeed confirmed her original analysis. On $\frac{2}{3}$ of the total trials the chord was larger than a radius. For a while she worried about the fact that her extreme turtles had probabilities quite far away from $\frac{2}{3}$ but eventually convinced herself that this was OK and that it was the average turtle “that mattered”.

But Ellie was still bothered by the way the chords were generated.

E: OK, so we got $\frac{2}{3}$ as we should have. But what's bothering me is that if I generate the chords using the idea you had then I'll get probably get $\frac{3}{4}$.⁵⁹ Which is the real way to generate random chords? (underline added)

Having to explicitly program the generation of the chords precipitated an epistemological shift. The focus was no longer on determining the probability but was moved to finding the “true” way to generate random chords. This takes Ellie immediately into an investigation of what “random” means. At this stage, as she was before for the probability, Ellie is still convinced that there can be only one set of random chords and the problem is to discover which set these are.

U: That's an interesting question.

⁵⁹ Ellie did go on and actually write the code to do this experiment just as a check of her insight. However she encountered difficulties because she could not generate a random real number less than a radius. She wrote the code uneasily with an integer distance and the result was much larger than $\frac{3}{4}$. She quickly saw the “bug” though and through varying the radius of the circle was able to convince herself that she would get $\frac{3}{4}$. Her new code is the same as the old code except for a rewrite of the procedure gen-random-chord:

```
to gen-random-chord2
make "dist random (:radius + 1) ;;; let turtle move integer distance from center to circumference
make "ang random 360
seth :ang
fd :dist
lt 90
make "len sqrt ((:radius * :radius) - (:dist * :dist))
fd :len
make "p1x xpos
make "p1y ypos
bk 2 * :len
make "clength distance :p1x :p1y
fd :len lt 90 fd :dist seth 0
end
```

E: Oh, I see. We have two methods for generating random chords - what we have to do is figure out which produces really random chords and which produces non-random chords. Only one of these would produce really random chords and that's the one that would work in the real world.

U: The real world? Do you mean you could perform a physical experiment?

E: Yes. I suppose I could.Say we have a circle drawn on the floor and I throw a stick on it and it lands on the circle. Then the stick makes a chord on the circle. We can throw sticks and see how many times we get a chord larger than a radius.

U: And what do you expect the answer to be in the physical experiment?

E: Egads. (very excitedly)We have the same problem in the real world!!! We could instead do the experiment by letting a pin drop on the circle and wherever the pin dropped we could draw a chord with the pin as midpoint. Depending on which experiment we try we will get either answer #1⁶⁰or #2. Whoa this is crazy. So which is a random chord? Both correspond to reality?.....

This was a breakthrough moment for Ellie, but she was not done yet. Though her insight above suggests that both answers are physically realizable, Ellie was still worried on the “mathematics side” that one of the methods for generating chords might be “missing some chords” or “counting chords twice”. Ellie needed to connect her insight about the physical experiment to her knowledge about randomness and distribution. She spent quite a bit of time looking over the two methods for generating chords to see if they were counting “all the chords once and only once”. She determined that in her method, once she fixed a point P, there was a one-to-one correspondence between the points on the circle and the chords having P as an end-point. She concluded therefore that there “are as many chords passing through P as there are points in the circle”. But there will be more chords of a large size than chords of a small size. As could be seen from her original argument, there will be twice as many chords of length between r and $2 * r$ as there are of chords of length between 0 and r . Now for the first time, Ellie advanced the argument that most of the interviewees had given first.

⁶⁰ I chose not to intervene at this juncture and point out that the first experiment Ellie proposed did not correspond exactly to her first analysis and method of generating chords.

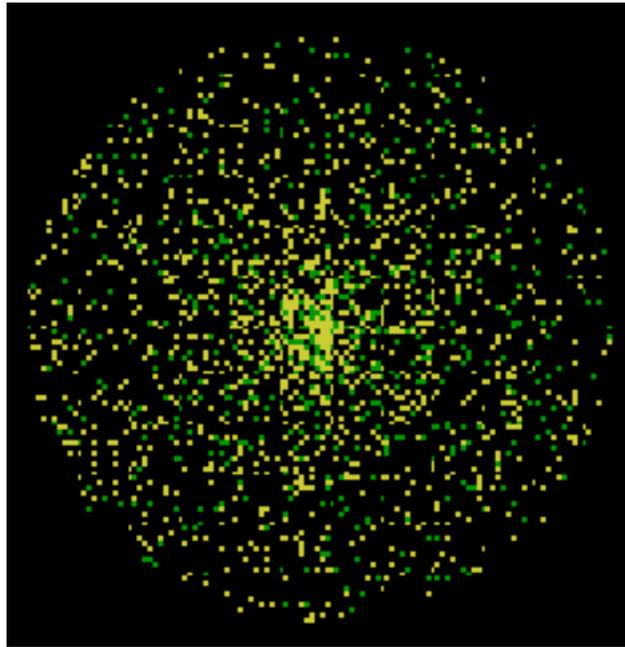
E: I never thought of the obvious. I've been sort of assuming all along that every chord of a given size is equally likely. But if that were true then I could have solved this problem simply. Each chord would have an equal chance of being of length between 0 and the diameter. So half the chords would be bigger than a radius and half smaller.

Ellie went on to see that, in the argument I advanced, for every chord of a given size (or more accurately a small size interval) there was a thin disk of points that would generate chords of that size by method 2. Choose a circle of radius r and an interval, “ a ”, small relative to r . Let's say “ r ” is large and “ a ” is 2. Then an interval of chord size of size “ a ” will correspond to an area of size P when the interval is near chord sizes of size 0. But an interval of size “ a ” will correspond to an area of size $P * (2r - 1)$ when the interval is near chord sizes of size $2r$. Thus large chords become increasingly more probable.

One other interesting feature of note: The program that Ellie wrote placed all the turtles at the origin and since Ellie, as a professional programmer, wrote state transparent code they stayed at the origin. Initially, she had placed the turtles at the origin because she recognized a potential bug in her program if she created the turtles randomly as is typical in *Logo - namely that turtles might “wrap”⁶¹ the screen when drawing their circles and thus produce incorrect calculations for their chord lengths. But, because the turtles remained centered at the origin, the program was not very visually appealing. While we were engaged in the interview, a student came by and watched. He asked us why nothing was happening on the screen. Ellie explained what she was investigating and then had an idea of how to make the program more interesting. She would spread the turtles out a bit so they each could be seen tracing their circles and have the turtles turn yellow if their chords were longer than radii and green if they were shorter. To spread the

⁶¹ In a typical Logo or *Logo screen, when a turtle goes off the screen to the right or at the top, it reappears at the left or bottom.

turtles out without getting too close to the screen edge, Ellie executed the code “fet [fd random (60 - radius)]” telling each turtle to move a random amount out from the origin. But in doing this, the result wasn’t quite what Ellie had hoped for. Near the origin there was a splotch of color [mostly yellow] as all the turtles were squeezed together very tightly, while near the edges the turtles were spaced out more sparsely (as in the figure below).



What had happened here quite by accident was a mirroring of the original dilemma. Ellie had used a linear random function to move points into a circular planar area. There were an equal number of turtles in each equally thick disk around the origin, but the outer disks had greater area than the inner disks and therefore appeared less crowded.

So Ellie’s function which worked to spread turtles out evenly (and what she then called randomly) along a line did not work to spread them out evenly on the screen. This experience was an important component of her subsequent “aha” moment - exposing her as it did to a crack in her solid and fixed notion of “random”.

Envelope Paradox

There are 2 envelopes in front of you. One contains x dollars and the other contains $2x$ dollars. You choose one, open it and get \$100. You are then offered a switch if you want it. You can take the other envelope instead.

Here are 2 arguments:

1) There were 2 envelopes, one contained x dollars and the other $2x$. You got either x or $2x$. Each was equally likely, so switching doesn't help or hurt.

(Call this the neutrality argument or “NA”).

2) The other envelope contains either \$200 or \$50. These are equally likely. So, the expected value of the other envelope is $\frac{\$200 + \$50}{2} = \$125$. You should switch.

(Call this one the switching argument or “SA”).

Which if any of these arguments do you believe? What is the bug in the other argument?

This paradox, of all the puzzles that came up in the second group of interview questions, provoked the most sustained interest. Since many of the interviews on this puzzle were conducted over email, many were in process concurrently. The volume of response was much greater than I expected with some respondents reporting many tens of hours thinking about it. Most of these had already worked through the Monty Hall problem either earlier in the interview or on a previous occasion. Eventually, unable to respond personally to each new insight, I created an email discussion list where people could discuss this problem. Through natural spread, I sometimes got responses to this problem from people not originally interviewed.

Some of the interviewees did not get very engaged with this question. In most cases this was because NA seemed clearly correct to them and SA seemed very dubious. As we have seen previously (in both Chapter IV and in the Monty Hall problem), in order for a paradox to take hold and engage the learner, both conflicting agents must be considered reliable and valued friends. NA embodies the most reliable of probability agents - the principle of insufficient reason - if you see no difference then there is no difference. SA involves ideas of expectation and utility, agents that are not particularly familiar or friendly with many of the interviewees. It is no surprise, then, that most interviewees initially favored NA.

Many of the interviewees found themselves quite “stuck” on this question. Some never got unstuck and eventually gave up. Others persisted and found a way to “resolve” the question to their satisfaction. Still others got deeply engaged with the paradox and found many related questions that they wanted to pursue. These latter interviewees often did not stop till they had “resolved” and connected many different variations and related questions.

In most cases, interviewees had a clear intuition right away as to which argument was true even if they had no argument to support it. Usually this intuition strongly constrained their approach to the paradox - they started by trying to prove that their intuition was correct.

In the remainder of this chapter, we will examine text fragments of four interviews about the envelope paradox.

Dale

Dale was typical of the interviewees in that he had a clear intuition which argument was right but was atypical in that he started by believing SA.

This is Dale’s first email message to me in response to the paradox:

The first argument doesn't seem very convincing; I guess it seems odd to make the argument that each was equally likely, therefore, you should pursue some course of action in the future.

On the other hand, here's a reformulation that seems to support the second argument -- (maybe it's basically an informal version of it):

By switching, you stand to either lose \$50 or gain \$100, so the upside looks better than the down side, and they're equally likely, so, go for the switch.

At this point, there is an interlude in which Dale expresses concerns about the concept of expected value and whether it can be applied in a single situation such as this.

Here's what's still bothering me: what do we mean by a good strategy? according to my understanding of these things, if you do the best "expected value" thing, then you win over the long term. But what does that say about a single instance? Think of it this way: suppose you really need the money, it's like a life or death situation, you're homeless or something. Maybe in that case it would be a bad idea to risk losing \$50 even if the upside is very good. So, for example, you obviously are losing in the "expected value" sense when you buy insurance, but most people think it's a good idea to buy it anyway, because the downside (of whatever situation) is so undesirable.

Dale is again bringing up the issue of relevance of probability to real life situations -- situations which you face only once. But Dale brushes aside these concerns implying that they are not really mathematical and therefore not appropriate to this context. He goes on to assert his original intuition:

All right, so let's forget about my problems with "expected value." Let's pretend that that problem doesn't exist, and take the view, as seems to be expected, that one argument is "right," and the other is buggy. I think that viewed that way, 2 is right, 1 is buggy. Now I'm going to try to prove it.

Unfortunately, in this case I did not notice this dismissal of concerns perceived as extra-mathematical and therefore ignored the opportunity to open up discussion of these important connections.

Dale goes on in his attempt to prove SA. He starts by marshaling an argument against
NA:

I think that arg 1 [NA]has a bug. I could say something like, " it doesn't make sense to argue about what was equally likely or whatever, when you're given the next choice, it's a new situation, and you have to deal with it. But maybe I can construct another example that makes this even more obvious.

Dale had spent some time thinking about the Monty Hall problem and like many of the interviewees initially classified the envelope paradox as a Monty Hall like situation. The above paragraph seems to drive a wedge between “envelopes” and “Monty” in that in Monty previous state was important (the way Monty eliminated one box was crucial for the resultant probability), but Dale asserts that only present state is relevant in this problem. However, looking more closely, Monty seems to be deeply influencing Dale here. On a superficial level, Dale knew that in Monty it paid to switch and he suspected that the same would be true in envelopes. But, also, Dale had learned from Monty to distrust arguments based on the principle of insufficient reason. The argument of “equiprobable if can’t find a reason not” is immediately overshadowed by any argument that could provide a reason for favoring one choice over another.

One more similarity that seemed to be important in Dale’s thinking is that in both “Monty” and “envelopes” it might appear that no new information has flowed to the system - in Monty the present situation of apparent equiprobability at first seems uninformed by its history and in envelopes the original equiprobable situation seems to have gotten no new information by opening an envelope. Yet, as Dale had learned through Monty, there was more information in the system than met the naive eye, and Dale suspected that opening the envelope created new information in envelopes as well.

In the language of agents developed in Chapter IV, NA is seen to be the agent of the principle of insufficient reason. This agent has been weakened through Monty and is now less reliable for Dale. Therefore, unlike most interviewees, Dale leans strongly towards SA.

Once he is convinced of SA, he goes about transforming the problem into one where SA is even more persuasive. To do this he focuses on the factor of 2 in the original formulation. He names the factor "a" and asks what if "a" were 100 instead of 2.

Back to the main line: is there a way of showing that I is buggy, yes!

Let's use $a = 100$. That means that the other envelope contains either \$1 or \$10,000. Now, who wouldn't risk losing \$99 bucks for a 50-50 chance of winning \$10,000? Well, maybe some people wouldn't but I believe arg. 2.

I guess that's all for now.

PS, one additional thought, maybe the "trick" in arg 1 is a confusion or "pun" on the factor of 2, confusing it with the fact that there are two choices -- in fact, they're independent.

Dale has transformed the problem into one where NA is irrelevant. The origin of the opened envelope is ignored and instead the problem is made equivalent to having \$100 and having the option of trading it for a 50-50 chance of \$1 or \$1000. This last problem admits a clear solution and the issue is resolved for Dale for the time being.

Barry

Barry is a research physicist and computer scientist. He has had considerable schooling in probability and statistics and works daily with these mathematical areas in his research.

Barry, right away, leans towards NA. However, he is not totally dismissive of SA and feels that there is some paradox in the situation that must be resolved. Like Dale, influenced by his belief in NA, Barry transforms the problem.:

I was thinking about it over the weekend, and here's a succinct way of stating the resolution to the paradox:

Let envelopes A and B contain a positive amount of money. With probability 50%, either may contain the smaller amount or the larger amount. The smaller amount is uniformly distributed between minval and maxval. The larger amount

is exactly twice the smaller amount, and is then uniformly distributed between $2*\text{minval}$ and $2*\text{maxval}$. Denote the amount in envelope A by X_A , and that in envelope B by X_B .

