POST-SECONDARY SCIENCE STUDENTS'
CONCEPTIONS OF RANDOMNESS AND ENTROPY

by

JOHN J. WEBER III

(Under the Direction of Denise S. Mewborn)

ABSTRACT

Thermodynamics and entropy are important aspects of secondary physical science. The molecular approach to thermodynamics provides insight into how the random motion of molecules results in regular properties of substances. The purpose of this research was to understand post-secondary science students' conceptions of randomness and entropy and how they integrated these concepts.

The Probabilistic Thinking Framework of Jones, Langrall, Mooney and Thornton and the Attribution of Randomness Framework of Metz were used as a theoretical framework to construct how the participants used the concepts of probability and randomness in their discussion of entropy.

Case studies focused on eight post-secondary science students while enrolled in a science education methods course. Four semi-structured interviews and a questionnaire informed the case studies. Qualitative analyses of the data used the constant comparative method.

Three key findings resulted from the analysis of the data. First, a new method of measuring the complexity of two-dimensional grids was proposed. This new measure expands on the measure proposed by Klinger and Salingaros by including specific characteristics of the grids identified by the participants. Second, the participants exhibited subjective or transitional conceptions of probability and sample space. The findings demonstrated that they did not make any connections between probability and sample space with randomness. Third, the participants had an incomplete or instrumental knowledge of entropy. They did not make any connections between the underlying molecular properties and the measure of entropy.

Based on these findings, it has been suggested that secondary science teacher education programs include statistics and entropy in the science education curriculum or as part of a science education methods course. It is important for science teacher educators to emphasize the connection between molecular behavior and macroscopic properties. Finally, it is important for textbook publishers to consider carefully how textbooks represent molecular systems.

INDEX WORDS: Preservice teachers, Secondary Mathematics Education, Secondary Science Education, Randomness, Complexity, Entropy

POST-SECONDARY SCIENCE STUDENTS' CONCEPTIONS OF RANDOMNESS AND ENTROPY

by

JOHN J. WEBER III

B. A., La Salle University, 1987
M. Ed., Loyola University of Chicago, 1993
M. A., De Paul University, 1996

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2009

POST-SECONDARY SCIENCE STUDENTS' CONCEPTIONS OF RANDOMNESS AND ENTROPY

by

JOHN J. WEBER III

Major Professor: Denise S. Mewborn

Committee: Lynn A. Bryan
Charles Kutal
James W. Wilson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2009

ACKNOWLEDGEMENTS

I want to express my appreciation for several people who have helped me throughout the many years of doctoral study and dissertation research and writing.

First, I want to thank my wife, Mary Kate, for the constant encouragement throughout the process and for the occasional push that I so desperately needed. I am so very grateful for your patience and for allowing me to teach part-time so that I could concentrate on completing my doctoral work. Now we have to plan that trip to Ireland.

I also want to thank my doctoral advisor, Denise Mewborn. Her continued support and encouragement helped me through the rough times and kept me moving toward the final product. She never abandoned me, despite many disappointments, pauses, and false starts. Likewise, I want to thank my wonderful doctoral committee that has provided sustained support throughout the years. The feedback provided by Lynn Bryan, Chuck Kutal, and Jim Wilson has been invaluable to me as a budding researcher.

Finally, I want to thank my family. To my sister, Maria: Thank you for always being available to listen to me when I needed someone with whom to talk. You always knew the right words of encouragement that I needed to hear. To the McCaffrey's: Thank you for understanding why I needed to spend time to work on my dissertation while I was supposed to spend time visiting with you during the holidays.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION TO PROBLEM

Random Processes in Secondary Science and Mathematics

My chemistry background and my experience teaching secondary chemistry provided a foundation for my interest in studying students' conceptions of randomness within the context of physical science (i.e., statistical mechanics and kinetic molecular theory [kmt]). In my eight years experience of teaching secondary chemistry, I noticed that many students have difficulty recognizing how the random motion of molecules results in regular behavior (e.g., constant temperature) and how the random motion helps explain everyday phenomena (e.g., the flow of odors from the kitchen). The students also have difficulty with the second law of thermodynamics, which states the entropy of the universe is increasing (Atkins, 1994, 1998; Reif, 1999) where entropy is a measure of the molecular arrangement (Dickerson, 1969; Gell-Mann, 1994; Reif, 1999; Styer, 2000).

My experiences agreed with Nakhleh and Samarapungavan's (1999) claim that students of all ages had difficulties in moving from a macroscopic worldview to a microscopic worldview (see also Lin, Cheng, & Lawrenz, 2000). One difficulty lies with students having problems making the connection between the distribution of the microstates of a system of molecules and the macroproperties of the system (see Wilensky, 1999a; Wilensky & Resnick, 1999). Ben-Naim (2007, 2008) provided an interesting argument claiming entropy and the second law of thermodynamics can be easily understood with the following three pieces of knowledge: how to count (i.e., finding the number of combinations), the particulate nature of matter, and the idea that atoms of the same element are indistinguishable from one another.

Finding the number of combinations is a mathematical concept found in statistics. Ben-Zvi and Garfield (2004b) argued that secondary students have difficulties with the mathematics that underlie statistics. Furthermore, secondary students may have problems with statistical reasoning (see delMas, 2004; Kahnemann & Tversky, 1982; Pfannkuch & Wild, 2004). Ben-Zvi and Garfield (2004b) provided some possible reasons for secondary students' difficulties in statistics: students do not have a great deal of preparation in statistical literacy (i.e., the basic skills needed to interpret data), statistical reasoning (i.e., how individuals interpret data) and statistical thinking (i.e., how the various statistical concepts fit together). Ben-Zvi and Garfield also suggested that secondary instruction focuses more on the procedural aspects than on conceptual understanding (see also Gal, 2004). Another reason may be that students are challenged by the complexity of statistical ideas (see Ben-Zvi & Garfield, 2004b) or that they hold misconceptions about concepts of probability and statistics (see Fischbein & Schnarch, 1997; Pfannkuch & Brown, 1996). Finally, the concept of randomness is difficult for students to understand in general (Falk & Konold, 1994), and some students have misconceptions about random phenomena (Batanero, Green, & Serrano, 1998; Falk & Konold, 1994; Metz, 1997, 1998).

Another challenge is the presentation of entropy and the second law of thermodynamics in secondary chemistry textbooks. In general, chemistry textbooks (e.g., Hill, 1992; Holtzclaw, Robinson, & Odom, 1991; Whitten, Davis, Peck, & Stanley, 2004) mention the importance of randomness within the context of kmt and entropy. Unfortunately, they typically discuss the connection of randomness with entropy in a static manner (see Rhodes, 1992). For example, Whitten et al., (2004) used the image shown in Figure 1 to show how the entropy of a system increases when two gases are mixed. However, the authors provided no explanation as to why the gases necessarily must mix and why the gases will never spontaneously separate. In other words, the authors did not provide any discussion of why spontaneous changes occur in the direction that leads to greater entropy. In addition, the textbooks did not provide any discussion of the quantitative aspects of entropy (see Whitten et al., 2004). As Batanero, Green and Serrano (1998) claimed "the meaning of randomness in … [mathematics] textbooks is not clear and unequivocal" (p. 113); it seems that science textbooks did not provide a clear explanation of entropy. In order to provide an unambiguous explanation of entropy, secondary science teachers must provide some discussion of the qualitative aspects of entropy while using quantitative ideas for support (see Ben-Naim, 2007, 2008).

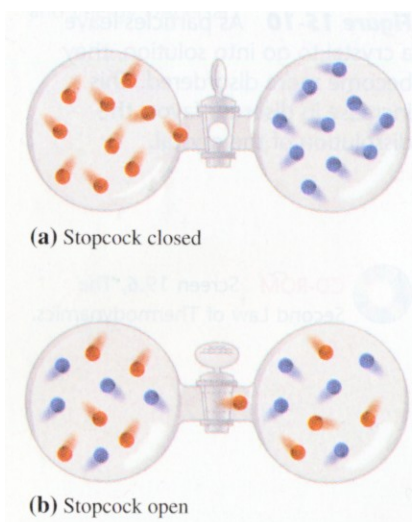


Figure 1. Textbook representation showing the increase in entropy when two gases mix (Whitten et al., 2004, p. 614).

Background

There are few, if any, studies on educational implications of teaching statistical mechanics at the secondary or university level (Canada, 2004, 2006; Resnick & Wilensky, 1998). How do post-secondary science students[1] eventually make the connections between statistics (e.g., combinations, probability, and randomness) and a given scientific phenomenon? In other words, how do students move from their current conceptions of matter to a scientifically valid explanation of a phenomenon that incorporates the ideas of statistical mechanics? The answers to these questions will allow me to determine how different conceptions of randomness interact with statistical mechanics, probability, and kmt within the SPSTs' explanations of complex systems. This will aid science educators in developing effective teaching strategies on the nature and properties of gases. This study will inform secondary science teacher education by studying how SPSTs use molecular reasoning and randomness to construct a working model (i.e., capable of prediction and explanation) of systems of gas molecules (and other complex systems). Lastly, there is the issue of the static two-dimensional diagrams that authors use to represent systems of molecules in textbooks. There is nothing inherently wrong with using these drawings; however, authors need to consider carefully how they represent supposedly "random" systems.

Traditionally, students learn thermodynamics without considering the molecular level (Dickerson, 1969; Reif, 1999; Smith, 1993). There are shortcomings to this method of learning thermodynamics. Reif claimed, "the macroscopic approach is not easy for students because it is rather abstract and provides no readily visualizable mental models" (1999, p. 1051). However, statistical mechanics provides an explanation for the results in thermodynamics (Dickerson,

---

[1] The participants were secondary preservice science teachers (SPSTs) selected from a science education methods course.

1969; Reif, 1999; Smith, 1993). Reif claimed the molecular approach to thermodynamics is better because "it introduces some basic notions of probability that are widely relevant in statistics" (1999, p. 1051) and are applicable in many domains. Some of the basic notions of probability are finding the combinations, calculating probability, and comparing probabilities (see Ben-Naim, 2007, 2008).

Specifically, there is a relationship among probability, randomness, and entropy. Each molecular system has a very large number of different microstates of a closed system (i.e., a system in which there is no exchange of matter or energy with the surroundings). Each microstate has a large number of configurations (i.e., the locations of the specific molecules in the energy levels of the system). Each configuration has the same probability of being observed. Even so, the system will likely be observed in an equilibrium state corresponding to the most probable microstate. Entropy is a measure of the number of configurations of a microstate. The most probable microstate has an extremely large number of configurations so that it dominates all other microstates. In other words, the system will randomly change configurations due to the many collisions of the molecules in the system; however, the system will continue to be observed in the most probable microstate (and appearing to be in equilibrium in the macrostate). Prior to equilibrium, spontaneous changes of systems occur in the direction of higher entropy. This is because one microstate has an extremely large number of configurations and thus has a higher probability of being occupied by the system. The larger number of configurations corresponds to a higher entropy. Once the most probable microstate is occupied, the system will randomly move among the various configurations of the miscrostate, but the properties of the macrostate (determined by the miscrostate) will look constant (and entropy will no longer change).

There have been several suggestions for how to include and/or improve the teaching of statistical mechanics and kmt in university-level introductory physics class (Fuchs, 1987; Lee, 2001; Moore & Schroeder, 1997; Reif; 1999; Wilson, 1981). Reif (1999) claimed there are four advantages to teaching thermodynamics from the atomic perspective: first, this approach "show[s] how [thermodynamics and atomic structure] are complimentary" (p. 1051); second, this approach "is more in tune with the kinds of thinking prevalent in contemporary physics and other fields" (p. 1051); third, "the approach is more interesting to many students … and [it] build[s] on the knowledge of mechanics that students have previously [studied]" (p. 1051); and last, "the approach introduces ideas that prepare students to deal with future courses in thermodynamics, statistical mechanics, physical chemistry and other fields" (p. 1051). Chabay and Sherwood claimed the molecular approach "allow[s] students to gain a sense of mechanism which can help make sense of thermal phenomenon" (1999, p. 1049; see also Reif, 1999). Finally, Fuchs (1987), Lee (2001), Moore and Schroeder (1997), Reif (1999), Smith (1993), and Prentis and Zainiev (1999) provided anecdotal evidence to suggest that the molecular approach to statistical mechanics is helpful to first-year college physics students. Lee stated that "it is not so difficult to teach the statistical foundations of thermal physics in introductory college physics courses if *we know how to teach statistical concepts*" (2001, p. 75, emphasis mine). Wilensky has made the same arguments about the ability of secondary students to understand statistical mechanics with the appropriate methods and use of the particulate nature of matter (1999a; see also Wilensky, Hazzard, & Froemke, 1999). Furthermore, delMas argued that "concrete physical activities … help students develop an understanding of abstract concepts and reasoning" (delMas, 2004, pp. 91-2). Statistical mechanics is a prime example of the need for an activity to help individuals develop an understanding of the concept of entropy.

Statistics textbooks typically present probability using idealized spinners, fair dice, or marbles in an urn (e.g., Moore, 2001, 2004; Schaeffer, Watkins, Witmer, & Gnanadesikan, 2004; Watkins, Schaeffer, & Cobb, 2004). Even though most of the examples used in textbooks were unrealistic (Greer & Mukhopadhyay, 2005), the textbooks provided these situations as a starting point for classicist-type probability. All of these situations dealt with discrete distributions with sample spaces that were relatively easy to enumerate. With small sample spaces, students can use basic counting procedures to determine the total size of the sample space. This study addressed classicist-type probabilities by considering complex systems with extremely large sample spaces. When considering samples of gas that contain 10^{23} molecules, counting (without the aid of a computing device) is no longer an option. Terry claimed "[s]tudents must … realize that the regular and predictable behavior of systems, which involve large number of particles …, is underpinned by the random, unpredictable motions of these particles" (1995, p. 328). I wanted to determine if a qualitative and conceptual understanding of these large complex systems could be based on the knowledge of smaller systems. We need to determine if a student's interpretation of randomness changes when considering larger systems of molecules. Furthermore, we need to know more about how students interpret a phenomenon, how they make the connection between the microstructure and the macroproperties of matter, and how possible alternative conceptions affect their conceptual understanding. Lastly, Greer and Mukhopadhyay (2005) claimed there is a need to study the connection students make among direct physical experiences, models, and simulations.

Rationale for Study

The overarching reason for this study is that these concepts of randomness, statistics, and probability are important for describing natural scientific phenomena as well as other common phenomena that scientists and science students encounter daily. Gal (2004) also claimed, "people … have been shown to hold many misconceptions and discontinuities in understanding and reasoning about stochastic phenomena" (p. 61). Furthermore, secondary science students need to understand (i.e., conditional knowledge) how to use these concepts when explaining why a particular natural phenomenon occurs (California State Board of Education, 1998b; Georgia Department of Education, 2005; Illinois State Board of Education, 1997a; New York State Education Department, n.d.a.; Pennsylvania Department of Education, 2002b; Texas Education Agency, 1998). Thus, this study examined how SPSTs make connections between their notion of randomness and the properties of a sample of gas molecules.

Random phenomena were the area of interest for this research study. Two aspects that underlie random phenomenon are "the uncertainty and unpredictability of the single event, in conjunction with the patterns that emerge across a large number of repetitions of the event" (Metz, 1997, p. 224). Falk and Konold (1997) suggested these two aspects of randomness (i.e., randomness as a process and randomness as a pattern) were closely related. Random phenomena are processes in which "individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions" (Moore, 2001, p. 348). Moore also stated that "'random' … is not a synonym for 'haphazard' but a description of the kind of order that emerges only in the long run" (2001, p. 347). In a random process, each individual outcome occurs by chance; however, after observing the phenomenon over time regular behavior may develop. Moore stated, "probability … gives us a language to describe the long-term regularity of random behavior" (Moore, 2001, p. 348). Furthermore, the concept of randomness is essential in statistics (Metz, 1997; Moore, 1990, 2001; Venn, 2006/1866) and physical science (Ben-Naim, 2007, 2008; Venn, 2006/1866; see also Eigen & Winkler, 1993; Greer & Mukhopadhyay, 2005; Hill, 1992; Holtzclaw, Robinson, & Odom, 1991; Pratt, 2005; Whitten, Davis, Peck, & Stanley, 1998, 2004). Lastly, the word random occurs in our daily conversations (Batanero, Green, & Serrano, 1998, Ben-Naim, 2007; Le Coutre, Rovira, Le Coutre, & Poitevineau, 2006).

Connection to Physical Science

Some physical science concepts rely on probability and randomness including quantum mechanics (Ware, 2003/1997; Whitten et al., 2004), reaction kinetics (American Chemical Society [ACS], 2003; Russell, 1996; Tinnesand, 2003/1997), statistical mechanics (Atkins, 1994, 1998; Ware, 2003/1997; Wilson, 1981), kinetic molecular theory [kmt] (Ben-Naim, 2007; Poincare, 2004/1908), and entropy[2] (Ben-Naim, 2007, 2008; Eigen & Winkler, 1993; Styer, 2000; see also ACS, 2003; Atkins, 1994; Hill, 1992; Holtzclaw, Robinson, & Odom, 1991; Tinnesand, 2003/1997; Ware, 2003/1997; Whitten et al., 2004). A sample of gas molecules is an example of a complex system (i.e., random phenomenon) that requires the concepts of probability, statistics (especially combinations), and randomness for understanding. There are a large number of molecules in a sample of gas, approximately 10^{23} molecules[3], which move in a constant random motion[4]. A microstate is a particular configuration of molecular energies over the eligible energy levels (Ben-Naim, 2007, 2008; Kauzmann, 1966). Each macrostate (i.e., the

---

[2] Entropy is a measure of molecular arrangement (Dickerson, 1969; Gell-Mann, 1994; Reif, 1999; Styer, 2000). Boltzmann defined entropy as $S = k \ln W$ where *W* is the number of microstates of the system (i.e., number of combinations of molecules arranged over the various energy levels) and *k* is a constant (McDowell, 1999; Pauling, 1970). Furthermore, Styer (2000) suggests the 'entropy as a measure of disorder', as is used in many textbooks and some state standards (see California State Board of Education, 1998c, New York State Education Department, n.d.a.), may lead to misconceptions about entropy.

[3] In approximately 22.4 liters at a pressure of 1 atmosphere and a temperature of 298 kelvin.

[4] This random motion (Brown, 1827) is Brownian motion.

measurable properties of the system)[5] corresponds to a large number of microstates. Each of the microstates is equiprobable (Bauman, 1995; Ben-Naim, 2007, 2008; Lee, 2001; Moore & Schroeder, 1997; Reif, 1999; Riveros, 1995), i.e., the probability of any "individual" microstate is $1/W$ (Ben-Naim, 2007, 2008; Moore & Schroeder, 1997). The molecules exchange energy through elastic[6] collisions. The extremely large number of collisions among the molecules changes the molecular speeds and direction. As a result, the system frequently changes the individual microstate it occupies. More interestingly, even when a system appears to be in equilibrium, there is randomness on the molecular level[7] (Stieff & Wilensky, 2003; see also Ben-Naim, 2007, 2008; Eigen & Winkler, 1993). In other words, the system constantly moves among the various possible individual microstates (Bauman, 1995). Despite the high frequency of collisions and the resulting changes in the molecular speeds, the macrostate appears to reach an equilibrium. Of the many possible microstates, one microstate is most likely[8] (Ben-Naim, 2007, 2008; Reif, 1999) because the number of combinations for this microstate far exceeds the number of combinations of any other possible microstate. The odds favor the system occupying one of the many individual combinations of the more likely microstate. Thus, the macrostate appears to be stable. In summary, despite the random motion of the molecules, a sample of gas has very ordered and well-understood macroproperties (e.g., thermodynamic observables[9]).