Then:

$$\langle X_A \rangle = \langle X_B \rangle = (\text{minval} + \text{maxval})/2,$$

(By the “ $\langle \rangle$ ” notation, Barry is denoting expected values)

so

$$\langle X_A \rangle / \langle X_B \rangle = \langle X_B \rangle / \langle X_A \rangle = 1.0.$$

Nevertheless, it is also true that:

$$\langle X_A / X_B \rangle = \langle X_B / X_A \rangle = 1.25.$$

(Why is the expected value of the ratio not equal to the ratio of the expected values? Because the quantities involved are strongly correlated. Expectation values of ratios are only equal to ratios of expected values if the quantities involved are uncorrelated. That is certainly not the case here.)

Thus, suppose that you choose envelope A, open it, and find that it contains \$100. It is true that $\langle X_B / X_A \rangle = 1.25$. This does not, however, mean that you should switch. You should switch if and only if $\langle X_B \rangle / \langle X_A \rangle > 1$, and this is **not** true.

To test this, the following Mathematica routine fills the envelopes:

```
fillenvelopes[minval_,maxval_,type_:Integer] :=
  Block[{lil,big},
    lil = Random[type,{minval,maxval}];
    big = 2*lil;
    If[Random[] > 0.5,
      {lil,big},
      {big,lil}]
```

while the following one does a little Monte Carlo run:

```
trial[ntrials_,minval_:1,maxval_:500,type_:Integer] :=
  Block[{ratsuma=0,ratsumb=0,expsuma=0,expsumb=0,env},
    Do[env = fillenvelopes[minval,maxval,type];
      expsuma = expsuma + N[env[[2]]];
      expsumb = expsumb + N[env[[1]]];
      ratsuma = ratsuma + N[env[[2]]/env[[1]]];
      ratsumb = ratsumb + N[env[[1]]/env[[2]]],
      {ntrials}];
  expsuma = expsuma/ntrials;
  expsumb = expsumb/ntrials;
  ratsuma = ratsuma/ntrials;
  ratsumb = ratsumb/ntrials;
```

```
Print["<X_A>/<X_B> = ",N[expsuma/expsumb]];
Print["<X_B>/<X_A> = ",N[expsumb/expsuma]];
Print["<X_A / X_B> = ",N[ratsuma]];
Print["<X_B / X_A> = ",N[ratsumb]]]
```

The results for 10000 envelopes are as follows:

```
bmb% math
Mathematica 2.0 for SPARC
Copyright 1988-91 Wolfram Research, Inc.
-- X11 windows graphics initialized --
```

```
In[1]:= << "envelopes.m"
```

```
In[2]:= trial[10000]
<X_A>/<X_B> = 1.00301
<X_B>/<X_A> = 0.996998
<X_A / X_B> = 1.2464
<X_B / X_A> = 1.2536
```

```
In[3]:= Quit
93.5u 2.7s 3:03 52% 0+2192k 7+0io 325pf+0w
```

confirming the above argument.

Barry has found a feature of this situation that corresponds to a familiar and well loved “agent” - that probability situations often involve noticing the “bug” that products and ratios of expected values do not necessarily equal the expected values of the products and ratios. This is in stark contrast to sums and products of expected values which do behave linearly.

I had the feeling when reading his email that Barry was very familiar with this particular agent and was happy to put it to work in this situation without much reflection on whether it applied. (In the context of “bugs” with expected value, the product bug agent is dominant for Barry.) In this way, the problem is classified according to familiar categories much like a perception might be.

Let me stress right here that by describing Barry’s thought process in this way I do not mean to put it down or devalue it. To the contrary, the move I describe here is the norm for how we all first respond to new and unfamiliar mathematical territory. To find a

connection or analogy between the new situation and a familiar agent is the only strategy that could possibly work for navigating new mathematical terrain. The problem here that leads to a “bug” is that one agent being so trusted and familiar takes over the connecting process and controls the whole situation leaving other agents, that might usefully be brought to bear, disconnected and out in the cold.

Later, I asked Barry to reflect on why it was that he transformed the problem into one which did not involve opening the envelope and getting a particular value. At first he did not see that he had transformed the problem, but when he suddenly did see it he remembered that he had explicitly transformed it but never considered that the change could make a difference since he was seeing the problem through the lens of NA. Barry reconstructed his thought process as saying that before opening the envelopes the expectations of the two envelopes are the same, so for Barry the puzzle is: why then does SA sound plausible? He then remembered the familiar “bug” and jumped to seeing it as the explanation for the plausibility of SA. I remarked that I always find it amazing how each person brings their own familiar friends to bear on new problems. Barry then said:

Well I guess I have to admit then that this idea (of the non-multiplicativity of expectations) is one that I am very familiar with. It's near and dear to my heart. For the past few years, I have been working on lattice gas simulations and the concept of connected correlation functions. Connected correlation functions are defined as the expectation of products minus the product of expectations. It turns out that if you make the mistake that these cancel out to zero then you get persistent errors in your simulations. So, you're right, this idea or tool was right at hand for me. I guess as they say, "if all you have is a hammer then all the world looks like a nail".

Before we leave Barry one final point. Barry ended his first email message by producing the Mathematica (Wolfram, 1990) code included above. Note here that the computer served in this instance to confirm an already strongly held belief. The code was written to describe the transformed problem and therefore did not surprise in its results. We might expect that, like Ellie in the previous interview, doing a computer simulation would provide insight for Barry because it would require him to write an explicit generating

function for the money in the envelopes. This expectation is not confirmed, the potential insight obscured by Barry's reformulation. When I asked Barry about this later, with his characteristic humor, he said he was following the rule "when in doubt simulate". But as we see here, there can sometimes be a danger in the too easy use of computer simulations as they can confer authority on incomplete arguments.

So far we have seen two attempts at grappling with the envelope paradox. While Dale and Barry reach opposite conclusions, they both have a strong initial hunch as to what the resolution should be and then go on to "prove" their hunches and reach resolve. However, some interviewees do not reach resolve. One response to the envelope paradox that often leads to giving up on it is a "wild oscillation" between NA and SA. The oscillation is experienced as very uncomfortable and therefore often leads to the interviewee's stopping to work on the question so as to avoid the discomfort.

Mark

Mark is a research mathematician at a corporate "think tank". As such, he is deeply enculturated in the mathematical culture described in Chapters II and IV. However, Mark has not thought much about probability since taking a university course many years before. Mark began writing a response to the question before finishing reading it:

I could have x and the other has \$200, or I could have $2x$, and the other has \$50. Normally, I would suspend thinking until I finished reading, but since this isn't just a puzzle, but a "first thoughts" type puzzle, I want to do as well as I can as early as I can. I expect a paradox coming about the chance to gain \$100 vs. the chance to lose \$50. Usually, the bet is double or nothing, 50-50 whether I win or lose \$100. That's even. This is better, but I'm having trouble believing that I have changed something by opening one box.

At this point Mark seems to believe NA. But Mark is not sure of his allegiance to NA. As he begins to doubt his belief, he starts to evaluate the process of his thinking and compares

his ambivalent conflictual experience it to his wife Karen's process which is described much more like the univalent responses of Dale and Barry:

How would Karen tackle this? She thinks completely differently than me. Maybe she would just jump to the right conclusion by some kind of intuition. (Not stopping to answer that, but brainstorming away.) What is the answer that I truly believe? (Can't be good to switch, but it was good to switch in the goats and car puzzle, so maybe it is. I dunno.) Any more strange ways of tackling this? No. Go back for a full picture with these wider perspectives in mind.

And now Mark seems to lean towards SA.

The first argument as you worded it seems too glib, and thus suspect. It uses arguments about previous state which must be where the confusion lies. I should ignore previous state and think only about where I am now.

Here I go and ignore my advice. What is different about having some money to lose and playing the game Vs not having the money and playing the game? Yes, that must be where the fallacy lies. I am being asked to play the same game as before with the same odds, but I must be making some mistake about what I have.

Here the revelation of his thoughts about the puzzle brings up shame for Mark. He is very faithfully (in a much more detailed and self-revelatory way than most) recording his thoughts in thinking about the paradox, but the unfamiliarity of the effort is taxing and the exposure of his "naked" thoughts is acutely embarrassing - an embarrassment which results in anger towards himself and to the interviewer:

There is something weird about justifying my argument steps to you, which is slowing down my thinking. If I didn't have to pop out and say what I just thought I wouldn't lose track of my arguments so much and would follow some of these trains of thought through. Excuses.

I hate you, Uri. I want to stop and go Xmas shopping, but I don't trust myself to record my thoughts, so I am trapped here trying to solve this at the terminal. I am sure this is easy, so I am embarrassed. But then the arguments about the goats lasted a long time, so this is exactly the kind of puzzle that good people get mixed up on, and I'll just have to admit that I got mixed up on it.

Despite my many reassurances to Mark that “envelopes” is hard, that others including myself have also struggled with it, that I am very interested in and not judgmental of his partial understandings, that I do not have an expectation that he will “solve” the paradox, Mark is angry and embarrassed at having his messy thought process exposed. The cultural taboo against revealing our incomplete understandings leave powerful inhibitions in learners.

In the next excerpt, Mark, again, reflects on his own thinking process, acknowledging inner conflict and comparing it to Karen’s freedom from conflict.

One source of paradox is the tendency to want $1.5x$ to equal y , but it doesn't, but I don't see the right things to equate in order to quiet my mind.

Karen wouldn't worry about this. She would believe in one and refuse to consider the other. I wouldn't be able to interest her in the paradox.

Note that Mark’s comparison of himself to Karen is a negative judgment. Karen would not have (or admit) confusion and conflict in thinking about “envelopes” and neither, Mark implies, should he. From a post-conservation perspective this seems strange. From this perspective, the conflict Mark is undergoing leads towards more complete understanding. Should a child, almost conservational, beginning to experience conflict between TALLER and WIDER, feel inferior because she has not the certainty of a child who “knows” the taller glass has more liquid?

Mark switches back again to NA.

I believe that there can't be any value in switching because I don't even have to open the box to be in this paradox, so there has to be a bug in the calculation in the second argument.

....

But the resolution of the bug has to be that the value of the other envelope is also \$100. How?

Here Mark initiates a new approach - getting out of the tacit frame of the paradox and making light of it. This approach is done with creativity and humor, but it sometimes recurs in the context of giving up on the paradox -- it's creating too much discomfort.

A cutesy argument for not switching is that you could be lying and there might be nothing in the other envelope, or even a court summons.

Another cutesy argument is that I need exactly \$100 to pay a bill, and I cannot take any risk that I will end up with less. I should have said "more than \$50" instead of "exactly \$100."

Another cutesy argument is that the first argument says that I won't change my net worth, and the second argument says that it will improve, and I'm stuck unable to resolve this, so I ought to switch just in case I'm wrong.

But, the discomfort here isn't just that Mark is uncomfortable with so much inner conflict. Here, he is also, like other interviewees we have seen, negotiating with the background question of why we should make choices according to probabilities. As a mathematician, Mark "knows" that he should not be concerned with this question and should pay attention only to the formal context. Nevertheless he is concerned with it and so raises it playfully. As in Dale's interview, I missed this aspect of Mark's thinking. I therefore failed to explicitly address his concern and, unfortunately, implicitly supported the view that questions of this sort are not important in a mathematical context.

But Mark doesn't give up yet. Here he transforms the situation trying to get new insight:

Say the puzzle was this: There are two envelopes. One is transparent and shows \$100 in it. The other is opaque and contains something. What does it contain?

case A: The transparent envelope contains $x = \$100$ and the opaque, $2x = \$200$.

case B: The transparent envelope contains $2x = \$100$ and the opaque. $x = \$50$.

I wanted to write two other cases, but that's it.

Now the puzzle is which envelope should you choose?

First, what is the expected value of the whole game? No, too early to tell.

The expected value of choosing the transparent envelope is \$100. The expected value of choosing the opaque envelope is the weighted sum of the two cases, and they are equally likely, so it is \$125.

Another oscillation - Mark now leans towards SA.

I believe this to be a good argument for this version of the game. There isn't any paradox about the value of the choices changing. So the second argument in your puzzle must be correct, and there must be a bug in the first argument. It must be that the game changes, but how? Why?

Say I choose the second envelope and don't open it; now argue about the expected value of switching to the envelope with the \$100 in it. The value of sticking stays at \$125. Shit.

Go back over my argument that the paradox exists without opening either envelope. I have y and the expected value of the other envelope is $1.25y$. Shit.

Well, it's been 5 hours since I started on this. Very embarrassed. Got to get some sleep. I don't think I can avoid thinking about this.

And the next morning when Mark looks at the paradox again, he swings back to NA. But

being as he is “un-Karen-like”, he recognizes that he has not resolved the paradox.

In problem P, when I open the envelope, I am not working with extra knowledge, and I do not gain extra knowledge. I don't change the puzzle except to put a name on the amount in the envelope, so it can't possibly change the value of the other envelope. Now I just have to refute the other half of the argument, but I can't.

(two days later)

I sent Mark a long message back explaining to him my understanding of his process in thinking about the paradox. At various points I quote some of his text and ask him about it.

Uri: (quoting from Mark)⁶²

*>So what is the resolution of the paradox, that switching now
>seems to have a positive expected value of $.5(100 - 50)$, which is, let's
>see, \$25? The fact that this number turned out not to be \$50 leads me to
>guess that it must have something to do with zero-sum assumptions. I never
>did read that part of game theory very carefully.*

Uri: What do you mean by zero-sum assumptions in this situation? Why would you expect the expgain [expected gain] to be \$50?

⁶² It is an email convention when including text of another's message on which you wish to comment, to prefix each line with a “>” symbol.

With characteristic honesty, Mark reveals that he has attributed his lack of resolve to a specific mathematical content deficit. This further weakens his confidence that he can resolve the paradox.

Mark: I'm just grasping at straws. I think of an area of game theory that I am unsure about, and guess that the resolution of my dilemma may lie hidden in there. No calculation to the \$50, just a feeling that it would have been have of \$100. No way for me to justify these remarks.

In response to another of my detailed questions, Mark finally decides that he has come to a dead end and decides to stop thinking about the paradox for now.

Yeah. But what I'm really doing here is casting about for a hint about what could be wrong. Using what felt like consistent logic failed me, so I am brainstorming for a stream of thought that could lead me somewhere else. I'm not being logical at this stage.

The reason I argue that the expected value should be constant is that I don't see what could change it. In the Monty Hall problem, it seems at first that seeing the goat hasn't changed anything, but when I look closer, I see that it has revealed more information, so the expected value can change. In this problem, I don't see anything that has changed by opening the envelope, because I get into the same conundrum without even opening the envelope.

I'm still stuck..

One point that particularly stands out for me about Mark's interview is the degree to which his formal knowledge (and his knowledge that he lacks other formal knowledge together with his belief that in that formal knowledge must be the key) gets in the way of his good mathematical instincts.