---

[5] Let me provide an analogy. Suppose I flipped six fair coins. Any specific result of the coin flips corresponds to a microstate of the system. There are seven macrostates of the system: zero Hs, one H, …, six Hs. The macrostate with zero Hs has only one microstate: TTTTTT. The macrostate with one H has six 'individual' microstates (HTTTTT, THTTTT, TTHTTT, TTTHTT, TTTTHT and TTTTTH). Each microstate in the coin flip analogy is distinct and identifiable. In statistical mechanics, the molecules are indistinguishable; however, we need to consider them as distinguishable in order to find the number of microstates (Ben-Naim, 2007, 2008; Gell-Mann, 1994).

[6] Elastic collisions are collisions in which kinetic energy is conserved.

[7] This is called dynamic equilibrium.

[8] In the coin flip analogy, the macrostate with three Hs has the most microstates (20), i.e., combinations of 3 H and 3T, than any other macrostate. Keep in mind as the number of coin flips increases the number of 'individual' microstates of the most likely macrostate will dominate all other macrostates.

[9] Pressure (P), volume (V), temperature (T), and number of molecules (n) are examples of thermodynamic observables.

Connection to Chemistry

"Chemistry is the science of substances – their structure, their properties, and the reactions that change them into other substances" (Pauling, 1970, p.1). For secondary chemistry, the National Science Education Standards (NSES) suggested, "the relationship between properties of matter and its structure [is] a major component of study in 9-12 physical science" (National Research Council [NRC], 1996, p. 177). More specifically, the NRC argued the importance that "high school students develop the ability to relate the macroscopic properties of substances … to microscopic structure of substances" (NRC, 1996, p. 177). One important topic in secondary chemistry is the nature and behavior of gases (California State Board of Education, 1998b; Georgia Department of Education, 2005; Illinois State Board of Education, 1997a; Kendall & Marzano, 2004; New Jersey Department of Education, 1996a; New York State Education Department, n.d.a.; NRC, 1996; Pennsylvania Department of Education, 2002b; Texas Education Agency, 1998; Wilensky, 1999a)[10]. The nature and behavior of gases is an important topic of study for several reasons. First, because gases behave differently from liquids or solids (Pauling, 1970), gases are an essential topic in the secondary chemistry curriculum. Second, Lin, Cheng and Lawrenz (2000) claimed that "teachers and students hold similar alternative conceptions[11] of gases" (p. 238). One of the common misconceptions held by elementary students and college chemistry students is that gases behave similar to liquids (Benson, Wittrock, & Baur, 1993; Lin, Cheng, & Lawrenz, 2000). Third, kmt is an essential theory that provides an explanation for the properties of gases (Georgia Department of Education, 2005; Kendall & Marzano, 2004; NRC, 1996; Wilson, 1981). KMT is a scientific model that provides a

[10] In this proposal I refer to the standards from states with large populations of students (California, New York and Texas), states in which I have taught (Georgia and Illinois) and states in which I will seek teaching positions (Illinois, New Jersey and Pennsylvania).
[11] Gilbert and Swift define alternative conceptions as "meanings for words … which differ from the standard interpretations" (1985, p. 682; see also Taber, 1998).

connection between the microstates of a system of molecules with the macroproperties of the system (Kauzmann, 1966). Furthermore, scientific models are an important part of the secondary science curriculum (Bhushan & Rosenfeld, 1995; Justi & Gilbert, 1999; NRC, 1996; Schamp, 1990). Fourth, the study of the nature and behavior of gases uses concepts from all four branches of chemistry: quantum mechanics, thermodynamics, statistical mechanics, and reaction kinetics. Let me briefly describe some of these concepts[12].

Thermodynamics

Thermodynamics is the study of energy and energy transfer (Whitten, et al., 2004). Thermodynamics enables scientists to "*observe*, *measure*, and *predict* energy changes" (Whitten, et al., 2004, p. 584, italics in original) from physical and chemical changes. There are several important laws of thermodynamics. The first law of thermodynamics (law of conservation of energy) states that the total energy of the universe is constant[13] (Atkins, 1998). Energy cannot be created nor destroyed in ordinary chemical reactions but can only be transformed from one form (e.g., chemical, kinetic and electrical) to another. The second law of thermodynamics states that the entropy of the universe is increasing (Atkins, 1994, 1998; Reif, 1999). In other words, systems with higher entropy[14] are more probable. In other words, spontaneous chemical or physical changes typically occur in the direction of increased entropy. This study placed an emphasis on the second law. The third law of thermodynamics states that as the temperature of a perfectly ordered system tends toward absolute zero (0 Kelvin or −273° Celsius), the entropy of the system tends toward zero (Atkins, 1998). In general, as the amount of heat added to the system increases, the entropy increases.

[12] Quantum mechanics and reaction kinetics are not important for this study so I will not provide a description of them here.
[13] A more precise definition of the first law includes the concept of work, but this is beyond the scope of this study.
[14] Systems that are more probable are the macrostates with the most microstates (Ben-Naim, 2007, 2008).

Thermodynamics is fundamental to physical science (Linn & Songer, 1991) because it is easy to experiment with the variables that represent the thermodynamic observables (Linn & Songer, 1991; Thomsen, 1998; see also Ben-Naim, 2007, 2008). Thermodynamics provides a "set of relations between *macroscopic* properties that we can measure in the laboratory …" (Smith, 1982, p. 1, italics in original). For example, thermodynamics describes the relationships among temperature (*T*), entropy (*S*), enthalpy (*H*) and free energy (*G*) that help determine the spontaneity of a reaction between two substances. In addition, thermodynamics provides the mathematical relationships (called gas laws) among *P*, *V*, *T* and *n* of a sample of gas in equilibrium. Poincare suggested the gas laws are mathematically simple[15] because the "velocities of the gaseous molecules vary irregularly…by chance" (2004/1908, p. 43). From these mathematical relationships (and some simplifying assumptions regarding the nature and properties of gases), scientists use statistical concepts to explain the thermodynamic observables and other properties of a system in terms of the molecular properties (the domain of statistical mechanics).

<u>Statistical Mechanics</u>

The ideas of statistical mechanics "are particularly useful in chemistry …" (Kauzmann, 1967, p. 3). Statistical mechanics is a fundamental idea underlying many of the concepts in the physical sciences such as reaction rates, entropy, and kinetic molecular theory (Dickerson, 1969) and is beneficial in understanding thermodynamics (Dickerson, 1969; Kauzmann, 1967; Pathria, 2001/1972; Prentis & Zainiev, 1999; Reif, 1999; Smith, 1993). Moreover, it is possible to use statistical mechanics to derive thermodynamic properties of substances (Widom, 2002). In other words, statistical mechanics provides the scientist with information about macroproperties of

---

[15] For example, Boyle's law, one of the gas laws, is represented by the simple mathematical formula $PV = k$ where k is a constant.

gases (i.e., thermodynamic observables) based on the properties of its constituent molecules (Dickerson, 1969; Prentis & Zainiev, 1999). On the other hand, Lee claimed the purpose of statistical mechanics is “to relate Boltzmann’s distribution to thermodynamic functions in a simple and clear manner” (2001, p. 68).

Kinetic Molecular Theory

The kinetic theory of gases[16] is a model that describes the motion of molecules in a gas. Statistical mechanics and kmt “explain a broad range of phenomenon” (Chabay & Sherwood, 1999, p. 1049). KMT provides a link between the concepts of statistical mechanics and the concepts of thermodynamics (Dickerson, 1969). In other words, kmt is one of the first successful theories that used molecular properties to explain thermodynamic observables (Dickerson, 1969). This model[17] has several underlying assumptions:

1. The molecules of a gas have relatively large distances between them. In other words, the distance between molecules is so much larger than the size of the molecules that the molecules are considered point-masses and are assumed to have no volume.
2. The molecules are in constant, *random*, straight-line motion.
3. The temperature of a substance is a measure of the average kinetic energy. More specifically, temperature is directly proportional to average kinetic energy.
4. The collisions of molecules with other molecules or with the walls of the container are completely elastic. That is, there are no forces between two molecules of a gas so that there is no loss of energy when two molecules collide.

[16] I will use the term kinetic molecular theory (kmt).
[17] An ideal gas obeys all five assumptions. Although typical real gases do not follow all the assumptions of kmt, real gases at low concentration and low pressure and high temperature do behave similar to an ideal gas (Pauling, 1970).

5. The pressure of a gas is directly proportional to the number of collisions per unit area of the wall of the container; at the same temperature, the molecules of two different gases have the same average kinetic energy.

Although different introductory chemistry texts state these assumptions differently, the above assumptions are the basic ones that many agree upon (see Atkins, 1998; Bailar, Moeller, Kleinberg, Guss, Castellion, & Metz., 1978; Dickerson, 1969; Hill, 1992; Holtzclaw, et al., 1991; Pauling, 1970; Whitten, et al., 2004; Wilbraham, Staley, Matta, & Waterman, 2004; Zumdahl, 1989).

KMT is important to this study for four reasons. First, the second assumption explicitly uses the term random when describing the motion of the gas molecules. The random motion of the molecules provides explanations for the properties of gases (e.g., diffusion, effusion, and viscosity). The relationship between the random motion of gas molecules and the properties of gases is important for secondary science students in Georgia (Georgia Department of Education, 2005; see also California State Board of Education, 1998b). Second, several state science standards recommend kmt as an important topic in the secondary chemistry curriculum (e.g., California State Board of Education, 1998b; Georgia Department of Education, 2005; Illinois State Board of Education, 1997a; New Jersey Department of Education, 1996a; New York State Education Department, n.d.a.; Pennsylvania Department of Education, 2002b; Texas Education Agency, 1998). Third, some state science standards recommend entropy as another important topic in the secondary physics and chemistry curriculum (e.g., California State Board of Education, 1998c; Georgia Department of Education, 2005; New York State Education Department, n.d.a.; Pennsylvania Department of Education, 2002b), and some include the second law of thermodynamics (see California State Board of Education, 1998c; Georgia Department of

Education, 2005; New York State Education Department, n.d.a.). Fourth, students studying kmt can observe how effective models are in explaining natural phenomena. The connection between kmt and the properties of gases provides students an opportunity to view a scientific model and readily test the model with natural phenomena (Kendall & Marzano, 2004; Lesk, 1974; NRC, 1996; Rhodes, 1992). Finally, Nakhleh & Samarapungavan (1999) claimed, "an understanding of the kinetic molecular theory of matter forms the basis for much science learning in physics, chemistry, and biology" (p. 779). Nakhleh (1993) suggested that kmt "should be covered in enough depth to show students that the theory accounts for all of the important aspects of ideal gas behavior" (p. 16).

Connection to Mathematics

Combinatorics

Combinatorics is included in the mathematics curriculum (California State Board of Education, 1998a; Georgia Department of Education, 2005; Illinois State Board of Education, 1997b; New Jersey Department of Education, 1996b; New York State Education Department, 2005; Pennsylvania Department of Education, 2002a; Texas Education Agency, 2006) and is important for chemistry and physics (English, 2005). Combinatorics has an important role in the secondary curriculum because of its applicability to other courses, including chemistry and physics (English, 2005). In statistical mechanics, the number of microstates for any given macrostate is analogous to finding the number of ways (i.e., combination) N indistinguishable items can be placed into T positions, where each position either is filled or not filled (Dickerson, 1969; see also Ben-Naim, 2007, 2008; Moore & Schroeder, 1997; Pauli, 1973). The formula for calculating the number of combinations is $\frac{T!}{N!(T-N)!}$ [18] (Ben-Naim, 2007). Due to the large number of molecules in a mole of gas, there are a large number of microstates for any given macrostate.

Probability and Statistics

According to the National Council for Teachers of Mathematics [NCTM] (1989, 2000) and NRC (1996) standards, probability and statistics are important topics within secondary school mathematics and science (see also Stohl, 2005). Concepts of probability are important in areas outside of physical science and mathematics (Gal, 2004, 2005; Greer & Mukhopadhyay, 2005; Pratt, 2005). Gal (2004) claimed, "there is a need for adults to be familiar with the notion of randomness" (p. 61). Probability underlies various contexts that people experience daily including the state-run lottery (Albert, 2006; Kaplan & Kaplan, 2006), weather forecasts (Albert, 2006; Kaplan & Kaplan, 2006; Paulos, 1995; Pratt, 2005), health risks (Albert, 2006; Gal, 2004, 2005; Pratt, 2005), economics (see Wilensky, Hazzard, & Froemke, 1999) and opinion polls (Gal, 2005; Pfannkuch, 2005). Furthermore, the concept of randomness is essential in statistics (Metz, 1997; Moore, 1990, 2001; Venn, 2006/1866).

One of two approaches to probability, classicist and frequentist, underlie these phenomena. Metz (1998) defined the difference between these two approaches as follows:

> *Classicist* probabilities involve the derivation of probabilities from an analysis of the symmetries in chance-generating devices, such as coin tosses or spinners. *Frequentist* probabilities involve the frequency of a given outcome over an infinite number of

[18] This study considered Maxwell-Boltzmann statistics of classical particles. Furthermore, the particles under consideration will be indistinguishable. In other words, this study will not consider Fermi-Dirac statistics or Bose-Einstein statistics.

repetitions of the event, as approximated by relative frequencies that emerge across many repetitions. (p. 287, italics original)

The mathematical definition of classicist probability of event *E* with *t* equally possible outcomes of the event and *n* total possible outcomes is $P(E) = t/n$. The probability of winning in a lottery is an example of classicist probability in which the probability of winning is the ratio of favorable outcomes (i.e., the numbers on which you placed your money) to the total number of equipossible cases. On the other hand, weather forecasters use frequentist probability to determine the chance of a weather event. In other words, the probability of a weather event is the frequency of a particular event (for a given set of specific [measured to a particular precision] conditions) over the distribution (i.e., an ensemble) of calculated possible events (Kaplan & Kaplan, 2006). This study considered both approaches to probability.

Teacher Education

Gal (2004) suggested teaching statistics is in need of a change. Shaughnessy (1992) stated there are three barriers to improving the teaching of probability concepts: “(a) getting stochastics into the mainstream of mathematical science school curriculum … (b) enhancing teachers’ background and conceptions of probability and statistics, and (c) confronting students’ and teachers’ beliefs about probability and statistics” (p. 467). The NCTM standards have addressed Shaughnessy’s first barrier. Both NCTM Standards documents (1989, 2000) suggested that probability concepts should be a part of the K-12 curriculum. More specifically, the NCTM’s *Curriculum and Evaluations Standards* (NCTM, 1989) recommend that probability receive increased attention in mathematics classrooms. In addition, NCTM argued that secondary school “students should gain a deep understanding of the issues entailed in drawing conclusions in light of variability” (NCTM, 2000, p. 325). The *Principles and Standards for School*

Mathematics (NCTM, 2000) claimed that students' "understanding of statistics and probability could provide them with ways to think about a range of issues" (p. 288). Among these issues are the concepts within statistical mechanics and thermodynamics (including kmt and entropy).

The Standards for Science Teacher Preparation suggested secondary preservice physical science teachers (SPSTs) have core competency (required for licensure) in thermodynamics, an advanced competency (required for specialists in the field) in advanced concepts of thermodynamics, and supporting competency (for licensure in more than one field) in statistics (National Science Teachers Association [NSTA], 2003). NSTA also suggested SPSTs in the biological sciences have supporting competencies in physics (including thermodynamics) and supporting competencies in mathematics (including probability and statistics). Because SPSTs will eventually assist secondary students to learn the ideas of statistical mechanics, they need a deeper understanding of thermodynamics and statistics (NSTA, 2003). The National Association of Biology Teachers ([NABT], 2004) published a position paper that argued that pre-service programs in biological sciences contain physics (including first and second laws of thermodynamics) and mathematics (including probability and statistics). Furthermore, Lin, Cheng, and Lawrenz (2000) argued that teachers need a sound conceptual understanding of how the particulate nature of matter underlies the macroproperties of substances. More importantly, teachers need to be able to apply their knowledge of statistical mechanics and mathematics "in different situations" (Lin, Cheng, & Lawrenz, 2000, p. 235).

Purpose of Study

The purpose of this research was to describe how SPSTs understand (declarative, procedural, conceptual, instrumental and relational)[19] randomness and how they use probabilistic

[19] In this study, I will describe the participants' declarative, procedural, conceptual, instrumental ("knowing both what to do and why" Skemp, 1977, p. 258; see also NCTM, 2000) and relational understanding ("rules without

and statistical reasoning to generate explanations of the properties of gases (i.e., complex systems), especially entropy. In particular, this study focused on SPSTs' conceptions of the pattern aspect of randomness using two-dimensional grids (arrays) that have been used to represent molecular systems (see Atkins, 1994; Ben-Naim, 2007, 2008; Eigen & Winkler, 1993; Styer, 2000).

Statement of Problem

Although the concepts of statistical mechanics are difficult (Wilensky, 1999a), there have been several different arguments on how to make these concepts more accessible to university-level introductory physics classes (Fuchs, 1987; Lee, 2001; Moore & Schroeder, 1997; Reif, 1999; Smith, 1993). Chabay and Sherwood claimed the molecular approach "allow[s] students to gain a sense of mechanism which can help make sense of thermal phenomenon" (1999, p. 1049; see also Reif, 1999). According to Fuchs (1987), the challenge in presenting statistical mechanics lies with the definitions of heat and entropy (see also Ben-Naim, 2007, 2008). On the other hand, Lee suggested, "teaching statistical mechanics in college introductory physics courses is not at all difficult if we apply a proper method" (2001, p. 68). Lee (2001) proposes comparing statistical mechanics to casting a large number of dice (see also Ben-Naim, 2007, 2008; Eigen & Winkler, 1993).

Greer and Mukhopadhyay (2005) noticed a lack of correspondence between national mathematics standards and state mathematical standards. A disparity also exists not only between the national science standards and the state science standards but also among the various state science standards. Despite these differences, statistical mechanics is a very powerful set of ideas that explains the relationship between the properties of a gas and the properties of the

---

reasons" Skemp, 1977, p. 258) (Kvatinksky & Even, 2002; see also Bruning, Schraw, Norby & Ronning, 2004; Skemp, 1977). When I refer to understanding, I will refer to conceptual understanding (NCTM, 2000).

constituent molecules. Given the importance of statistical mechanics and the requirements of thermodynamics and statistical mechanics in the secondary physical science teacher education program at the university where this study was conducted, to what extent can SPSTs in physical science understand statistical mechanics?