Gordon

Gordon is a research computer scientist with strong interests in physics and philosophy. In this section's final interview, I will examine the development of Gordon's thinking about the envelope paradox. We shall see that it takes many twists and turns, and proposes many related problems along its path. Of the interviews I present about the envelope paradox, in Gordon's interview we will see the

greatest degree of development of intuitions and concretion. By the end of the interview, Gordon has concretized the ideas he constructs through “envelopes” that he makes connections between them and research he is doing in an entirely different field.

Gordon’s first (or as he says, zeroth) reaction:

Hm, a little reminiscent of the Let's Make a Deal problem, but with only 2 boxes instead of 3...

0th reaction: it's got to be neutral to switch. you've got zero info about which box has more.

This response is a bit atypical in that in the same breath Gordon mentions the Monty Hall problem in which it pays to switch and yet has a first intuition that switching is neutral in this situation.

Gordon continues:

let's see, argument 2 supposes a probability distribution over the actual amounts involved... or does it? there's a 50% chance that your box is the 2x box, so... I guess that really does translate into a 50% chance of 200 or 50 in the other. hm.

...no, that's right, you don't know the probability that your box is the winner, cuz that's a function of both the 50-50 choice and the probability distribution over the absolute amounts, which distribution is unknown. so then there >isn't< a 50% chance of having the winner? hm. well, certainly if you >did< know something about the pdist [probability distribution](e.g. that the probability of 100/200 was negligible), then that would affect the odds now.

Here Gordon notices that the distribution of possible envelope amounts has not been specified. For some interviewees, this observation was sufficient to dispel the paradox with a feeling of resolve.

OK, I guess that's right. given the new info, there might be > or < than a 50% chance of having the winning box. you don't know which, which ignorance reestablishes a symmetry. the most intuitive default distribution of possible contents would assign exponentially lower probabilities to higher amounts, which would cancel the expected-gain advantage associated with the \$200, resulting in an even expected gain.

Here Gordon seems to want to find a way to support NA - his intuitive choice. In the previous paragraph he has found an argument against SA, namely that the probability of having the winning envelope is not necessarily equal to 0.5, but depends crucially on the

a priori probability distribution of possible envelope amounts, hence the expected gain calculation is not valid. Since Gordon seems to want to believe NA, it seems that any hole in an argument against NA becomes an argument for NA. Gordon continues in his support for NA by making the seemingly implausible suggestion that the intuitive distribution is such that it would *exactly* cancel the expected gain advantage of the larger amount.

I asked Gordon about this in my next message:

U: I'm not entirely sure what your conclusion is here. Are you making the quantitative assertion that the exponential distribution that you think is intuitive exactly cancels the expected-gain?

G: sorry, i stated my answer vaguely.

the real answer is: the likelihood as to which envelope contains more, and hence the expected gain of the choices, and hence the correctness of switching, depends on information not given, namely the a priori probability distribution for the value of X. so, in any event, argument 2 is wrong: its assertion that the other box is equally likely to contain \$50 or \$200 is unfounded.

in real life, there'd be some commonsense criteria that could be invoked to get some intuitive estimate of the missing probability distribution. one criterion is that the odds get lower as the amount gets higher, more or less exponentially. that would run counter to the advantage that the \$200 gets in the expgain calculation--though of course whether it exactly cancels out depends on the details of the distribution function.

so part of the problem is just a violation of math-problem conventions: answering the question requires invoking additional real-world knowledge of the sort that math problems conventionally abstract away from. my actual behavior would depend on what i guessed the a priori odds were for various values of X. if, say, you were to actually enact this with me, i'd assume that the \$100 was already a ridiculously high amount of money to be spending on this exercise, and that \$200 would be even less plausible, so I'd assume the other envelope probably had \$50 and I'd keep the \$100.

if additional real-world knowledge is disallowed (as is normal for math problems), then the problem is tantamount to: let X be a positive integer. what are the odds that X is greater than 100? clearly unanswerable based on the stated info.

Here Gordon decides that neither argument NA nor argument SA apply, but that the probability of having the winning envelope and hence the rationality of switching is

indeterminate. This is another common stopping point (or point of resolve) for interviewees. The puzzle has been shown to be meaningless. However, since Gordon's reduction of the paradox depends on the unknowability of the distribution, I decide to take away that assumption and send him one more message before the weekend.

U: So, to pursue one fork of our prior conversation: [email conversation] Suppose then, that the "a priori" pdsitribution [probability distribution] is known to be flat (uniform). Does argument 2 work then?

On Monday, back on email, Gordon replies:

G: I was thinking about this some more over the weekend (sorry, I couldn't resist). I realized I was mistaken to say there was insufficient info if the a priori distribution is unknown. After all, it was unknown before opening the box, but the question was answerable then: switching is neutral (or, to put it more forcefully, it would be incorrect to agree to pay a commission of say 1% to be allowed to switch, or to not switch). So if you do open the box but know nothing about the ap dist, [a priori distribution] opening the box doesn't tell you anything relevant, so it should still be neutral.

Here Gordon rejects indeterminacy and advances a positive argument for NA: Before opening the envelope no information was needed about the distribution of amounts to know that the choice was neutral and opening the envelope does not change anything relevant.

Moreover, the expgain [expected gain] argument applies even before the box has been opened. Let X be the amount of money in the box. If there's a .5 chance of that being the greater amount, then the expgain of the other box is $1.25X$, arguing for switching. But given the other box, there's the same argument for switching back.

Here Gordon discovers the infinite switching consequence of SA, almost always experienced as a clear sign that something is very wrong.

But treating the choice as neutral is tantamount to assuming that $P(X/2)=2P(X)$ -- that's what's needed to make the expgain come out equal. But it can't be right to make that arbitrary assumption...especially since if the utility of cash were nonlinear, then the equal-expgain assumption would then dictate a different assumption about the relative a priori probabilities.

But NA is also seen as "weird" - though not so weird as infinite switching.

And, as you now point out, it could be postulated that the ap dist is flat. Or simply that $P(50)=P(100)$, regardless of the rest of the distribution.

So let's see. If the only a priori choices were $X=50$ or $X=100$, with equal probability, then I guess the expgain argument does dictate switching if the box contains \$100.

A transformation of the puzzle which makes NA stronger.

But then the same should apply if there were many a priori choices, because once you see the \$100 then $X=50$ and $X=100$ are the only possibilities left, and all that matters is their relative a priori probabilities.

And an argument for why the transformed situation is equivalent to the original.

But then that argument dictates switching (even for say a 1% fee) no matter what's in the box, without even having to look (modulo boundary conditions if the range of possible values for X is finite; but that can be made negligible if the range is very large). And then, if you still haven't looked, the same argument dictates switching back (even for another 1% fee), which is obviously wrong.

Here the parenthetical remarks are especially interesting. Exploring boundary conditions would probably lead to seeing the impossibility of an infinite flat distribution. But that path is shunted into the parentheses. Could it be that the “obvious wrongness” of the infinite switching situation is less firmly entrenched than it appears? Perhaps infinite switching is more plausible than rejecting infinite flat distributions? At this point, Gordon seems to be inclined to choosing one argument or strategy for all possible cases, but this inclination is beginning to crack.

So I'm firmly in the grip of the paradox right now. More later.

After this message I saw Gordon in the hall and he told me that he couldn't yet see a way to resolve the paradox. But, he said, he was taking it very seriously. “I can't understand”, Gordon said, “how some people don't get bothered by such paradoxes - don't they recognize their importance?”

This last comment was a bit reminiscent of Greg's comment about his wife Karen. For Gordon both of the arguments, NA and SA were familiar friends.⁶³ There could be no question that their disagreement had to be resolved.

I felt sure that Gordon would come back to the parentheses. I did not send further questions and waited for his next message.

G: Oh, I see. I was being sloppy about whether to think of x as an integer or real. If it's an integer, then the box content conveys info as to whether it could be a doubled value--if it's odd, it can't be, so definitely switch; if it's even, hm, then it's 50-50 (assuming flat ap dist), so the expgain argument again says to switch. hm. so that doesn't help.

Sure enough, Gordon begins to look at the "data type" of the possible envelope amounts.

oh, right, if the range is finite, then if you're in the upper half you know not to switch, even though the amt [amount] is even. so the two pieces of info are even/odd and lo/hi half of the range. but switch iff [if and only if] it's in the low half, regardless of parity. no neutral case.

Once you look at the type of the data, it's clear that the meaning of "pick any x " or "flat distribution" needs concrete elaboration. Different "flat" distributions give different results.

ok, but then what if it's continuous and/or unbounded. or if it's discrete and bounded, but you're not told the bound. cuz then the switch-for-any- X argument resurfaces.

well in the latter case there's immediately the question of the a priori distribution of the possible bounds.

spoze it's unbounded and discrete. then can there be a distribution that assigns equal probability to all values of X ? if not, then we don't have to worry about switch-for-any- X resurfacing. same for unbounded and continuous. i'd guess there isn't any such distribution.

Here for the first time, Gordon clearly, though still tentatively, prefers to reject infinite switching and consequently the possibility of infinite flat distributions. Confidence in this

⁶³ In fact Gordon had been writing an ethics paper which involved making expected utility arguments. One level up, paradoxes were also friends of Gordon's and attracted him to their community.

rejection seems to have been buoyed by the defeat of infinite switching in the bounded cases.

and meanwhile, back at the original problem, the equal-probability postulate does indeed require switching to realize the better expected gain. but without any postulate about the ap dist, it's neutral.

Here, by the “original problem”, Gordon means when you are told that the amount, N , in the envelope you chose is just as likely to be the smaller amount as the larger.

but then what about the argument that neutrality implies acting as though $P(x/2)=2P(x)$? or a different dist for nonlin utility? well, that's weird, but not necessarily paradoxical; i guess if you've got 0 info about something then it's pretty much at the mercy of arbitrary considerations.

and given a flat distribution over some finite but unknown range (with unknown distribution over the possible bounds), the choice is also neutral, with or without looking in one box (except that you should of course switch if x is an integer and the amt is odd; and if you're told the odds are equal for the particular amt observed, then switch). the any- X expgain argument fails because the expgain is unknown without knowing the ap dist of the bound.

At this stage, Gordon had reached a potential point of resolve. I'm a bit unsure how he is thinking about infinite flat distributions, so I decide to ask.

U: I'm not sure what your meaning is here. You seem to be suggesting that the stipulation of flat a priori distribution is impossible?

My explicit question seems to prompt Gordon into solidifying his impossibility claim. In the next day's message he writes:

I'm saying a flat distribution is impossible if there's no upper bound on the possible values of X (where the envelopes contain $X, 2X$).

Say X ranges over the positive integers. If infinitely many values of X have nonzero probability, then the distribution can't be flat, because for any nonzero p , $p \cdot \text{infinity} > 1$; but the probabilities have to sum to unity.

Similarly, if X ranges over the positive reals, a flat density function is impossible except within a finite bound on X , because the integral from 0 to infinity of $p(x)$ is infinite for any nonzero constant value of $p(x)$.

Gordon is now at a resting place but he continues to elaborate the puzzle, finding new situations and connecting them up, checking whether they fit into his new perspective.

Suppose you have a version where you're told that given the amount N in the envelope you chose, the a priori probability $P(X=N)=P(X=N/2)$, where X is the smaller amount in the two envelopes. Then you should switch, by the expected gain argument, even if you don't open the envelope. But without knowledge of the rest of the a priori distribution, the two probabilities aren't necessarily equal for other values of N , so the switch-for-any- N argument doesn't follow. E.g. if $X=50$ and $X=100$ were the only two possibilities, each with .5 a priori chance, then $P(X=N)=P(X=N/2)$ iff $N = 100$.

A situation in which SA works but doesn't lead to infinite switching:

Suppose you're told instead that there was a flat a priori distribution, which, by the earlier argument, is only possible if X has a finite bound (i.e. the distribution can't be flat over all the +integers or +reals). The switch-for-any- N paradox disappears because if X has a finite range, then it's not true that for any amount N in the envelope you chose, $P(X=N)=P(X=N/2)$ --that's only true for $N \leq \text{bound}$, cuz otherwise $P(X=N)=0$. Further, since you don't know that for your particular N , the a priori $P(X=N)=P(X=N/2)$, there's no argument for switching in the first place, whether or not you open the envelope. So just being told that there was a flat, bounded a priori distribution with unknown bound (or a known bound with unknown N , i.e. unopened envelope) still leaves switching neutral.

Suppose you're told instead that there's a flat, bounded a priori distribution, and that N in the opened or unopened chosen envelope is \leq the bound; or you're told the bound and you open the envelope and see that N is \leq to it. Then, by the expected gain argument, you should switch.

Now, suppose you're told instead that $P(X=N)=P(X=N/2)$ and $P(X=M)=P(X=M/2)$ (where M is the amount in the unchosen envelope), without being told anything else about the a priori distribution (e.g. it's not necessarily flat). This might seem to argue for switching back and forth, by the expected gain argument. But in fact, the expected gain argument for switching fails, whether or not you see the value of N , because e.g. $P(X=N) \neq P(X=N)/(P(X=M)=P(X=M/2))$. So switching is again neutral.

Similarly, suppose you're told only that there was a flat bounded a priori distribution, and that both envelopes contain an amount \leq bound. It follows that the a prioris $P(X=N)=P(X=N/2)$ and $P(X=M)=P(X=M/2)$. But given that N and M are both below the bound, you can deduce a new flat distribution with half the bound of the original. The size of N and M compared to the new bound is unknown; hence, the expgain argument again fails and switching is again neutral.

More situations are elaborated and shown not to fall prey to infinite switching.

I am now convinced that Gordon is ready to end this conversation. I ask a wrap up question:

U: Do you feel like you have answered the original question and its related

puzzles to your satisfaction? Or are there still some puzzles left for you in it?

G: I'm satisfied with my understanding of it now.

I ask Gordon if he has any closing comments:

G: Well, the paradox was that there were too [sic] apparently valid arguments leading to contradictory conclusions. Refuting the expgain argument leaves standing the argument for neutrality: whatever the distribution, randomly choosing between the two boxes makes subsequent switching neutral in the absence of further relevant information; and knowing N without knowing the distribution isn't relevant information about whether to switch.

It's interesting that, given the paradoxical arguments, it's much easier to determine that the neutrality argument must be right (because of the switch-for-any- N absurdity that follows from the alternative) than to actually find the flaw in the other argument (which requires noticing the boundary problem and the impossibility of a flat unbounded distribution).

The value of $P(x > 100)$ is indeterminate, i.e., there isn't enough info given to answer the question. Probabilities and expected gains are meaningful only with respect to underlying distributions, and none is specified here. Neutrality is a unitary ratio of expected gains. That ratio can be knowable even if the distribution and hence the expected gains themselves are unknown (as is the case in the envelopes problem when the distribution isn't given) or infinite.

Gordon concludes our email conversation in declarative sentences. He is now quite confident in his understanding, which is framed less in terms of the envelope paradox and more in general epistemological terms – he has built a concrete relationship with the ideas of probability, expectation, distribution and neutrality.

Note that only in this last fragment does Gordon assert that one can know a ratio of expected gains without knowing the expected gains themselves. Therefore one can know that the “envelopes” choice is neutral without knowing the distribution and hence without knowing the probabilities or expected values. Most interviewees who came to this same conclusion did so by doing substantial calculations. Some interviewees, for example, calculated the value of the sticking and switching strategies for each value of a discrete bounded distribution. Only when the gain gotten from switching with a low number was exactly balanced by switching from a high number did they feel comfortable

with neutrality. Gordon, however, seems to infer the compensation of the calculation from the necessity of neutrality. This difference is reminiscent of Piaget's analysis of water volume conservation. As a child is transitioning from pre-conservational to post-conservational, they need to experiment to see that the extra width of the wide glass compensates for the tallness of the tall glass. This compensation is the explanation for the equality of the liquids. Only when conservation is more solid do children make the opposite inference -- since no water was lost in the pouring, we can infer the compensation from the necessity of the equality of volume.

I have run into Gordon since our interview was concluded. He told me that he is continuing to elaborate his thinking about the paradox. He now sees that the resolution of the envelope paradox fills an essential gap in a paper that he is writing about a theory of ethics. He asks if it is alright for him to include it in his paper. When a mathematical object or situation is deeply elaborated and concretized then it is part of a community of ideas and will be called upon again and again to take an active role in that community.

The envelope paradox is a particularly powerful learning paradox. Getting engaged with it leads us to see that probabilities depend crucially for their meaning on distributions. As you may recall, recognition of this dependence was missing for the students who used the NSF-sponsored probability software which was designed to bring home this point. But, as we have argued, learning the "true" conceptions is unlikely if you don't get to construct and mess around with many alternate conceptions. All "envelopes" interviewees constructed many different transforms and variants of the "envelope" paradox and used their intuitions about these variants to connect to and critique "envelopes" arguments. Learners engaged with the envelope paradox also learned how easy it is to be seduced by an expression such as "pick a random number". They learned that, when choosing numbers, it is very important to consider how the choice is made, what mechanisms govern the choice, or what process or distribution generates it.

Additional Starlogo Programs

In this section I describe two additional *Logo programs that were designed and coded by interviewees and myself in order to investigate some probabilistic questions.

Binomial and Normal Distributions Revisited

One of the interviewees, Alan, was intrigued by the second interview question - why is height normally distributed? Although, he is mathematically quite sophisticated, Alan said he “never really understood normal distributions”.

Alan made a conjecture as to why height is normally distributed:

A: You start out with this one couple, Adam and Eve say, and they're a certain height. No, make this simpler, we just have Adam and he's a certain height. Now let's suppose we have parthenogenesis, (is that the word?) and Adam has kids on his own. And suppose his kids are a bit off from where he is (that is different heights) due to copying errors or something. Then they sort of form a little bell curve - anyway a binomial distribution. Now suppose they have children and that process continues many generations, then wouldn't you wind up with a normally distributed population?

To explore this model, Alan decided to write a *Logo program with me.

The program is initialized to have thousands of “rabbits” at the bottom middle of the graphics screen. (Call the initial location of the rabbits the “origin”. For reference, a green vertical line is placed at the origin.) Rabbits can hop right or left from the origin on a line at the bottom of the screen. (Call it the hopping line) Each rabbit has a set of possible step sizes which it can hop in one clock “tick”. Associated with each step size is a probability - the probability that the rabbit will choose that step size. To make the interface easier for users who might not be comfortable with probabilities (expressed as real numbers between 0 and 1), the probabilities of rabbit steps are expressed as integer relative odds. The rabbits are initialized with the command “setup-rabbits” which takes

three arguments: the number of rabbits, a list of step-sizes and a list of probability-ratios. So, for example, if one wanted to create 1000 rabbits that could hop to the right 2 units with probability $3/4$ and hop to the left 1 unit with probability $1/4$, one would type the command

```
setup-rabbits 1000 [2 -1] [3 1].
```

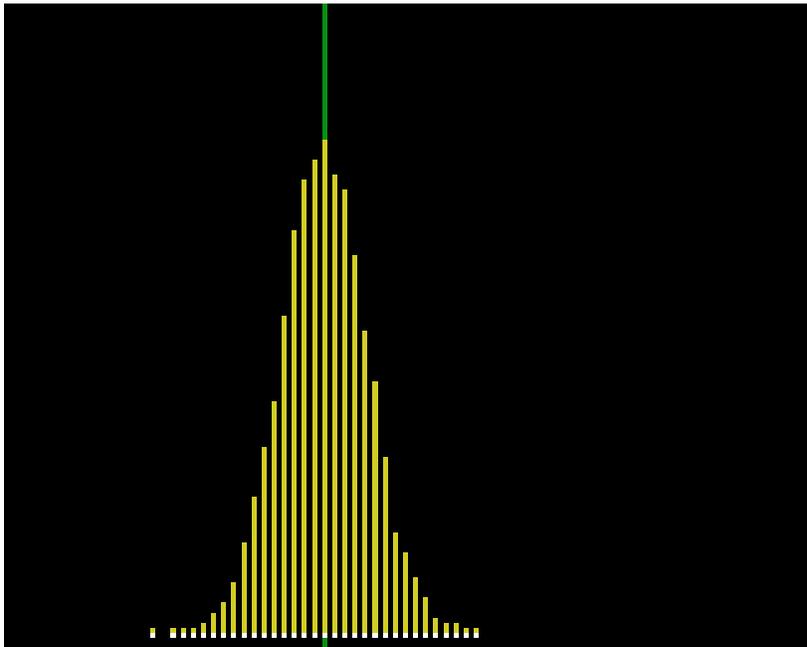
After executing this command, the rabbits would be all piled on top of each other at the middle bottom of the screen. It is then possible to perform experiments and let them hop for a while. At each clock cycle each rabbit picks a step to hop according to the probabilities and hops that distance on the hopping line. Once the rabbits have hopped for one clock tick, they are no longer all in the same place. They are still on the hopping line, but at different x positions on that line. Above each position at which there are rabbits present, we display a yellow column - the height of which is proportional to the number of rabbits at that location. In this way an evolving “living” histogram of the distribution of rabbit locations on the hopping line appears on the screen.

To make this a bit clearer, let’s take a simple example:

If we initialize the rabbits with the command `setup-rabbits 1000 [1 -1] [1 1]`, then 1000 rabbits will appear at the origin. Each rabbit has been initialized so that it can only hop either one unit to the left or one unit to the right with equal probability. If we let them hop one step, then because the probability of moving left is the same as the probability of moving right, approximately half the rabbits will move one unit left and another half will move one unit right. As a result two roughly equally high yellow columns will appear on the screen above locations $x=1$ and $x=-1$. If we now let the rabbits hop one more step, there will be a tall yellow column in the middle and shorter, roughly equally high, yellow columns to the right and to the left.



setup-rabbits 8000 [1 -1] [1 1] after 10 ticks



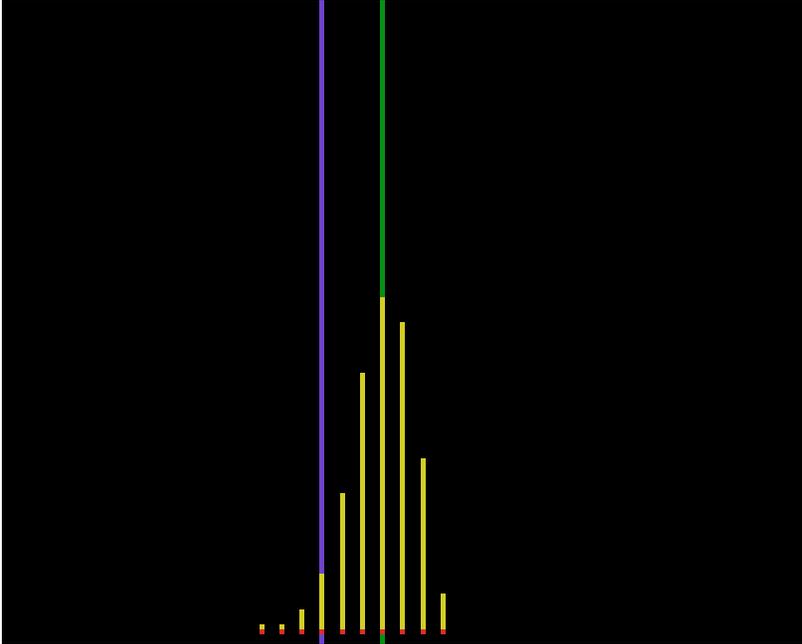
setup 8000 [1 -1] [1 1] after 40 ticks

In the above simple example, the average location of the rabbits does not move from the origin. To help visualize the movement of the rabbits when the average location does change, a purple column is displayed above the average location and a green column “remembers” where the rabbits started. In this example the histogram of rabbits is also symmetric. Natural questions that arise in this environment, then, are: what will cause the average location of the rabbits to change and what will cause more of them to be one side of the average?

As you may recall, the reason Alan and I created this mini-environment was to investigate Alan’s theory of height distributions. One way to think of the rabbits when they’re all at the origin is as a population of individuals all of a standard height. When rabbits take a step they represent new individuals in the population who have deviated slightly from their parents’ height. In our simple example the average height of the population does not change and the population heights are symmetric about this average.

When viewing this example, Alan exclaimed: “I see why population heights are symmetrical, the rabbits hop the same distance to the right as to the left”. He then typed in the command `setup-rabbits 8000 [2 -1] [1 1]`. “Now”, he said, “the graph won’t be symmetrical, it will skew to the right”. But to Alan’s surprise, when the rabbits hopped, the histogram remained symmetrical, but the purple line representing the average rabbit location or average population height moved to the right.

Alan quickly understood the “bug” in his thinking: “I see. Rabbits hop farther to the right when they hop but they’re still just as likely to hop right as left. So, the distribution will remain symmetric, it will just move to the right one unit on average for each step. To make the distribution asymmetric, you need to change the probabilities not the step sizes.”



setup 8000 [2 -2] [3 1] after 10 hops

From the above experiment Alan concluded: ⁶⁴

If children are just as likely to be shorter as they are taller than their parents but when they are taller their difference from their parents is larger, then the population as a whole will slowly⁶⁵ get taller. But since we know that height of the population as a whole is symmetrically distributed, children must be equally likely to be taller as they are to be shorter than their parents. I see why normal distributions are so common. Whenever we make a measurement, we're just as likely to make a mistake in one direction as the other, so the resultant distribution of measurements will be normal. Its average will be the true value of the measurement and the spread of the graph will depend on the accuracy of the measurement.

Alan then went on to investigate what would happen to the distributions if he gave the rabbits a much larger “palette” of step sizes.

In summary, this *Logo mini-environment allows the user to explore distributions from the point of view of the individual constituents as opposed to the more standard macroscopic view in which distributions come to be known through their global parameters. Through exploring variations in the behavior of the individual constituents,

⁶⁴ This is a paraphrase of Alan's conclusions reconstructed from my notes.

⁶⁵ Presumably slowly because the difference of differences (or asymmetry) is expected to be small.

we see which mechanisms must be in place to create a normal distribution and how altering those mechanisms changes the shape of the distribution. We see the normal distributions as part of a family of distributions which emerge from different local rules or mechanisms. No longer is the epistemological status of this distribution in question as it was for Lynn in the first interview. We have constructed it and understand what must characterize phenomena which it describes.

Particles in Collision - *LogoGas

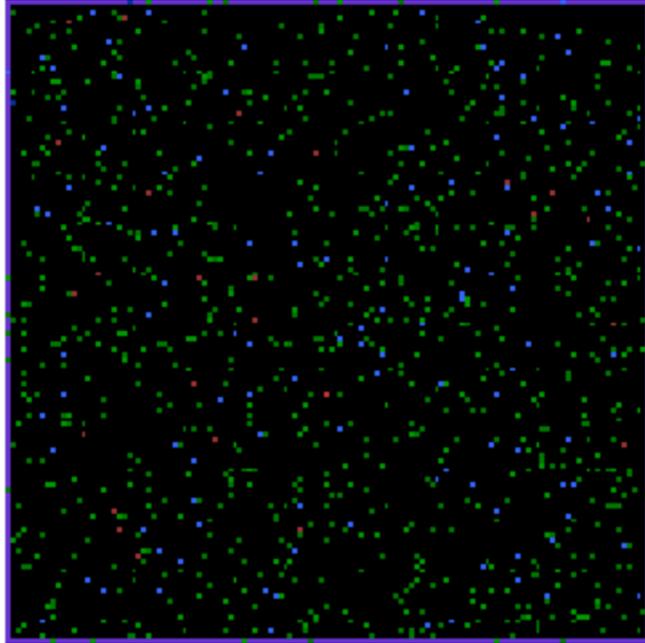
Harry was very interested in the behavior of gas particles in a closed chamber. He wanted to develop a *Logo environment in which he could visualize what happens over time to a mass of particles as they collide with each other and the edges of their container. Harry had learned that the energies of the particles, when graphed, form what is known as a Maxwell-Boltzman distribution. But he had no “feel” for why this should be the case and he hoped that the *Logo environment would give him insight into this phenomenon. Harry, a graduate student in education, is also a science teacher in the Boston public schools and he wanted this environment to be useful for thinking about large scale ensembles such as frequently appear in physics and chemistry. In accordance with this desire, the “*LogoGas” environment was used by a number of students and some of their experiences and reactions will also be described below.

In typical use, the program is initialized to display a box with a specified number of gas “molecules” randomly distributed inside it. The user can then perform “experiments” with the molecules.

Typically, the molecules are initialized to be of equal mass and start at the same speed (that is distance traveled in one clock tick) but at random headings. By typing the command `diff :n`, the molecules move “n” steps”. Collisions with the container or between molecules are “elastic”, (i.e., no energy is “lost” to heat).

There are several ways of displaying information associated with the distribution of speeds and energy. Choosing a display method turned out to be the most difficult design decision in the project. What display would most afford intuitive access to the information? Which display would encourage questions and explorations? Would choosing a display focused on the individual molecules be more engaging than a display that summarized information in a graph? In the end, of course, we chose to have many methods of display. Chief among these are three:

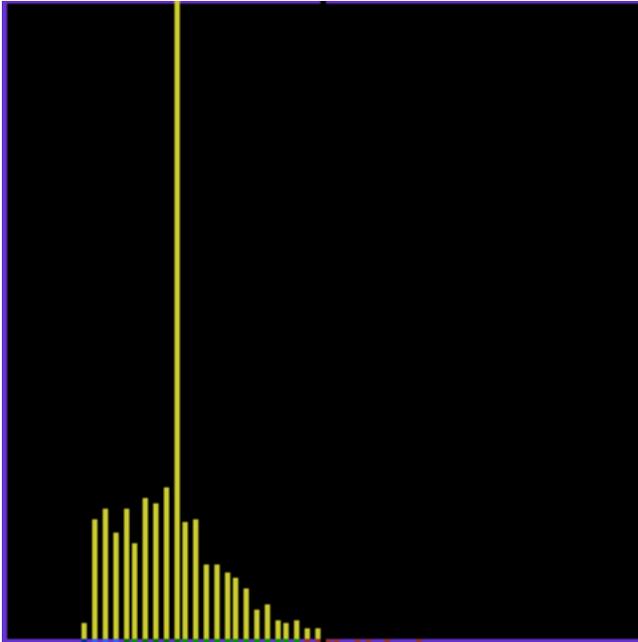
1) Faster molecules get brighter colors, slower molecules darker colors.