Research Questions

Specifically, the research questions for this study are:

1. How do SPSTs interpret randomness in two-dimensions (i.e., randomness as a pattern)? In other words, how do SPSTs determine whether a figure showing the molecules of one or two gases exhibits randomness? More specifically, the data from the SPSTs' interpretations of randomness will support the development of a measure describing the complexity of two-dimensional representations of molecules.
2. How do SPSTs make sense of entropy (i.e., randomness as a process)?
3. How do SPSTs make the connection between sample space (i.e., combinations), randomness, and entropy?
4. What are the salient features of the complex physical phenomena that SPSTs in chemistry/physics use to determine the definition of randomness needed to conceptualize random phenomena?

CHAPTER 2

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Literature Review

The standards published by the National Council for Teachers of Mathematics [NCTM] (NCTM, 1989, 2000) stress the importance of probability in the secondary mathematics curriculum. Furthermore, several researchers have claimed the importance of context, especially real-world situations, in developing statistical reasoning (delMas, 2004; Greer & Mukhopadhyay, 2005; Pfannkuch & Wild, 2004; Wilensky & Resnick, 1999). Statistical mechanics and kmt are contexts in which both mathematics and physical science are intertwined (see Resnick & Wilensky, 1998). Mathematics (i.e., probability, statistics and the concept of randomness) are important to understand statistical mechanics. On the other hand, statistical mechanics is a context to reinforce mathematical concepts. Thus, for those interested in becoming a secondary science teacher, this connection between mathematics and statistical mechanics is extremely important. In this section, I will discuss research on students' probabilistic reasoning, statistical reasoning, combinatorial reasoning, conceptions of randomness and understanding statistical mechanics that are relevant to this study.

Probabilistic Reasoning

Summary of Existing Research

Although there has been research on the probabilistic reasoning of elementary students (e.g., Falk, Falk, & Levin, 1980; Piaget & Inhelder, 1975; see also Langrall & Mooney, 2005), middle school students (see Watson, 2005) and secondary students (see Batanero & Sanchez, 2005), the research that is more appropriate to this study regards the probabilistic reasoning of university students, secondary teachers and SPSTs. Secondary mathematics and science teachers may be unprepared or under prepared for teaching statistics (see Pfannkuch & Wild, 2004; Stohl, 2005). Teachers may have poor understanding of probability or hold some misconceptions about probability concepts (Jolliffe, 2005; see also Stohl, 2005).

There are some important research results on secondary students that may provide some insights to the probabilistic reasoning of the participants in this study. Munisamy and Doraisamy (1998) studied the probabilistic reasoning of secondary students. The key findings were that secondary students "had procedural knowledge [of probability], but lacked the conceptual knowledge [of probability]" (p. 44); "probabilistic reasoning is not an easily acquired skill for most [students]" (p. 40; see also Konold, 1991); "[students] understanding of probability is far from satisfactory" (p. 45) and "probability concepts are unlikely to develop either incidentally or through maturation" (p. 43; see also Konold, 1991).

Batanero and Sanchez (2005) list several probability misconceptions commonly held by secondary students including: representativeness, i.e., short runs show appropriate proportion of outcomes (Tversky & Kahnemann, 1974; see also Batanero & Sanchez, 2005; Fischbein & Schnarch, 1997; Fischbein, Sainati Nello, & Sciolis Marino, 1991); equiprobability bias, i.e., if there were two possible outcomes, then each occur 50% of the time (see also Fischbein, Sainati Nello, & Sciolis Marino, 1991); the "outcome approach" to interpret frequentist probabilities, and not identifying common mathematical structures in different situations (pp. 248-9). Fischbein & Schnarch (1997) have identified additional misconceptions held by students of various ages including: negative recency effects, i.e., the gambler's fallacy in which the probability of the next outcome changes depending on the recent outcomes (see also Batanero &

Sanchez, 2005; Fischbein, Sainati Nello, & Sciolis Marino, 1991); positive recency effects, i.e., the tendency to think the next outcome will be the most common one in the last set of outcomes; confusion of compound and simple events, i.e., rolling two 5s have the same probability as rolling a 5 and a 6 on fair die (see also Fischbein, Sainati Nello, & Sciolis Marino, 1991); availability, i.e., individuals base probability on personal experience as opposed to random sample of population (see also Fischbein, Sainati Nello, & Sciolis Marino, 1991); and the effect of sample size. Results from the cross-age study conducted by Fischbein and Schnarch (1997) has shown that representativeness, negative and positive recency effects appear to decrease with age; confusion between compound and simple events appears to be frequent and stable across ages; and the availability misconception appears to increase with age. Recent studies have expanded upon these results. Metz (1998) considered both the classicist and frequentist approaches of objectivist probability. Metz studied 36 participants (twelve kindergartners, twelve third-graders, and twelve undergraduates). One of the key findings was "children and adults frequently fail to integrate uncertainty with patterns" (Metz, 1998, p. 349). Integrating uncertainty and randomness with patterns is important when science students consider visual representations of molecules (see Figure 3). Another finding was that all of the participants failed to "consistently evoke randomness where appropriate" (Metz, 1998, p. 350) and that the participants tend to "over attribute determinism" and "under attribute chance" (Metz, 1998, p. 350). Metz (1998) argued that the "findings reflect a complex and largely unknown interaction of experiential, instructional and developmental factors" (p. 351).

Gaps and Potential for Future Research

Most of the research has been centered on discrete probability based on spinners (e.g., Falk, Falk, & Levin, 1980; Green, 1989; Jones, Langrall, Thornton, & Mogill, 1997, 1999; Metz, 1997, 1998; Piaget & Inhelder, 1975), compound probability (not needed for this study), coin flips (e.g., Falk & Konold, 1994; Gal, 2005; Green, 1991; Piaget & Inhelder, 1975; Rubel, 2006; Tarr, Stohl Lee, & Rider, 2006); casts of die (e.g., Albert, 2006; Tarr, Stohl Lee, & Rider, 2006); selecting marbles from urns using both classicist approach to probability (e.g., Jones, et al. 1997, 1999; Piaget & Inhelder, 1975) and frequentist approach probability (e.g., Green, 1989; Jones, et al. 1997, 1999). Casts of die seem to be appropriate (see Ben-Naim, 2007, 2008; Lee, 2001; Styer, 2000) in determining SPSTs' conceptions of randomness within statistical mechanics. Lastly, Stohl (2005) claimed, "there has been significantly less research on teachers' knowledge *of* probability and their knowledge *for* teaching probability" (p. 351; italics in original).

Statistical Reasoning

Ben-Zvi and Garfield (2004b) define statistical literacy, statistical reasoning, and statistical thinking. Statistical literacy "includes basic and important skills that may be used in understanding statistical information" (p. 7) such as organizing data using tables and charts, "understanding concepts, vocabulary, and symbols, and includes an understanding of probability as a measure of uncertainty" (p. 7). Statistical reasoning is how people "make sense of statistical information" (p. 7) including making interpretations of data, representing data and using statistical summaries of data. Statistical thinking "involves an understanding of why and how statistical investigations are conducted and the "big ideas" that underlie statistical investigations" (p. 7). Statistical thinking "includes an understanding of how models are used to simulate random phenomena" (p. 7). Ben-Zvi and Garfield (2004a) suggested statistics education emphasize the following big ideas in statistics: data, distribution, models, trend, variability, association, samples and sampling, and inference. Only the first three ideas and the concept of variability were appropriate to this study.

Summary of Existing Research

Jones, Langrall, Mooney and Thornton (2004) studied elementary and middle school students and concluded, "students' statistical reasoning from elementary through college is diverse and often idiosyncratic" (p. 113). delMas (2004) has provided a summary of research showing that students have difficulty with statistical reasoning. Research on statistical reasoning needs to include "the use of analogy, metaphor and imagery" (delMas, 2004, p. 91).

Gaps and Potential for Future Research

There are very few research studies on pre-service teachers' statistical reasoning (Canada, 2006). delMas (2004) claimed, "one of the most neglected areas is research devoted to understanding students' statistical reasoning" (p. 92). delMas (2004) claimed "there is very little consensus on what is involved in statistical reasoning and … research on statistical reasoning is still in a state of development (p. 85). Finally, it is important to determine how statistical reasoning affects SPSTs' conception of statistical mechanics, thermodynamics and kmt.

Combinatorial Reasoning

Combinatorics and sample space is a very important concept in the mathematics curriculum that underlies probability (Batanero, Henry, & Parzysz, 2005; Batanero & Sanchez, 2005; English, 2005; Langrall & Mooney, 2005). Combinatorial reasoning is also important in statistics (Batanero & Sanchez, 2005; English, 2005; Langrall & Mooney, 2005), for the concept of randomness (Batanero & Sanchez, 2005; Greer & Mukhopadhyay, 2005; Piaget & Inhelder, 1975) and for the concept of entropy (Ben-Naim, 2007, 2008; Styer, 2000).

Summary of Existing Research

According to English (2005), individuals used several different procedures to find the number of combinations. These procedures included logical procedures (e.g., systematic enumeration), graphical procedures (e.g., tree diagrams), numerical procedures (e.g., combinatorial and factorial numbers), tabular procedures (e.g., tables and arrays), and algebraic procedures (e.g., generating functions). Batanero and Sanchez discuss three models of combinatorial problems: the selection model in which *n* objects are selected from *m* distinct objects, the distribution model in which *n* objects are placed into *m* cells and the partition model in which *n* objects are separated into *m* subsets (2005; see also Batanero, Godino, & Navarro-Pelayo, 1997; English, 2005). Batanero and Sanchez (2005) list four different sampling procedures used within the selection model: with replacement and ordered, $AR_{m,n}$, with replacement and without order, $CR_{m,n}$, without replacement and with order, $A_{m,n}$, and without replacement and without order, $C_{m,n}$ (see also English, 2005).