8000 gas particles after 30 ticks

2) Colliding molecules appear in red, molecules “bouncing” off the box appear in green, and all other molecules appear in a shade of blue, the faster the molecule, the brighter the shade of blue.

3) A “live” histogram of the distribution of molecule speeds or of molecule energies.



a histogram of molecule energies after 40 ticks

*Logo naturally facilitates inquiry into the molecules' states. So users can ascertain the mass, speed, energy or any function of these for individual molecules. One can also track individual molecules through color labels. Also built into the environment were procedures for computing average physical quantities, such as average molecule speed and average molecule energy.

When this environment was finally completed⁶⁶, its most striking feature was its capacity to attract, captivate and engage random passers-by. People whose idea of fun did not include working out physics equations nonetheless were mesmerized by the motion of the shifting gases, their pattern and colors. And they were motivated to ask questions - why do more of them seem to slow down than speed up? What would happen if they were all released in the center? in the corner?

⁶⁶ And this took a while due to instabilities in the *Logo environment which resulted in numerous bugs for Mitchel Resnick and myself to fix.

This enthusiasm was pleasing to us. We decided to put in many more facilities for experimentation such as different box shapes, multiple boxes, different gases mixing together etc. If hot or cold plates were inserted into the box, how would that affect the ensemble? What if there was some energy lost at each collision? We imagined learners playing with the environment, watching molecules heat up and become more gas-like or cool down and become more solid-like.

But while we were writing new procedures and primitives for the environment, we encountered a puzzle that distracted us from our immediate pursuits. We had been keeping track of the average speed of the molecules and it was printing out on another screen. When looking at the other screen, we noticed that the average speed was going down. This surprised us. After all, we reasoned, the collisions are all elastic, so no energy is lost from the system. Since the number of molecules isn't changing, the average energy or $\frac{\text{total-energy}}{\text{number of molecules}}$ should also be a constant. But energy is just proportional to the mass and the square of the speed. Since the mass is constant for all molecules, the average speed should also be constant. Why, then, was our program saying that average speed was decreasing? In Harry's words:

The IMPLICATION of what we discovered is that the average length of each of the individual vectors does indeed go down.

PICTURE IT!! .. I visualize little arrows that are getting smaller.

These mental vectors are just that. Little 2 (or 3)-dimensional arrows.

*The move to the scalar is in the calculation of energy (with its v^{**2} terms.)*

Doesn't it seem difficult to reconcile the arrows (vectors) collectively getting smaller with a scalar (which is a quantity that for a long time was visualized as a fluid) 'made up' from these little vectors NOT getting less!

We were dismayed by this new “bug” and set out to find what “had to” be an error in the code.

But there was no error in the code. After some time unsuccessfully hunting for the bug, we decided to print out average energy as well. To our surprise, average energy stayed constant.

At this point we realized the bug was in our thinking rather than in the code. After some reflection, we identified two sources for our bug.

The first bug had to do with our knowledge that momentum⁶⁷ is conserved. Since mass is constant, this means the average velocity as a vector is constant. Since the average velocity is constant, shouldn't its magnitude, the average speed, be constant? Again, we missed the obvious – that the magnitude of the average vector is not the average speed. The average speed is the average of the magnitudes of the vectors. And the average of the magnitudes is not equal to the magnitude of the average.

The second bug was thinking that if $\sum_i v_i^2$ is a constant then it follows that $\sum_i \text{abs}(v_i)$ will also be constant. But it is not hard to show that if the former sum (corresponding to energy) is constant then the latter sum (corresponding to speed) is maximal under our initial conditions. As the speeds diverge, the average speed decreases just as we “observed”. For a fixed energy, the maximum average speed would be attained when all the speeds were the same as they were in our initial state. From then on, more particles would slow down than would speed up. This was readily visible on the computer screen where the blue dots far outweighed the red.

Although both these bugs are now obvious and slightly embarrassing, this confusion between the vector and scalar averages was lurking in the background of our thinking. Once brought to light, it could be readily dispensed with using standard vector

⁶⁷ A source of confusion in many a physics classroom: why do we need these two separate quantities, energy = mv^2 and momentum = mv . The formalism masks the big difference between the scalar energy and the vector momentum.

arithmetic. But being immersed in an environment where questions about the large ensemble were naturally both asked and answerable brought this confusion to the surface. It also enabled Harry to go back and forth between different levels, the level of the whole ensemble, the gas, and the level of individual molecules.

Reflecting on this confusion gave Harry the insight he had originally sought from the *LogoGas environment. Originally, Harry had thought that because gas particles collided with each other randomly, they would be just as likely to speed up as to slow down, so the average speed should stay roughly constant. But now, Harry saw things from the perspective of the whole ensemble. He saw that high velocity particles would “steal lots of energy” from the ensemble. The amount they stole would be proportional to the square of their speed. It then followed that, since the energy had to stay constant, there had to be many more slow particles to balance the fast ones.

This new insight gave Harry a new perspective. But it had answered his question only at the level of the ensemble. What was going on at the level of individual collisions? Why were collisions more likely to lead to slow particles than fast ones? This led Harry to conduct further investigations into the connection between the micro- and macro-views of the particle ensemble.

The most frequently asked first question users of *LogoGas asked was: Why do the particles change speed at all? When teased out, this question could be framed as: “But the collisions between particles are completely symmetric – why does one particle change speed more than the other? To answer this question, users experimented with the environment, repeating collisions of two particles in fixed trajectories. After seeing two particles collide at the same angle again and again, but emerge at different angles each time, users realized that “randomness was going on here”. It is not long from this realization to the discovery of the broken symmetry: “in different collisions, the same particles coming at the same angle, can have different axes of collision -- and their trajectories may not be symmetrical with respect to their collision axis. The bug here was

that in physics classes, particles are usually points, while these particles had extent and therefore could collide in different ways.

In summary, *LogoGas provided an environment for playing with and exploring large ensemble behavior. Through creating experiments in *LogoGas, learners can get a feel for both the macro level, the behavior of the ensemble as an aggregate, and its connections to the micro level - what is happening at the level of the individual gas molecule. In a typical physics classroom, learners have access either only to the micro level - through , say, exact calculation of the trajectories of two colliding particles, or also to the macro level, but only in terms of pre-defined summary statistics selected by the physics canon. In *LogoGas, learners can visualize ensemble behavior all at once, sometimes obviating summary statistics. Furthermore, they can create their own statistical measures and see what results they give at both the micro and the macro level. They may, for example, observe that the average speed of the particles is not constant and search for a statistical measure which is invariant. In so doing, they may construct their own concept of energy. Their energy concept, for which they may develop a formula different than the formula common in physics texts, will not be an unmotivated formula the epistemological status of which is naggingly questioned in the background. Rather, it will be personally meaningful, having been constructed with a purpose by the learner.

The idea of “average” is then seen to be another method for summarizing the behavior of an ensemble. Different averages are convenient for different purposes. Each has certain advantages and disadvantages, certain features which it summarizes well and others that it doesn't . Which average we choose or construct depends on what about the data is important to us.

Another affordance of the *LogoGas environment is its highlighting of the question of what kind of thing is the randomness of the gas. Is it a property of the gas itself, or is it just a shorthand for our inability or unwillingness to compute each exact

trajectory, so that we summarize our ignorance by allocating collisions axes randomly. This leads naturally into questions about our abilities to measure particle states exactly and to exploring systems that behave stably under perturbation as opposed to systems that demonstrate sensitive dependence to initial conditions.

One final methodological note: I, the researcher of Harry's development in *LogoGas, acted entirely as a collaborator – developing the environment and learning in it together with Harry. Indeed, on the issue of average speed, I was more confused than Harry. I will not go into detail here to what degree “expert” domain knowledge is important for facilitating a connected mathematics approach. I believe it is important to have thought hard about the domain, although the evidence from *LogoGas suggests it is not essential.

Chapter IX - Conclusion

Summary and Discussion

In this thesis, we have seen how an adherence to an outdated rationalist paradigm has been inappropriately imported into the practice of mathematics education. Through its concern with securing the validity of mathematical arguments and its concomitant formalization of basic mathematical concepts, this old paradigm has failed to address the mind of the learner. It has relegated to the background the fundamentally creative meaning making activity of doing mathematics.

We have seen that doing mathematics is a messy process. A process which involves constructing new representations, forging links between representations and contrasting alternate representations. A process which is informed by much more than formal mathematical considerations and which requires for its sustenance connections to ideas and situations outside of mathematics.

Through analyzing the concretion process, and by analogy to the conservation experiments of Piaget, we have seen that the development of mathematical intuitions takes time. Simply giving our conceptions to the learner or even providing simplified environments where our conceptions must be constructed by the learner does not suffice to develop mathematical intuitions in the learner. We cannot short circuit the process of constructing many different meanings for mathematical objects and developing the critical practices which allow evaluation and choice among these alternate conceptions. We have seen that distinguished mathematicians throughout history were regularly engaged in this kind of meaning making. I have argued that to expect the learner to accept the product of this meaning making activity rather than to undergo a similar process is fundamentally disrespectful to the learner. In any case, we have seen that failing to address the construction of meaning in the learner is not effective.

This ineffectiveness is evident through the probability interviews: both negatively -- the uneasiness of the interviewees with basic probabilistic notions despite their advanced background, and positively -- their engagement with and progress in constructing and making meaning of probabilistic concepts. The difference between interviewees' levels of engagement in a traditional mathematics class and a connected mathematics environment can also be gauged through the reaction of the teachers to the mathematician's recursion lesson.

In the probability study, we have witnessed learners blocked from developing their probabilistic knowledge by concerns about the epistemological status of concepts such as probability, random, and distribution. Though they have taken advanced courses in probability, in particular, and mathematics in general, these concerns were not addressed. They stayed in the background, leaving the interviewees unsure about the meaning of their formal knowledge and whether it applies to the interview situations they are in and, indeed, to their non-interview lives.

As noted in the "divorce" interview, many of the interviewees encounter great difficulties in the interviews and some might view them as floundering. I am not discouraged by these learners' difficulties -- their floundering in search of meaning. And, I do not conclude from them, pessimistically, a la Tversky and Kahneman, that learners cannot develop probabilistic intuitions. Rather, in these interviews, I have witnessed the growth and development of these intuitions. The progress is slow, but to my "eye" quite significant. Though they have taken university courses in probability, for many of these interviewees, the interview was their first experience of grappling with and making progress in understanding fundamental concepts of probability. Like Amy and Doug, they came to understand the ambiguity of statistics in their day to day world, learned to construct and interpret such statistics, find connections among them, and by so doing attain a measure of power over their daily lives. Like Alan and Lynn, they asked questions about the epistemological status of probability distributions and initiated a

negotiation of meaning for such distributions. Like Chris and Gary, they constructed such concepts as probabilistic box worth and came to see how “values” can depend on more than a single outcome. Like Arne, they came to see “randomness” as connected to information and to our knowledge of the world, and not, either a declaration of complete ignorance, nor a well defined mathematical description. Like Harry, they began to see more clearly the subtle relationships that can obtain between an ensemble and its constituents, and that descriptions of those relationships are not given a priori, but can be constructed and selected according to criteria of use. Like Ellie and Gordon, they came to understand the intimate connections among probabilities, distributions, and physical mechanisms. Like Dale and Mark and Brian, they grappled with the projectibility of life situations and when it makes sense to live one’s life, to act, according to probabilities. These developments are not negligible, and I argue here are of far reaching significance.

Nevertheless, it is difficult to find “objective” measures for assessing the progress achieved by the interviewees. We have witnessed these learners over periods that, while perhaps long for interviews, are quite short on the time scale of conservation. It is undoubtedly true that over the same period of time, the interviewees could have acquired a substantial body of formal knowledge, the accumulation of which we could measure with pre- and post- tests. But, this knowledge would be of the same kind as they have already acquired in school and that did not help them overcome the obstacles they encountered in the interviews. As I have argued in the introduction to this thesis, perhaps the best measure we can use to evaluate their progress, is obtained through taking the time to read the interviews as if we were the interviewees, or at least to get a more than superficial understanding of what the interviewees were thinking in the interview fragments presented. To become convinced that the interviewees made substantial mathematical progress, that they in fact were doing good mathematics, it is a good idea to work through the mathematical questions oneself. This is one reason that I do not, in

general, present an easily digested lesson after each interview topic with my “real answer” to the question.

As we saw in the study of fractional division, there may be some mathematical situations where learning a formal rule and having a formal definition of mathematical objects and operations does not get us into serious difficulty. But, probability is not an area where we can get by with such a brittle understanding. As we saw, in many of the interviews, but perhaps most clearly in the Monty Hall interviews, formal procedures simply do not work; they do not resolve the difficulties that the learners face in the interviews.

Since probability is an area of mathematics with great relevance to our lives, it is particularly important to provide learners with an opportunity to overcome the obstacles to developing their probabilistic intuitions. Recall the students who used the NSF software specially designed for working with probability distributions. Despite months of experience using this software to manipulate distributions, they do not see the connection between probabilities and distributions – probabilities and randomness are not seen to be governed by distributions. Typically, people dichotomize, seeing phenomena as “wholly random” by which they mean that they cannot discern any pattern in it and that all possibilities seem equally likely, or as deterministic, governed by predictable lawlike rules. This dichotomy is an obstacle to understanding many phenomena in the world which emerge from the *interaction* of random and lawlike processes. The kinds of constructions made by the interviewees, the negotiation of meaning for randomness, probability and distributions, are the kinds of bridges necessary to a less dichotomized view.

Response to Tversky and Kahneman

Let us now respond to some of the questions raised earlier in this thesis concerning the claims of the psychologists Amos Tversky and Daniel Kahneman. As we

have seen, Tversky & Kahneman have uncovered systematic biases in reasoning about probability exhibited by most adults. According to Tversky & Kahneman, most adults fall victim to the conjunctive fallacy whereby a conjunction of two events is seen to be more likely than just one of those events if the conjunction is seen to be more representative. And people fail to take into account base rates or prior probabilities when classifying people or descriptions. Tversky & Kahneman have detailed many such examples.

Suppose we accept Tversky & Kahneman's explanation for this data (although there are many critics who would not grant this much (see, e.g., Lopes, 1991 and Gigerenzer, 1991)), what then is the implication for learning probability? Does it follow that people cannot develop correct intuitions concerning probability judgments? Are we doomed to repeating these errors unless we keep a formal guard at our door?

Tversky & Kahneman themselves are ambiguous about this issue:

“There are three related reasons for the focus on systematic errors and inferential biases in the study of reasoning. First they expose some of our intellectual limitations and suggest ways to improve the quality of our thinking. Second, errors and biases often reveal the psychological processes that govern judgment and inference. Third, mistakes and fallacies help the mapping of human intuitions by indicating which principles of statistics or logic are non-intuitive or counter-intuitive” (Tversky & Kahneman, 1982)

From the above quotation, one could conclude that Tversky & Kahneman's objective is to catalogue the different errors people make so as to help them develop better intuitions and “improve the quality of their thinking”. Or conversely, they can be read to be saying our basic human intuitions are fixed and all we can do is “map” which intuitions are correct and which are invalid.

Unfortunately, many readers have drawn the second conclusion. Like the probability professor we encountered in Chapter VII, they say:

“We consider it self-evident that by far the best remedy for such errors is to use the formal rules that lead to the right answers rather than unaided intuition to answer questions to which they apply” (Edwards & von Winterfeldt, 1986)

This last quotation is a strong statement. It is “self-evident” that one should use formal rules instead of “unaided intuition” when making probabilistic judgments.

Throughout the previous chapter, in numerous interviews, we have seen interviewees develop more robust probabilistic intuitions, deepen their understanding of basic probabilistic ideas, connect these ideas to each other and create a library of resources and links with which to think about probabilistic situations. How can it be “self-evident” that we should rely on formal rules?

As we have seen whenever there is such a clear dichotomy between two views, it is an opportunity to look deeper at the meanings of the fundamental terms. In this case, the idea of an “intuition” seems to be one in need of clarification. What is an intuition? This is a difficult question and I will not answer it here. But we can infer some differences between the idea of intuition between those who hold the first reading of Tversky & Kahneman and those who hold the second.

In the second reading, intuitions are immediate judgments unaided by formalisms, physical or mental tools. As such they are seen as static. In the first reading, intuitions admit of many more variations: they can lead to approximate answers, act as “pointers” to resources, physical tools, or plans. They need not be immediate and most importantly they develop over time.

The Monty Hall problem is as robust if not more robust a case of people’s systematic departure from valid probabilistic reasoning as any of the Tversky & Kahneman situations. Unlike most Tversky & Kahneman situations, individuals making an incorrect judgment in “Monty Hall” do not shake off this “cognitive illusion” when it is explained to them. Yet, through substantial interaction with this problem, people do

develop intuitions that lead them to see situations of this kind differently. Proponents of the second reading might insist that no such thing is happening, Monty Hall puzzlers are just learning to ignore their intuitions and flag certain situations as counter-intuitive and dealt with by formal procedures.

But if we look at both “ontogeny” and “phylogeny” the evidence is in support of the second view. Children do find it completely “counter-intuitive” that taller glasses can hold the same amount of water as shorter wider glasses. Yet once they are conversational, it is equally intuitive to them that this is the case. Cultural evolution, also, leads to development of intuitions. At one time, professional arithmeticians performed calculations for business people who found the calculations difficult. Today, many adults find many arithmetical calculations intuitive. Note that in this latter case, it was the construction of both physical tools (books, abacuses, calculators) and mental tools (zero, place value, decimal notation) that led to the developing of better arithmetical intuitions. The tools may have first been used in a formal way but were eventually assimilated and connected to other mental and physical resources.

Looked at from this lens, the concept of “intuition” seems intimately related to the concept of “concrete” that we discussed in chapter IV. Whether something is intuitive for us depends a lot on the way we have connected it to other things that we know. As we have seen, children learn to conserve water volume, thus concretizing the idea of volume and as a concrete object volume acquires permanence, and its unchangingness through pouring becomes intuitive. Our intuitions depend on which objects and concepts we have made concrete. Those who hold a standard view of “concrete”, seeing “concrete” as a property of objects, which certain ones have and others do not, will also hold the first view of intuitions -- seeing them as a fixed property of a person. But, as we have seen, new objects become concrete through being multiply connected to other objects, through entering the network of mathematical knowledge. As probabilistic concepts such as expected value become concrete, new ways of probabilistic thinking become intuitive. If

for example, we see “probabilistic box worth” as being a concrete substance, then its conservation becomes intuitive.

The process of building intuitions, of building concrete relationships with new knowledge, takes time. It is not the kind of change that can be measured by a pre-test and then a post-test after being instructed in the correct approach. As a result of the relative stability of intuitions thorough time, many people forget that these intuitions have developed. As we saw in Chapter IV, once something has become intuitive for you, if it is understood immediately and without effort, it becomes very difficult to conceive of how it ever was not that way. As a result, it is easy to make the mistake of assuming what is intuitive is built-in, is always there, and does not change.

Once we accept the view that intuitions develop, we can use the results of Tversky & Kahneman to point at areas where current intuitions are weak and to work on building them. We can give up the pessimistic view that we must always be on guard so that our intuitions won't lead us astray and instead find how to connect our current intuitions to new knowledge.

But there are still a few puzzles to work out here. One of the pieces of evidence that Tversky & Kahneman adduce in support of the non-educability of intuitions is that even experts in probability fall victim to these errors when they think about Tversky & Kahneman like problems in informal contexts. If intuitions were educable then surely these experts would avoid these errors?

But let us remember the account of society of more. The post-conservation individual does not lose the TALLER agent. While it is not listened to when in conflict with WIDER, TALLER will surface when its activation pattern arises and no other agent conflicts with it. Agents and intuitions do not disappear. They are not eradicated when individuals develop, they are just connected differently to the other agents and intuitions in their society of mind. It is not surprising, then, that experts can display the same errors

as novices. We would expect that in contexts in which the experts have not built connections, the old agents will resurface.

Another puzzle that arises is why probability of all the content areas is subject to this peculiar claim that it is not learnable except by non-intuitive means. After all, the research literature is replete with descriptions of people's systematic errors in arithmetical tasks such as long division, fractional division, subtraction of negative numbers, etc. In addition, there is a whole literature showing that most adults do not have a Newtonian model of physics, but instead rely on an intuitive physics (e.g., diSessa, 1981) which shares much in common with Aristotelian physics. Here, too, people who have mastered courses in Newtonian physics often err when asked to think about physical situations in non-school settings. And it is common knowledge that mathematicians don't start out with intuitions about the quaternion group or curves of fractional dimension, yet we hear them refer to their intuitions about such seemingly "abstract" mathematical objects. We can turn Tversky & Kahneman's question around and ask why are not all these other areas thought to be indicative of flaws in the human hardware. We don't hear claims that say people's systematic errors in division prove that we did not evolve to do division and that we cannot divide intuitively.

One answer to this puzzle is rooted, I think, in a confusion of levels. The literature on the psychology of errors in human reasoning is profoundly committed to a view of optimal human performance as logical and rational. Rationality has been defined in that literature as maximization of expected utility. But expected utility is most crucially dependent on assessing probabilities and likelihoods. It is therefore a strong expectation in that literature that probability calculation will be *built in to the mechanisms* of the human mind. When the data show that people make systematic errors in assessing probabilities, it is then easier to let go the assumption that innate mechanisms for probabilities are calculated correctly (and thus humans are only capable of an "imperfect" "bounded rationality") than to jettison the fundamental assumption that the

mechanisms of human thinking are composed of logic and its derivative probability theory.

But even if we are to grant that rationality is an optimal criterion for human thinking and behavior, it does not follow that the mechanisms of human cognition are logical. If one accepts a society of mind theory of human cognition, then it is clear that logic is not only not on top, but it is not a substrate either. The stuff of human cognition is made up of many “clunky” mechanisms which through their multiple connections create an illusion of a planned rational agent. But even if we accept that this account of human cognition is completely correct, it imposes no limit on the degree to which humans are capable of creating networks of agents which act and reason rationally. In particular, these mechanisms impose no barrier to the acquisition of robust probabilistic reasoning.

The above arguments, I believe, give a decisive refutation of the evidence for the strong version of the Tversky & Kahneman claim -- that human beings are “hardwired” with intuitions that make them incapable of learning probability. They also strongly suggest that the weaker claim -- the injunction to distrust your intuitions about probability -- is an artifact of the poor learning environments for probability.

There remains one puzzle left in this whole story. If there are no innate barriers to learning probability and statistics, what are the nature of the barriers? Why do so many people report that they can't learn probability or that they know whenever they encounter a probability question, to immediately distrust their intuitive analysis of the situation?

Obstacles to Learning Probability

This last question can be somewhat sharpened if we ask it in this form: Our whole lives are nothing but experiments in probability, we are constantly assessing likelihoods and acting on the basis of these likelihoods, why, then, have we not developed agents that can handle these situations with ease?

As I've argued, I believe that human difficulties with probability are of the same kind as their difficulties with other mathematical ideas. That said, we can still specify what are the specific nature of the obstacles to developing concrete intuitions about probability.

Here are some important obstacles:

- Probabilities often require reasoning back in time - we live forwards so miss making the connections. As in the Monty Hall problem, we often have to reason backwards to a situation before our current situation. People have general troubles with backwards reasoning even in non-probabilistic contexts such as in classic fallacies.⁶⁸
- The ideas of probability and randomness are still new for our culture and are still controversial and under debate. We are in an early stage of our social construction -- our cultural evolution in assimilating probability.
- When we are given statistical data, it is often detached from its meaning and its method of collection. If a statistic is left vague and disconnected in this way, we can not operate on it, transform it and compare it with others so as to make it concrete.
- Short term memory limitations make it difficult to collect lots of data on the fly. So instead, we construct summaries of the data. These summaries are not analyzable into their constituent elements. Note that these same constraints made arithmetic very difficult until memory saving technologies (such as writing and place value) were invented. The arrival of the computer could well be the new technology that does for probability what decimal notation does for arithmetic. By allowing large bodies of data to be visualized at

⁶⁸ A classic Greek fallacy deduces from "if there is fire there is smoke" that "if there is smoke, there is fire".

once in a small space and by enabling large numbers of repetitions to happen in short time, we get around memory resource limitations.

- It is very hard to get feedback of the adequacy of our probability judgments. If we assess that the probability is 30% that it will rain tomorrow and it does rain, what have we learned? We require many, many such observations in order to calibrate our judgments to the data. Again, memory limitations block us from receiving the necessary feedback, but even more than memory limitations, there is little opportunity for control. The world is constantly changing and we cannot repeat the experiments in a controlled manner.

- Seeing our lives as experiments in probability requires seeing our current situation as one of a large collection of similar situations. To make this large collection concrete is a massive construction job. It requires forging links between our current situations and situations forgotten in the past, situations not yet arisen, and counterfactual situations, that is situations which could have arisen but did not. The hurdle presented by this massive construction may be responsible for a large shift in perspective such as may occur in transitioning from a deterministic to a probabilistic “mindset”.

- A powerful way to think about probability situations is to think of multiple, interacting, distributed entities. Again we cannot keep so many entities in our heads at once. So we construct wholes out of the many interacting parts and understand the behavior of the ensemble in terms of its average behavior.

The above reasons show that, in fact, memory and resource limitations do contribute to why, until now, many people have not developed robust probabilistic

intuitions. There are very good reasons, however, for thinking that we have reached a time where these limitations will no longer hold sway.

In the “judgment and heuristics” literature itself, a countercurrent to Tversky & Kahneman like research has sprung up. The new research is optimistic about the acquisition of probabilistic intuitions. Newer studies seem to show people to be, on average, more sophisticated in their thinking about Probability:

“some of our subjects showed...an appreciation of statistical principles that in previous work other people failed to appreciate. It seems more reasonable to explain [this success] by saying that they are more skilled at statistical reasoning than the other subjects than by saying they saw through the experimenters’ tricks.... We see a powerful argument in the work we have reviewed for the role of cultural evolution. it does not require unusual optimism to speculate that we are on the threshold of a profound change in the way that people reason inductively.... Most people today appreciate entirely statistical accounts of sports events, accident rates, and the weather...Will our own descendants differ as much from us as we do from Bernoulli’s contemporaries?” (Nisbet et al, 1983, p 360-362)

Future Directions

The probability study focused on a cohort of individual learners, learning through personal interviews. For the most part, these interviewees did not substantially interact with each other. Unlike the recursion study, there was no built-in community and therefore no sustained probability culture. There was clear evidence that the interviewees felt a need for such a community. Amy and Doug, after having each interviewed with me separately, sought each other out in order to explore the statistics of divorce rates. Interviewees discussed their interview questions with friends and colleagues and often these latter would call me with their own ideas about the interview questions. In the case of the Monty Hall problem, post-conservational interviewees asked to meet pre-conservational interviewees so that they could “convince them”. Interviewees engaged in the “envelopes paradox” started an email discussion group to further explore the paradox.

It would be desirable to create an environment that combined the culture and community present in the recursion study with the depth of the focus of the probability study.

As noted above, the probability study focused on a short span of time relative to the time scale of development. To notice changes in understanding of the magnitude of full scale conversations, we must study learners engaged in connected mathematics environments over substantial time periods. Perhaps, the creation of a “connected probability” community would facilitate long term observations of learners developing probabilistic intuitions. Only then could we learn if, in fact, the massive construction project needed to build all the probability concepts in these interviews is completed. If these concepts are constructed and become intuitive, we could then see if they, in fact, synergize, working together to form a larger agency. If such an agency does arise, we might be justified in describing the transition as a new way of knowing, or making the qualitative shift from a deterministic to a probabilistic mindset.

Appendix A - Probability Interview

The first group of questions was asked each interviewee in the course of the conversation.

Group 1

Question #1 was usually the opening question of the interview. Quite frequently it took the longest amount of time of any interview question. In many interviews, discussion of question #1 lasted many hours and in some cases the whole interview was seen as being a continuation of question #1 with the other questions being brought up to support or disconfirm a particular branch of the conversation.

1) What do you mean by the words:

normal

average

certain

uncertain

probable

random

The last in the series of questions contained in question #2 emerged from the conversations. It was not planned in advance but it came up so explicitly in the first interview that I became on watch for it in subsequent interviews.

2) What is the probability of a man's height being over 5'8"? How about a woman? How is height distributed in the population? If normal, why is it normal? Is this a law of nature?

3) The Monty Hall problem and its many variants. Usually, version "2a" was the one first presented, but often the others would come up in the course of the interview.

2a) Monty Hall on the TV show "Let's make a Deal" gives you a choice of 3 boxes. One of the boxes contains a valuable prize (a Ferrari), the other two contains non-valuable prizes (goats) . You choose box #1. Monty, who knows what's in each box, opens box #2 intentionally revealing a goat. He then makes you an offer: you can keep box #1 or you can trade it for box #3. Which should you do?