Batanero, Godino, & Navarro-Pelayo (1997) suggested that many misconceptions in probability are caused by "a lack of combinatorial reasoning" (p. 242; see also Hawkins & Kapadia, 1984) in determining the sample space (see also Jones, et al., 1997, 1999). According to Horvath and Lehrer (1998) understanding sample space required three things: recognizing different possible ways to obtain an outcome; being able to systematically and completely generate the sample space (i.e., all possible outcomes); and being able to "map the sample space onto the distribution of outcomes" (p. 123).

Gaps and Potential for Future Research

There is "considerable debate about children's understanding of [sample space]" (Langrall & Mooney, 2005, p. 106). More importantly for this study, sample space has an important restriction: The participants need to consider the total energy of the system when finding the possible microstates of the system. There are no studies on how participants handle this type of restriction when computing combinations. Furthermore, most studies only considered very relatively small sample spaces. There is a need to determine how SPSTs conceptualize larger sample spaces, especially those that are needed when considering entropy.

Randomness

Le Coutre, Rovira, Le Coutre, and Poitevineau suggested the "concept of randomness is ambiguous and complex … and gives rise to various interpretations" (2006, p. 22). In a study of secondary-level students, college psychology and mathematics researchers with Ph.D.s, they concluded, "individuals hold a wide range of meanings for the concept of randomness" (Le Coutre et al., 2006, p. 30). They categorized the various conceptions into following types: probability implies randomness, no causality implies randomness, causality implies non-randomness, and probability implies non-randomness (Le Coutre et al., 2006). Only a small fraction of the participants held the latter conception of probability.

Falk and Konold (1994) suggested there is a "similarity between the concepts of randomness and complexity" (p. 9) and that "it might be possible to foster a more intuitive, yet mathematically sound, conception of randomness if it is introduced via the complexity interpretation" (p. 10). Thus, in the current study when the participants ranked the binary sequences, the binary grids and the tertiary grids, they considered the complexity of the grids. A more intuitive conception of randomness might be more advantageous to SPSTs due to extremely large sample spaces that are involved with entropy.

"Randomness involves two distinct ideas: process and pattern" (Falk & Konold, 1997). Brownian motion is randomness as a process. The distribution of physical particles in two- or three-dimensional space is randomness as pattern. Randomness as pattern includes mathematics such as cellular automata (Wolfram, 2002) and chaos theory. In this study, SPSTs will judge the randomness, i.e., complexity (Klinger & Salingaros, 2000), of two-dimensional grids with either two or three colors.

Konold and Pollatsek (2004) suggested data can be viewed either as a sample of a population or as the output of a random process. When viewing data as the output of a random process, the distribution is an "emergent entity" (Konold & Pollatsek, 2004, p. 170) from the underlying randomness of the data. More importantly, the arithmetic mean and root mean square are identifying features of the distribution that are not found in the individual pieces of the data. The "mean is a stable property of a variable system" (Konold & Pollatsek, 2004, p. 172). In this study, we are considering the Maxwell-Boltzmann distribution of speeds and the temperature (a measure of the mean energy). However, there are various interpretations of the mean of variable data: data reduction, typical value, signal in noise and fair share (Konold & Pollatsek, 2004). The data reduction interpretation occurs when a student replaces all the individual values with the average value. The typical value interpretation is similar to the concept of mode. The authors claimed this is the most common interpretation used by secondary students. In the signal in noise interpretation, "each observation is viewed as deviating from [some] actual [value] by some measurement error which is viewed as 'random'" (Konold & Pollatsek, 2004, p. 179). The remaining interpretation, fair share, is not valid for this study. It consisted of redistributing quantities evenly.

Summary of Existing Research

Most of the research on randomness was conducted using tilt box (Piaget & Inhelder, 1975; Metz, 1997, 1999), binary sequences of H and T representing coin flips (Batanero, Green, & Serrano, 1998; Batanero & Serrano, 1999; Falk & Konold, 1994; Flores, 2006; Green, 1989; Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993) and snowflakes/raindrops on a grid

(Batanero & Serrano, 1999; Falk & Konold, 1994; Green, 1989; Piaget & Inhelder, 1975; Paparistodemou, Noss, & Pratt, 2002).

Batanero and Serrano (1999) concluded, "age and instruction have little influence on students' conception of randomness" (p. 563). They suggested four categories students used to justify or reject randomness: regular pattern, irregular pattern, representativeness (i.e., frequency of possible results is similar), frequency of possible results is different, long runs (in binary sequences or in grids), no runs (in binary sequences or in grids), and unpredictability as randomness.

Falk, Falk, and Levin (1980) have shown that elementary students have difficulties with the deterministic and chance dichotomy. Pfannkuch and Brown (1996) claimed undergraduate students have the "tendency towards overly deterministic thinking" (n. p.; see also, Jacobson & Wilensky, 2006; Langrall & Mooney, 2005; Metz, 1999; Resnick & Wilensky, 1998; Wilensky, 1993, 1995). Adults (see Konold et al., 1993; Stohl, 2005) and professionals who depend on statistics (see Taleb, 2005, 2007) tend to think deterministically.

Students used several operational definitions of randomness (Falk & Konold, 1994; Metz, 1998; Langrall & Mooney, 2005; Moore, 1990). The constructs associated with randomness were unpredictability (Langrall & Mooney, 2005; Metz, 1998; Pratt, 2005), irregularity or patternless (Langrall & Mooney, 2005; Metz, 1998; Pratt, 2005), unsteerability (Langrall & Mooney, 2005; Pratt, 2005), uncertainty (Metz, 1998) and fairness (Langrall & Mooney, 2005; Pratt, 2005; Tarr, Stohl Lee, & Rider, 2006). Gal (2004) also suggested additional reasons for uncertainty: incomplete data regarding the situation (see also Metz, 1998; Polkinghorne, 1984), lack of familiarity with the background of the situation (see also Batanero, Henry, & Parzysz, 2005; Metz, 1998; Poincare, 2004/1908), beliefs and attitudes of the individual and personal experiences (see also Piaget & Inhelder, 1975; Batanero et al., 2005). Konold and Pollatsek (2004) suggested that unpredictability arises from large number of interactions (see also Batanero, Henry, Parzysz, 2005; Falk & Konold, 1994; Metz, 1998; Piaget & Inhelder, 1975; Poincare, 2004/1908). Lastly, Metz (1998) suggested randomness as chance as the definition that shows an understanding of randomness (see also Batanero et al., 2005). Poincare (2004/1908) suggested randomness must be more than ignorance.

Piaget and Inhelder's (1975) landmark study showed that there were different levels of probabilistic thinking in children. The researchers used a tilt-box to elicit children's understanding of chance (see Figure 2). At one end of the tilt-box were marbles of two colors separated by color. When the box was tilted, the marbles moved to the other side of the box and became mixed. The children were questioned about their conceptions of what occurred and asked to predict what would happen on additional tiltings of the box. The result of Piaget and Inhelder's experiments was a stage theory for students' conceptions of randomness. Stage 1 was typical of children of four to seven years of age. Children whose thinking was at this stage did not recognize the idea of chance; they continued to look for regularities within the mix of marbles. At this stage, the child also believed that the initial arrangement of marbles was preferred and thus believes in a return to this state. Lastly, the students at the first stage did not take into account the collisions of the marbles with each other or with the walls of the box. Stage 2 was typical of children of seven to eleven years of age. This stage was further divided into two stages. Children, whose thinking was in this stage, began to show conceptions of combinations and chance. The children recognized collisions but were not precise on the results of these collisions. Children at the beginning of this stage still believe in a return to the initial state, whereas children in the latter part of this stage do not believe in such a return. Stage 3 was typical of children eleven to twelve years old. Children's thinking within this stage includes the ideas of permutations and the children take into account collisions and the results of the collisions.

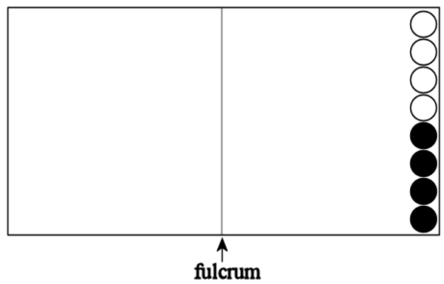


Figure 2. Drawing representation of Piaget's tiltbox experiment.

According to Piaget and Inhelder (1975), there were two necessary conditions for the development of probabilistic notions: The ability to determine combinatorial operations and the ability to construct a relationship between the individual cases and the sample space.

The tiltbox experiment is analogous to the entropy of a system (randomness as a pattern). For example, consider a system in which two gases are in different flasks connected by a stopcock. Once the stopcock is opened the gases will mix as the colored marbles in the tiltbox. The system will not appear to return to its initial state. The reason is the extremely large number of microstates of the mixed system compared to the two microstates of a system in which the gases are separated. Thus, the probability the system will return to its initial state is extremely small, as in the tiltbox experiment.

The idea of collisions is important KMT. The random motion of gas molecules is associated with the large number of collisions the molecules experience (Wilensky, 1999a). Students have difficulties in understanding the "statistical patterns arising from molecular collisions" (Resnick & Wilensky, 1998, p. 156). Furthermore, it is not only important to determine how students explain this random arrangement of marbles (or molecules or colored cells in a grid), but also to determine how students use the idea of random arrangement to model natural phenomena.