2b) Three prisoners who share the same cell are condemned to die at noon. That morning, a guard announces that one of the prisoners has been pardoned. When asked which one it was the guard replied: "I cannot tell a prisoner her own fate". Despite repeated pleas, the guard would not tell. One of the prisoners, Amy, persisted and was able to talk to the guard privately. She suggested to the guard that if she couldn't reveal Amy's own fate, at least she could say which of Betty and Celia were sure to die. The guard reasoned that this was OK, as she would not be revealing Amy's fate nor which prisoner was pardoned, so she told Amy that Betty would die. Amy then reasoned that her chances of survival had improved from $\frac{1}{3}$ to $\frac{1}{2}$. Had they improved?

2c) I have 3 pieces of paper. One is red on both sides, one blue on both sides, and one red on one side and blue on the other. I ask you to pick one, but only let you see one side of it. The side you see is red. I offer you an even money bet. Predict the color on the other side of the paper you chose. What's the best bet?

2d) (sometimes known as Bertrand's paradox after the French mathematician Joseph Bertrand who proposed it in 1889).

Imagine three identical boxes each of which contains two coins. One box has two gold coins, one box has two silver coins and the third has one gold and one silver coin. Select one box. It seems clear that the probability of selecting the box with two different coins is

$\frac{1}{3}$. Suppose you take one coin out of one box and it is gold. Now the remaining coin in the selected box is either silver or gold, so it seems clear that the probability of your having selected the mixed box has changed and is now equal to $\frac{1}{2}$.

The next problem isn't as close to Monty Hall as the others but it does illuminate some of the issues related to the problem.

2e) You run into an old friend at a train station. You ask her how she's doing. She says she's fine and has 2 kids. You say you have one child, a girl, and then ask her if one of her children is a girl. She says yes. What's the probability that she has two girls?

3) A law is passed in the US. Each family can have as many children as they like till they get one boy, then no more children. Will there be more boys or girls⁶⁹?

4) What does it mean when the weatherperson says there is a 70% chance of rain tomorrow?

5) Which year car are there the most of on the road in the US.? How 'bout in India? What is the shape of the distribution of cars by model year?

6) Suppose you have a table from an atlas of countries and their area. What percentage of the numbers in the table will end in 1? In 5, 9? What about a table of numbers selected at "random" by a class of kids?

7) Do you buy lottery tickets? Do you gamble? Why or why not?

⁶⁹ Questions 3 and 4 were first introduced to me by Mitchel Resnick.

8) The Elisa test is said to be 99% reliable in detecting the presence of the HIV virus. If Joan gets tested and comes out positive on the Elisa test, what is the probability that she has HIV?

The next question was presented differently in different interviews. Sometimes a way into the question was through a published newspaper statistic.

9) It is often said that the divorce rate these days is very high. How does one measure the divorce rate? Or if someone gets married this year, what is their chance of getting divorced?⁷⁰

10) What is the probability that a chord drawn at random on a given circle will be longer than a radius?

11) You are in the hospital with chest pains. The doctor tells you you have arteriosclerosis. She can either do open heart surgery right away and there is a 80% chance of full recovery, or she can wait and give you blood thinning medications. In the latter eventuality, there is a 90% chance of recovery without surgery but you won't know if it succeeds or not for 2 months. If it fails, then you can still have surgery but the chance of success goes down to 60%. What do you do?

Group 2

1) The Envelope paradox:

⁷⁰ I am indebted to Josh Mitteldorf for suggesting this statistic as a good one to think about.

There are 2 envelopes in front of you. one contains x dollars and the other contains $2x$ dollars. You choose one, open it and get \$100. You are then offered a switch if you want it. You can take the other envelope instead.

Here are 2 arguments:

A) There were 2 envelopes, one contained x dollars and the other $2x$. You got either x or $2x$. Each was equally likely, so switching doesn't help or hurt.

B) The other envelope contains either \$200 or \$50. These are equally likely. So, the expected value of the other envelope is $\frac{\$200 + \$50}{2} = \$125$. You should switch.

Which if any of these arguments do you believe? What is the bug in the other argument?

1a) X is a number. What's the probability that $X > 100$?

2) Sometimes known as the Allais paradox.⁷¹

Situation 1: choose between

gamble 1: \$500,00 with probability 100%

and

gamble 2: \$500,00 with probability 89%

\$2,500,00 with probability 10%

\$0 with probability 1%

Situation 2: choose between

gamble 3: \$500,00 with probability 89%

⁷¹ These situations can be found in Savage (1954).

\$0 with probability 11%

and

gamble 4: \$2,500,00 with $p = 10\%$

\$0 with probability 90%

3) The Aces paradox.⁷²

You are dealt four cards, face down. A person peeks at the hand and reports a fact from the following list:

- 1) I looked at one of your cards, and it's an ace.
- 2) I looked at all of your cards, and at least one is an ace.
- 3) I looked at one of your cards, and it's the ace of spades.
- 4) I looked at all of your cards, and one is the ace of spades.

Which fact improves your odds of having a flush?

4) A town has two hospitals, one big and one small. Every day the big hospital delivers 1000 babies and the small hospital delivers 100 babies. There's a 50/50 chance of male or female on each birth. Which hospital has a better chance of having the same number of boys as girls?

5) Simpson's paradox

Can the following situation arise: Each of studies A & B show that a medication is effective, but when the data from the two studies are pooled, the medication is shown to be ineffective?

6) *Logo activity - Painted Desert Challenge⁷³

⁷² Sometimes attributed to the British Mathematician Henry Whitehead.

⁷³ Developed together with Mitchel Resnick for the Third International Conference on Artificial Life.

A terrible windstorm swept through the Painted Desert. Blue, green, yellow, and white grains of sand were all mixed together. The ants in the desert don't like their new high-entropy environment. They decide to move the sand to form a more interesting pattern.

Each ant can perform four simple actions: it can take a step forward, it can turn left by an arbitrary angle, it can pick up a grain of sand (if there is a grain directly underneath it and it isn't already carrying a grain), and it can put down a grain of sand (if it is carrying one). Each ant also has several senses: it can detect whether there is a grain of sand directly under it (and if so what color) and whether there is a grain directly ahead of it (and if so what color). An ant also knows whether it is carrying a grain at the current time.

What rules should the ants follow to move the sand into a more interesting pattern? Try to create some of the following patterns:

- *Move grains of the same color near one another.*
- *Move the grains into concentric circles, according to color.*
- *Move the grains into four horizontal strips, according to color.*
- *Create your own abstract art.*
- *Draw a picture of Seymour Papert.*

References

- Abelson, H. & diSessa, A. (1980). *Turtle Geometry: The Computer as a Medium for Exploring Mathematics*. Cambridge, MA: MIT Press.
- Ackerman, E. (1988). Pathways into a child's mind: Helping children become Epistemologists. In P. Heltne and L. Marquart, (Eds.), *Proceedings of the Symposium on Science Learning in the Informal Setting*, 7-19. The Chicago Academy of Science, Chicago, IL.
- Ackerman, E. (1991). Who is designing what for whom? Who is learning what from whom? or The Clinical Method as a tool for rethinking learning and teaching. Proceedings of the annual meeting of the American Educational Research Association, 1991.
- Agre, P. (1988). *The Dynamic Structure of Everyday Life*. Ph.D. thesis, MIT, Cambridge, MA.
- Appel, Haken and Koch (1977). Every Planar Map is Four Colorable. *Illinois Journal of Mathematics*. XXI 84, September.
- Arbel, A., Rosin, M. & Dromi, D. (1992). *Geometry Inventor (LOGAL)*. Kiryat Shmona: Israel.
- Archibald, R. (1934). *Outline of the History of Mathematics*. Oberlin, Ohio: The Mathematical Association of America.
- Arons, A. (1990). *A Guide to Introductory Physics Teaching*. New York: John Wiley & Sons.
- Ayer, A.J. (1946). *Language, Truth and Logic*. New York: Dover.
- Ball, D. (1990). Prospective Elementary and Secondary Teachers' Understanding of Division. *Journal of Research in Mathematics Education*, vol. 21, No. 2, 132-144.
- Ball, D. (1990). With an eye on the Mathematical Horizon: Dilemmas of Teaching. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.
- Bateson, G. (1972). *Steps to an Ecology of Mind*. New York: Ballantine Books.
- Baulac, I., Bellemain, F., & Laborde, J.M. (1992). *Cabri: the Interactive Geometry Notebook*. Pacific Grove, CA: Brooks-Cole.
- Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's Ways of Knowing*. New York: Basic Books.
- Benaceraf, P., & Putnam, H. (Eds.) (1964). *Philosophy of Mathematics: Selected Readings*. Englewood Cliffs: Prentice Hall.
- Brandes, A., & Wilensky, U. (1991). Treasureworld: A Computer Environment for the Exploration and Study of Feedback. In Harel, I. & Papert, S. *Constructionism*. Norwood N.J. Ablex Publishing Corp. Chapter 20.
- Brandes, A. (in preparation). *Planting Seeds*. Unpublished dissertation proposal. MIT Media Lab.
- Brock, C.F., Cappel, M. (1992). *Geometry Inventor (Wings)*. Scots Valley, CA.

- Brouwer, L.E.J. (1981). *Brouwer's Cambridge lectures on Intuitionism*. New York: Cambridge University Press.
- Brown, S. (1983). *The Art of Problem Posing*. Philadelphia, Pa. Franklin Institute Press.
- Bruner, J. (1960). *The Process of Education*. Cambridge, MA: Harvard University Press.
- Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.
- Carnap, R. (1939). *The Foundations of Logic and Mathematics*
- Carnap, R. (1962). *The Logical Foundations of Probability Theory*. Chicago: University of Chicago Press.
- Charischak, I. (1993). Personal Communication.
- Chen, D., & Stroup, W. (in press). General Systems Theory: Toward a Conceptual Framework for Science and Technology Education for All. *Journal for Science Education and Technology*.
- Cohen, B. (1990). Scientific Revolutions, Revolutions in Science, and a Probabilistic Revolution 1800-1930. in Kruger, L., Daston, L., & Heidelberger, M. (Eds.) *The Probabilistic Revolution* Vol. 1. Cambridge, MA: MIT Press
- Cohen, P.J. (1966). *Set Theory and the Continuum Hypothesis*. New York. W.A. Benjamin.
- Collins, A. & Brown, J. S. (1985). The Computer as a Tool for Learning Through Reflection.
- Confrey, J. (1991). Learning to Listen: A Student's Understanding of Powers of Ten. In Von Glaserfeld, E. (Ed.) *Radical Constructivism in Mathematics Education*. Netherlands: Kluwer Academic Press.
- Cuoco, A.& Goldenberg, E. P. (1992). Reconnecting Geometry: A Role for Technology. Proceedings of *Computers in the Geometry Classroom* conference. St. Olaf College, Northfield, MN, June 24-27, 1992.
- Davis P., & R. Hersh (1981). *The Mathematical Experience*. Boston: Houghton Mifflin.
- Davis, R. (1983). Complex Mathematical Cognition. in Ginsburg, H. (Ed.) *The Development of Mathematical Thinking*. Orlando, FL: Academic Press.
- Dennett, D. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Dennis, D. (1993). Personal communication.
- Dewey, J. (1902). *The Child and the Curriculum*. Chicago: University of Chicago Press.
- Dewey, J. (1916). *Democracy and Education*. Chicago: University of Chicago Press.
- diSessa, A. (1981). Unlearning Aristotelian Physics: A study of knowledge based learning. *DSRE Working Papers 10*, Division for Study and Research in Education.
- diSessa, A. (1983). Phenomenology and the Evolution of Intuition. In D. Gentner & A. Stevens (Eds.), *Mental Models*. Hillsdale, N.J.: Erlbaum.

- diSessa, A. (1985). Knowledge in Pieces. Paper presented at the 5th annual symposium of the Jean Piaget Society, "Constructivism in the Computer Age" Philadelphia, PA.
- Drescher, G. (1991). *Made-up minds: A constructivist approach to artificial Intelligence*. Cambridge, MA: MIT Press.
- Edwards, L. (1992). A LOGO Microworld for Transformational Geometry. In Hoyles, C. & Noss, R. (Eds.) *Learning Mathematics and LOGO*. London: MIT Press.
- Edwards, W. & von Winterfeldt, D. (1986). On Cognitive Illusions and their Implications. *Southern California Law Review*, 59(2), 401-451.
- Einstein, A., Podolsky, B. and Rosen, N. (1935). Can Quantum Mechanical Description of Reality be Considered Complete?, in *Physical Review*, 47:2.
- Evans, J. (1993). Bias and Rationality. In Mantkelow & Over (Eds.) *Rationality* (in press) London: Routledge.
- Falbel, A. (1991). The Computer as a Convivial Tool. In Harel, I. & Papert, S. *Constructionism*. Norwood N.J. Ablex Publishing Corp. Chapter 3.
- Feller, W. (1957). *An Introduction to Probability Theory and its Applications.*, 1-2. Wiley.
- Feyerabend, P. K. (1975). *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: New Left Books.
- Finetti, B. de, (1974). *Theory of Probability*, 1-2. Wiley,
- Fox Keller, E. (1983). *A Feeling for the Organism: The Life and Work of Barbara McClintock*. San Francisco, CA: W.H. Freeman.
- Fredkin, E. (1990). Digital Mechanics. *Physica D*, 45, 254-270.
- Frege, G. (1968). *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*. Evanston, Ill: Northwestern University Press.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.
- Gigerenzer, G. (1991). How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases". in Stroebe & Hewstone (Eds.) *European Review of Social Psychology*, Vol. 2, Chapter 4. Wiley and Sons.
- Gilligan, C. (1977). *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Gilligan, C. (1986). Remapping the Moral Domain: New Images of Self in Relationship In T. Heller, M. Sosna, & D. Welber (Eds.), *Reconstructing Individualism: Autonomy, Individualism, and the Self in Western Thought*. Stanford: Stanford University Press.
- Gillman, L. (1992). The Car and the Goats. *American Mathematical Monthly*, volume 99, number 1, January, 1992.
- Gödel, K. (1931). On Formally Undecidable Propositions of Principia Mathematica.

- Goldenberg, E. P., Cuoco, A. & Mark, J. (1993). Connected Geometry. *Proceedings of the Tenth International Conference on Technology and Education*. Cambridge, MA, March 21-24, 1993.
- Goodman, N. (1983). *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press.
- Granott, N. (1991). Puzzled Minds and Weird Creatures: Phases in the Spontaneous Process of Knowledge Construction. In Harel, I. & Papert, S. *Constructionism*. Norwood N.J. : Ablex Publishing Corp. Chapter 15.
- Greeno, J. G. (1983). Conceptual Entities. In D. Gentner & A. Stevens (Eds.), *Mental Models*. Hillsdale, N.J.: Erlbaum.
- Greeno, J. G. (1991). Number Sense as Situated Knowing in a Conceptual Domain. *Journal for Research on Mathematics Education*, 22, 170-218.
- Hacking, I. (1990). Was there a Probabilistic Revolution 1800-1930? In Kruger, L., Daston, L., & Heidelberger, M. (Eds.) *The Probabilistic Revolution*. Vol. 1. Cambridge, Ma: MIT Press.
- Hardy, G.H. (1940). *A Mathematician's Apology*. New York: University of Cambridge Press.
- Harvey, B. (1992). Avoiding Recursion. In Hoyles, C. & Noss, R. (Eds.) *Learning Mathematics and LOGO*. London: MIT Press.
- Harel, I. (1988). *Software Design for Learning: Children's Learning Fractions and Logo Programming Through Instructional Software Design*. Unpublished Ph.D. Dissertation. Cambridge, MA: Media Laboratory, MIT.
- Harel, I. & Papert, S. (1990). Software Design as a Learning Environment. *Interactive Learning Environments Journal*. Vol.1 (1). Norwood, NJ: Ablex.
- Heyting, A. (1966). *Intuitionism: An Introduction*. Amsterdam: North Holland Publishing Co.
- Hofstadter, D.R. (1979). *Godel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books.
- Hofstadter, D.R. (1982). *Can Creativity be mechanized?* Scientific American, 247, October 1982.
- Honey, M. (1987). The interview as text: Hermeneutics considered as a model for analyzing the clinically informed research interview. *Human Development*, 30, 69-82.
- Hoyles, C. (1991). Computer-based learning environments for Mathematics. In A. Bishop, S. Mellin-Olson, and J. Van Dormolen (Eds.), *Mathematical Knowledge: Its growth Through Teaching*. Dordrecht: Kluwer, 147-172.
- Hoyles, C. & Noss, R. (Eds.) (1992). *Learning Mathematics and LOGO*. London: MIT Press.
- Jackiw, N. (1990). *Geometer's Sketchpad*. Berkeley, CA: Key Curriculum Press.
- Kahneman, D.. & Tversky, A. (1982). On the study of Statistical Intuitions. In D. Kahneman, A. Tversky, & D. Slovic (Eds.) *Judgment under Uncertainty: Heuristics and Biases*. Cambridge, England: Cambridge University Press.
- Kahneman, D. (1991). Judgment and Decision Making: A Personal View. *Psychological Science*. vol. 2, no. 3, May 1991.

- Kafai, Y. & Harel, I. (1991). Learning through Design and Teaching: Exploring Social and Collaborative Aspects of Constructionism. In I. Harel & S. Papert (Eds.) *Constructionism*. Norwood, N.J. Ablex Publishing Corp. Chapter 5.
- Kaput, J. (in press). Democratizing access to calculus: New routes to old roots. In A. Schoenfeld (Ed.) *Mathematical Problem Solving and Thinking*. Hillsdale, NJ.: Lawrence Erlbaum Associates.
- Kitcher, P. (1983). *The Nature of Mathematical Knowledge*. New York: Oxford University Press.
- Kleiner, I. (1990). Rigor and Proof in Mathematics: A Historical Perspective in *Mathematics Magazine*, 64.
- Kline, M. (1953). *Mathematics and Western Culture*. Oxford University Press.
- Kline, M. (1972). *Mathematical Thought from Ancient to Modern times*. New York: Oxford University Press.
- Kolmogorov, A.N. (1950). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Chelsea.
- Knuth, D. (1974). *Surreal Numbers*. Addison-Wesley.
- Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction*, 6, 59-98.
- Konold, C. (1991). Understanding Students' beliefs about Probability. In Von Glaserfeld, E. (Ed.) *Radical Constructivism in Mathematics Education*. Netherlands: Kluwer Academic Press.
- Kuhn, T. (1962). *The Structure of Scientific Revolution*. Chicago: University of Chicago Press.
- Kunda, Z. & Nisbett, R.E. (1986). The Psychometrics of Everyday Life. *Cognitive Psychology*, 18, 195-224.
- Lakatos, I. (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.
- Lakatos, I. (1978). The Methodology of Scientific Research Programmes. In *Collected Papers, Vol. I*. Cambridge, UK.: Cambridge University Press.
- Lampert, M. (1988). The Teacher's role in reinventing the meaning of mathematical knowing in the classroom. The Institute for Research on Teaching, Michigan State University. Research Series No. 186. East Lansing, MI.
- Lampert, M. (1990.). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. In *American Education Research Journal*, spring, vol. 27, no. 1, pp. 29- 63.
- Lampert, M. (1991). Managing the Tensions in connecting students inquiry with learning mathematics in school. In *Teaching for Understanding in the age of technology*. D. Perkins, J. Schwartz, & S. Wiske, (Eds.). Cambridge, MA: Harvard University Press.
- Lampert, M. (1992). Practices and Problems in Teaching Authentic Mathematics. In Oser, F, Dick, A, & Patry, J. (Eds.) *Effective and Responsible Teaching: the New Synthesis*. San Francisco, CA: Jossey-Bass Publishers.
- Latour, B. (1987). *Science in Action*. Cambridge, MA: Harvard University Press.

- Leron, U. (1985). Logo today: Vision and Reality. *The Computing Teacher* 12:26-32.
- Leron, U. & Zazkis, R. (1992). Of Geometry, Turtles, and Groups. In Hoyles, C. & Noss, R. (Eds.) *Learning Mathematics and LOGO*. London: MIT Press.
- Lave, J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge, England: Cambridge University Press.
- Lopes, L. (1991). The Rhetoric of Irrationality. *Theory and Psychology*, vol 1(1): 65-82
- Mason, J. (1987). What do symbols represent? In C. Janvier (Ed.) *Problems of representation in the Teaching and Learning of Mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McCloskey, M. (1982). Naive theories of Motion. In D. Gentner & A. Stevens (Eds.), *Mental Models*. Hillsdale, N.J.: Erlbaum.
- McCulloch, W. (1965). *Embodiments of Mind*. Cambridge: MIT Press.
- McDougall, A. (1988). *Children, Recursion and Logo Programming*. Unpublished Ph.D. Dissertation. Melbourne, Australia: Monash University.
- Minsky, M. (1987). Form and Content in Computer Science. in *ACM Turing Award Lectures: the First Twenty Years*. New York: Addison-Wesley.
- Minsky, M. (1987). *The Society of Mind*. New York: Simon & Schuster Inc.
- Minsky, M. & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
- Nemirovsky, R. (1993). Rethinking Calculus Education. *Hands On!* volume 16, no. 1, spring 1993.
- Nemirovsky, R. (in press). Don't Tell me how things are, Tell me how you see them.
- Neville, B. (1989). *Educating Psyche: emotion, imagination and the unconscious in learning*. Melbourne: Collins Dove.
- Noss, R. (1988). The Computer as a Cultural influence in Mathematical Learning. *Educational Studies in Mathematics* 19 (2): 251 - 258.
- Noss, R. & Hoyles, C. (1992). Looking Back and Looking Forward. In Hoyles, C. & Noss, R. (Eds.) *Learning Mathematics and LOGO*. London: MIT Press.
- Packer, M. & Addison, R. (Eds.) (1989). *Entering the Circle: Hermeneutic Investigation in Psychology*. Albany, NY: State University of New York Press.
- Papert, S. (1972). *Teaching Children to be Mathematicians Versus Teaching About Mathematics*. Int. J. Math Educ. Sci Technology, vol 3, 249-262
- Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. New York: Basic Books.
- Papert, S. (1987). *A Critique of Technocentrism in Thinking about the School of the Future*. Epistemology and Learning Memo #2, MIT.
- Papert, S. (1990). Personal Communication.

- Papert, S. (1991). Situating Constructionism. In I. Harel & S. Papert (Eds.) *Constructionism*. Norwood, N.J. Ablex Publishing Corp. (Chapter 1)
- Paulos, J. A. (1989). *Innumeracy: Mathematical Illiteracy and its Consequences*. Farrar, Straus & Giroux.
- Pea, R. & Kurland, M. (1984). On the Cognitive Effects of Learning Computer Programming. *New Ideas in Psychology* 2(1): 137-168.
- Perkins, D. (1986). *Knowledge as Design*. Hillsdale, NJ.: Lawrence Erlbaum.
- Perkins, Segal, & Voss (1991). *Informal Reasoning and Education*. Hillsdale, NJ: L. Erlbaum Associates.
- Phillips, J. (1988) *How to Think About Statistics*. New York: W.H. Freeman.
- Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: International University Press.
- Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.
- Piaget, J. (1964). *The Child's Conception of Number*. London: Routledge & Paul.
- Piaget, J. (1970). *Genetic Epistemology*. New York. Columbia University Press.
- Piaget, J. (1975). *The Origin of the Idea of Chance in Children*. New York: Norton.
- Pimm, D. (1987). *Speaking Mathematically*. London: RKP.
- Poincare, H. (1908). The future of Mathematics, *Revue generale des sciences pures et appliquees*. Vol. 19, Paris.
- Polya, G. (1954). *Mathematics and plausible reasoning*. Princeton, N.J.: Princeton University Press.
- Polya, G. (1962). *Mathematical Discovery*. New York: John Wiley and sons.
- Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.
- Putnam, H. (1975) *Mathematics, matter and method*. London. Cambridge University Press.
- Quine, W. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rabin, M. (1976). Probabilistic Algorithms. In Traub (ed.) *Algorithms and Complexity: New Directions and Recent Results*. New York. Academic Press.
- Rapoport, A. (1967). Escape from Paradox. *Scientific American*, July 1967.
- Resnick, M. (1991). Overcoming the Centralized Mindset: Towards an Understanding of Emergent Phenomena. In I. Harel & S. Papert (Eds.) *Constructionism*. Norwood, N.J.: Ablex Publishing Corp. (Chapter 11).
- Resnick, M. (1992). Beyond the Centralized Mindset: Explorations in Massively Parallel Microworlds, unpublished doctoral dissertation, MIT.
- Roberts, E. (1986). *Thinking Recursively*. New York: John Wiley and Sons.
- Rubin, A. & Goodman, B. (1991). *TapeMeasure: Video as Data for Statistical Investigations*. Presented at American Educational Research Association conference. Chicago, April, 1991.
- Russell, B. (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press.

- Sachter, J. (1991). Different Styles of Exploration and Construction of 3-D Spatial Knowledge in a 3-D Computer Graphics Microworld. In Harel, I. & Papert, S. *Constructionism*. Norwood, N.J.: Ablex Publishing Corp. Chapter 17.
- Savage, J. (1954). *The Foundation of Statistics*. New York.
- Schelling, T. (1978). *Micromotives and Macrobehavior*. New York: Norton.
- Schoenfeld, A. (1985). *Mathematical Problem Solving*. Orlando, FLA: Academic Press.
- Schoenfeld, A. ed. (1987). *Cognitive Science and Mathematics Education*. Hillsdale, NJ.: Lawrence Erlbaum.
- Schoenfeld, A. (1991). On Mathematics as Sense-Making: An Informal Attack on the Unfortunate Divorce of Formal and Informal Mathematics. In Perkins, Segal, & Voss (Eds.) *Informal Reasoning and Education*.
- Schwartz, J. & Yerushalmy, M. (1987). The geometric supposer: an intellectual prosthesis for making conjectures. *The College Mathematics Journal*, 18 (1): 58-65.
- Segall, R.G. (1991). Three Children, Three Styles: A Call for Opening the Curriculum. in Harel, I. & Papert, S. *Constructionism*. Norwood N.J. Ablex Publishing Corp. Chapter 13.
- Simon, J. (1991) . *Resampling Stats*.
- Steedman, P. (1991). There is no more Safety in Numbers: A New Conception of Mathematics Teaching. In Von Glaserfeld, E. (Ed.) *Radical Constructivism in Mathematics Education*. Netherlands: Kluwer Academic Press.
- Steen, L. (1986). Living with a new Mathematical Species. *Math Intell* 8:2 33-49.
- Stroecker, C. (1991). Elucidating Styles of Thinking about Topology Through Thinking About Knots. In Harel, I. & Papert, S. *Constructionism*. Norwood N.J. Ablex Publishing Corp. Chapter 12.
- Stroup, W. (1993). What Non-universal Understanding looks like: An investigation of results from a Qualitative-calculus assessment. Unpublished paper. Harvard Graduate School of Education.
- Tappan, M. (1987). *Hermeneutics and Moral development: A developmental analysis of short-term change in moral functioning during late adolescence*. Unpublished doctoral dissertation, Harvard University.
- Tarski, A. (1956). *Logic, Semantics, Metamathematics*. Oxford, England: Clarendon Press.
- Tobias, S. (1978). *Overcoming Math Anxiety*. New York: Norton.
- Toffoli, T. (1990). How cheap can mechanics' first principles be? in W..Zurek (Ed.) *Complexity, Entropy, and the Physics of Information*. Redwood City, Ca: Addison-Wesley. 301-318.
- Turkle, S. (1984). *The Second Self: Computers and the Human Spirit*. New York: Simon & Schuster.
- Turkle S., & Papert, S. (1991). Epistemological Pluralism and Reevaluation of the Concrete. In I. Harel & S. Papert (Eds.) *Constructionism*. Norwood N.J. Ablex Publishing Corp. (Chapter 9).

- Tversky, A., & Kahneman, D. (1971). Belief in the Law of Small Numbers. *Psychological Review*, 76, 105-110.
- Tversky, A. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Tymoczko, T. (1979). The Four Color Problem and its Philosophical Significance. *The Journal of Philosophy*. Volume LXXVI, no. 2, February.
- Tymoczko, T. (1985). *New Directions in the Philosophy of Mathematics*. Boston: Birkhauser.
- Vitale, B. (1992). Processes: A Dynamical Integration of Computer Science into Mathematical Education. In Hoyles, C. & Noss, R. (Eds.) *Learning Mathematics and LOGO*. London: MIT Press.
- Von Glaserfeld, E. (Ed.) (1991). *Radical Constructivism in Mathematics Education*. Netherlands: Kluwer Academic Press.
- Vygotsky, L. (1986). *Thought and Language*. MIT Press, Cambridge, MA.
- Wilensky, U. (1989). *Fractional Division*. Unpublished paper. Cambridge, MA: MIT Media Laboratory.
- Wilensky, U. (1991a). Abstract Meditations on the Concrete and Concrete Implications for Mathematics Education. In I. Harel & S. Papert (Eds.) *Constructionism*. Norwood N.J.: Ablex Publishing Corp. (Chapter 10)
- Wilensky, U. (1991b). Feedback, Information and Domain Construction. Unpublished paper. MIT Media Lab.
- Wilensky, U. (in preparation). Conservation of Worth: Agent Analyses of the Monty Hall Problem. MIT Media Lab.
- Wittgenstein, L. (1956). *Remarks on the Foundations of Mathematics*. New York: Macmillan.
- Wolfram, S. (1990). *Mathematica: A System for Doing Mathematics by Computer*. Redwood City, CA: Addison-Wesley Publishing Co.