Falk and Konold (1994) studied students' conception of randomness using binary sequences and the snowflake task (see also Falk & Konold, 1997; Green, 1979, 1989, 1991; Piaget & Inhelder, 1975). Falk and Konold (1994, 1997, 1998) concluded that the various definitions of randomness were highly correlated. These definitions included copying difficulty, apparent randomness, memorization time, and the sequence's probability of alternation. Using the definition of a random sequence as one "*which cannot be condensed*" (Falk & Konold, 1994, p. 3, italics in original), there is no easy way to encode the sequence to make recall faster. In other words, subjects "relate complexity of a [binary sequence] to the difficulty to memorize, reproduce, or concisely encode it" (Falk & Konold, 1994, p. 2). Thus, Falk and Konold used the probability measure of alternations of symbols to define the degree of randomness of a binary sequence of length, n. The probability of alternations was defined as $P(A) = \frac{\#\text{changes in changes in symbol type}}{n-1}$ (Falk & Konold, 1994, 1997) where the numerator was the number of times the symbol changes in the sequence and n was the length of the sequence. According to Falk and Konold (1997), binary sequences with equal numbers of each symbol type were random when the frequency of each symbol was equal and $P(A)$ was close to 0.5. Furthermore, binary sequences with $P(A) = 0.5 \pm k$, for $0 < k < 0.5$ have the same amount of randomness. However, Falk and Konold (1994, 1997) found their participants typically considered $P(A) = 0.6$ as random. Furthermore, the participants considered $P(A) = 0.7$ (i.e., more alternations than expected) to be more random than $P(A) = 0.3$ (i.e., less alternations than expected). They conclude that “subjective complexity mediates the judgment of randomness” (Falk & Konold, 1994, p. 4; see also Feldman, 2004). In other words, an individual’s perception of complexity affects what they consider as random (Feldman, 2004). The implication (which is important for this study) was “students and teachers ought to know something about how we are already thinking of randomness …” (p. 10). Similarly, research conducted by Green (1991) showed students used too many short runs when generating binary sequences (see also Flores, 2006). Falk and Konold (1997) argued that pre-generated grids rather than student-generated grids provided more insights into students’ thinking. There were at least two reasons: first, the student may perceive randomness but may not be able to reproduce it; second, performance variables may confound the results (see also Falk, 1975; Joliffe, 2005).

Grids containing various symbols (or squares of various colors) are more appropriate to this study. Textbooks and instructors draw patterns of molecules for students (see Figure 3). Understanding the particulate nature of matter is very important in chemistry (Noh & Scharmann, 1997). A study conducted by Benson, Wittrock, and Baur (1993) showed that almost two-thirds of college students did not draw particles with a correct spatial distribution. Styer (2000) argued that the pictures used by Atkins (see Atkins, 1994) showed spatial misconceptions. Furthermore, Noh and Scharmann (1997) claimed college students have difficulties with chemistry questions represented with pictures. Therefore, relating the pattern aspect of randomness to spatial representations of molecules is appropriate.

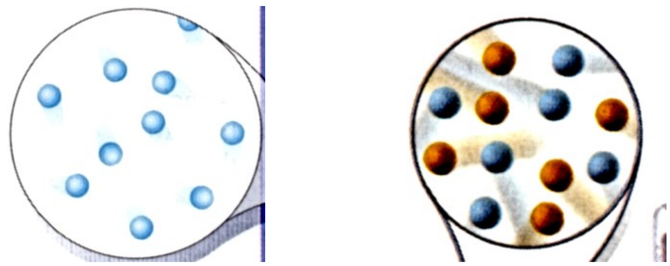


Figure 3. Textbook drawings of a sample of gas (Whitten, Davis, Peck, & Stanley, 2004, p. 458).

There are various definitions of randomness for grids. Falk and Konold (1997) suggested the following as a measure of randomness as a pattern in two-dimensions:

$$P(A) = \frac{\#\text{changes in color of cells in all rows and columns}}{2n(n-1)}$$

for $n \times n$ grids (Falk & Konold, 1997). Falk and Konold (1997) claimed the $P(A)$ for grids was consistent with various other measures of randomness and complexity in two-dimensions.

Klinger and Salingaros (2000) provided a more descriptive measure of the complexity of the symbols on a grid. Their randomness measure consisted of various calculations concerning the number of symbols and the symmetries of the patterns of symbols. The calculations are computed on different dimensions or scales of the grid. For example, consider the following 6 × 6 grid that consists of four disjoint 3×3 sub-grids and nine disjoint 2×2 sub-grids (shown in Figure 4).

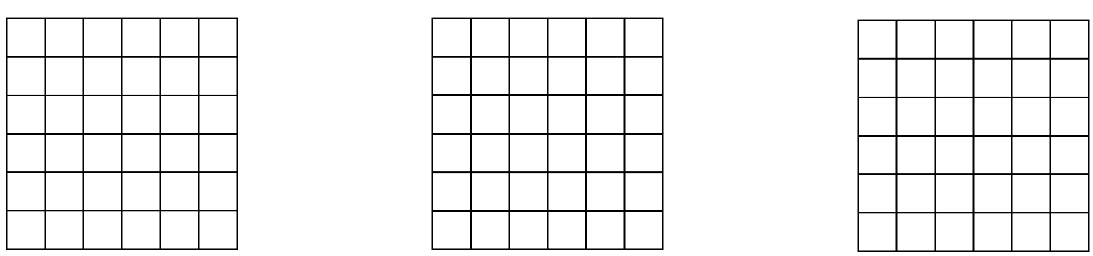


Figure 4. 6 × 6 grids showing the various disjoint subgrids.

The randomness measure proposed by Klinger and Salingaros (2000) was calculated using the mean number of symbols and the mean number of symmetries contained in the grid and subgrids (see Table 1). For example, $T(6{\times}6)$ is the number of different symbols in the 6×6 grid, $T(3{\times}3)$ is average number of symbols in the four disjoint 3×3 subgrids, and $T(2{\times}2)$ is the average number of symbols for the nine disjoint 2×2 subgrids. Finally, T is the average of $T(6{\times}6)$, $T(3{\times}3)$, and $T(2{\times}2)$. H is a measure of six different symmetries on the $n{\times}n$ grid and nine different symmetries on the disjoint subgrids. See Table 2 for a list and description of the symmetries used in the calculation of $H(n{\times}n)$. H is the average of $H(6{\times}6)$, $H(3{\times}3)$, and $H(2{\times}2)$.

Table 1. The measures on disjoint subgrids used in calculating the complexity of grids.

Measure	Formula	Definition
T		Average number of different symbols
H		Average measure of symmetry
H_{max}		Maximum sum of h_i for any [sub]grid
L	$L = TH$	Pattern Measure
C	$C = T(H_{max} - H)$	Randomness Measure

Falk and Konold (1997) suggested the most random grid occurs when $P(A) = 0.5$. However, as seen in Figure 3, gases have a low concentration (i.e., more empty space than space taken up by molecules). Thus, a grid with fewer black cells (corresponding to gas molecules) and more white cells (corresponding to empty space) is more appropriate to a study of molecular systems and entropy and, thus, more important to this study. Figure 5 shows Figure 3 with an overlaid 6×6 grid[20]. Figure 6 shows the black and white 6×6 grid corresponding to Figure 5. However, $P(A)$ is not appropriate for these types of grids. The various measures for the grid shown in Figure 6 are:

$P(A) = 0.38$	$T = 0.889$	$H_{max} = 9$	$H = 2.324$	$L = 2.066$	$C = 5.935$

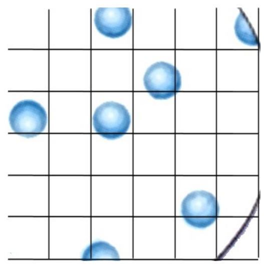


Figure 5. Textbook drawing of a sample of gas with an overlaid 6×6 grid (Whitten, Davis, Peck, & Stanley, 2004, p. 458).

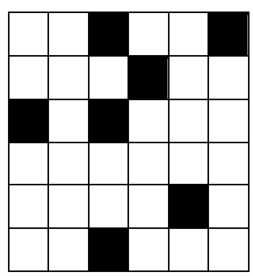


Figure 6. The black and white 6×6 grid corresponding to Figure 5.

The corresponding grid to Figure 3 is shown in Figure 7. The various measures for the grid shown in Figure 7 are:

$P(A) = 0.833$	$T = 2$	$H_{max} = 1$	$H = 0.5$	$L = 1$	$C = 1$

[20] I chose a 6×6 grid to compare it to other grids in this section of the research. In addition, I made the individual square size slightly larger than a 'molecule' and rotated the picture until most of the 'molecules' fit into a square.

The value of $P(A)$ for Figure 7 is further from 0.5 than $P(A)$ for Figure 6, which suggests Figure 7 is more patterned than Figure 6 (see Falk & Konold, 1994, 1997). In addition, the smaller value for the randomness measure suggests the same conclusion (see Klinger & Salingaros, 2000).

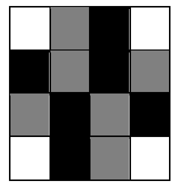


Figure 7. A 4×4 shaded grid corresponding to Figure 3.

Table 2. The symmetries used to calculate H in the randomness measure.

Symmetry measure	Symmetry
h_1	Reflectional symmetry over x-axis
h_2	Reflectional symmetry over y-axis
h_3	Reflectional symmetry over $y = x$
h_4	Reflectional symmetry over $y = -x$
h_5	±90° rotational symmetry
h_6	180° rotational symmetry
h_7	The existence of another $n \times n$ grid with same exact symmetries, for $n < 10$
h_8	After reflecting over x-axis or over y-axis and there is another $n \times n$ grid with same exact symmetries, for $n < 10$
h_9	After rotating ±90° or 180° and there is another $n \times n$ grid with same exact symmetries, for $n < 10$

Gaps and Potential for Future Research

However, $P(A)$ may be problematic when the ratio of colors is less than 1:1. Figure 8 shows some examples of 6×6 grids with 10 black cells[21]. Note grids III and IV have the same $P(A)$ but they differ in the randomness measure. An important question is whether these two grids are similar or different with respect to the participants' conceptions of complexity.

Grids with small number of black squares compared to white squares are more appropriate for this study since gases have low concentrations. Figure 9 shows two examples of 6×6 grids with six gray cells and six black cells. The two different colors correspond to the molecules of two different gases.

In addition, even though the tiltbox experiment is a complex system, it is not one of the naturally occurring complex physical phenomena studied by secondary students. The tiltbox experiment, selecting balls from an urn, spinners and coin flips (see Jones et al., 1999; Metz, 1998; Piaget & Inhelder, 1975) have provided insight into students' knowledge of randomness. What needs to be researched is whether individuals' understanding of these basic probability situations helps them understand more complicated probabilistic situations, like entropy.

[21] The black cells correspond to gas molecules (see Atkins, 1994; Styer, 2000).

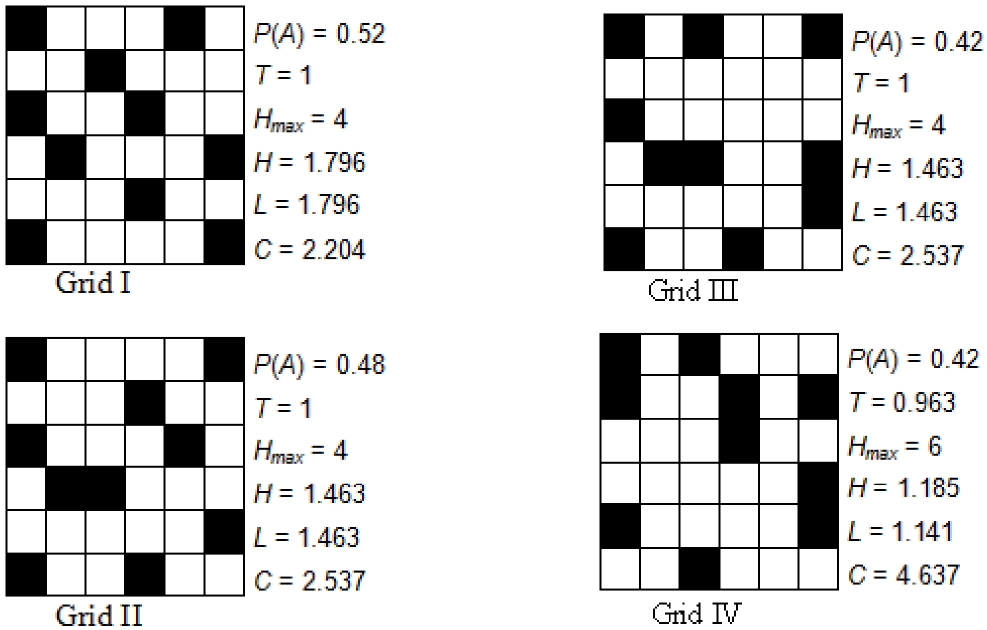


Figure 8. 6×6 grids with 10 black squares with various measures of complexity.

Resnick and Wilensky (1998) claimed there was “very little research in the developmental and cognitive psychology communities on how people make sense of complexity” (p. 170). Falk and Konold have expressed doubts about the “consistency and meaning of the findings” (1994, p. 5) regarding the randomness and probability thinking based on binary sequences. Flores (2006) suggested that teacher educators need to prepare teachers to teach about data and chance. Shaughnessy (1992) suggested research was needed “adolescents’ … notion of … chance, random events” (p. 489). This research fills at least two gaps. First, it addresses the conceptions of SPSTs who have not been studied. This study will help assess the conceptions of SPSTs so that any misconceptions can be addressed in the science education curriculum. Also, this study expands the notion of randomness/complexity to additional types of grids that are used in textbooks and drawn by college teachers.

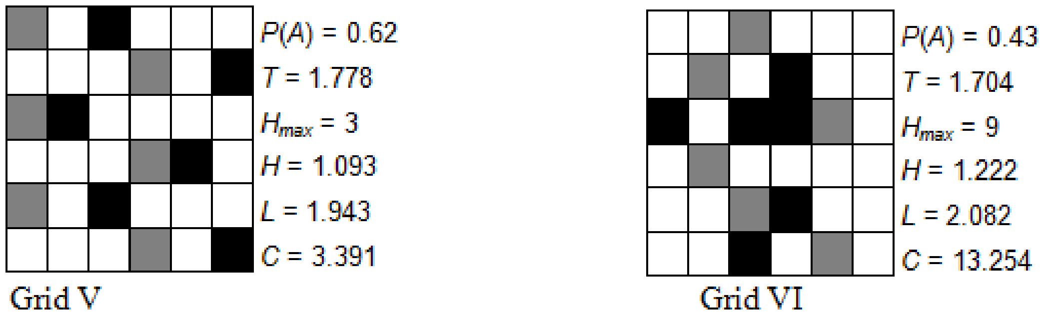


Figure 9. 6×6 grids with 12 shaded squares with various measures of complexity.

Statistical Mechanics, Thermodynamics and KMT

Summary of Existing Research

There have been studies investigating students' qualitative understandings of molecular properties (see Gabel, Samuel, & Hunn, 1987; Griffiths & Preston, 1992; Harrison & Treagust, 1996; Lee, Eichinger, Anderson, Berkheimer, & Blakeslee, 1993; Nakhleh & Samarapungavan, 1999) and thermodynamics (Styer, 2000); however, these studies did not consider the random motion of molecules. The random motion of molecules is an essential property of gases. It is this random motion that results in the need for statistical mechanics in order to explain the laws of thermodynamics (Dickerson, 1969) and thermodynamic observables (Kauzmann, 1966). For example, Styer (2000) anecdotally suggested that college physics students have problems with visualizing entropy. Benson, Wittrock and Baur (1993) claimed, "one-third of college students … tend to view the structure of gases similar to the structure of liquids" (p. 595).

Stieff and Wilensky (2003) claimed students have difficulties with the relationship between the properties of the gas molecules and thermodynamic properties on the macro-level (see also Wilensky, Hazzard, & Froemke, 1999; Wilensky & Resnick, 1999). Stieff and Wilensky (2003) claimed that if students were able to engage in "thinking in levels" (i.e., shift between the different levels), then college chemistry students' understanding of equilibrium and reaction kinetics improves (see also Wilensky, 1999a; Wilensky & Resnick, 1999). Wilensky (1999a) claimed that students investigating the properties of a gas using NetLogo (Wilensky, 1999b), an agent-based modeling software, while thinking in levels enables the student to cover "much of the territory of collegiate statistical mechanics and thermal physics" (n. p.; see also Wilensky, Hazzard, & Froemke, 1999). Furthermore, Wilensky (1999a) claimed the "traditional secondary chemistry curriculum segregates the micro-level and macro-level because the mathematics is out of reach of most secondary students" (n. p.).

Veal (2004) has shown that SPSTs consider chemistry as abstract since they cannot see atoms and molecules. Nussbaum (1985) summarized several studies which concluded that students from elementary school through college show evidence of difficulties with understanding the particulate nature of matter (see also Nussbaum, 1978). Veal (2004) concludes that the abstract view of chemistry hinders students in making the connection between atoms and molecules with thermodynamics. Novick and Nussbaum (1981) have shown students have problems with the intrinsic motion of molecules (see also Clough and Driver, 1985). Nakhleh and Mitchell (1993) concluded college chemistry students were more procedural than conceptual (see also Nakhleh, 1993). Furthermore, this thinking leads SPSTs to rely on traditional modes of teaching (e.g., lectures) (see Odgers, 2003).

Gaps and Potential for Future Research

Nakhleh (1993) and Nakhleh and Mitchell (1993) suggested further research was needed in students' learning and understanding of kmt and gas laws. Jacobson and Wilensky (2006) claimed, "there have been relatively few studies conducted regarding how students can learn complex systems concepts" (p. 15).

Several university-level physics professors have provided anecdotal evidence to the importance of molecular reasoning for understanding thermodynamics (Lee, 2001; Prentis & Zainiev, 1999) and entropy (Baierlein, 1994; Fuchs, 1987; Styer, 2000). Although it seems reasonable to expect molecular reasoning to assist in understanding thermodynamics, very little research supports this claim. The Center for Connected Learning and Computer-Based Modeling is conducting much of the research in student understanding of statistical mechanics.

The research on secondary students, secondary teachers, or SPSTs is lacking (Stieff & Wilensky, 2003). I have not found any research regarding SPSTs' understanding of statistical mechanics. Wilensky and his colleagues have performed a limited number of studies in this area. Their work mostly studies how the use of NetLogo (Wilensky, 1999b), an agent-based computer program, helps students understand various complex phenomena. Stieff and Wilensky (2003) studied undergraduate student understanding of the Maxwell-Boltzmann distribution. The purpose of the study was to determine if the students understood why the distribution is right-skewed. The authors claimed the students recognized that despite how they coded the molecular properties into NetLogo, the Maxwell-Boltzmann distribution will eventually emerge in their model (see also Wilensky, 1995; Wilensky, Hazzard, & Froemke, 1999). Wilensky and Resnick (1999) described a secondary mathematics and science teacher who used NetLogo to come to the same understanding as the students (see also Wilensky, 2001). However, these studies did not consider either of the following two important questions. First, using NetLogo, participants started the modeling of molecular systems at various non-equilibrium states. After running the model for a certain length of time, the system displayed the Maxwell-Boltzmann distribution. Therefore, once a system reaches the Maxwell-Boltzmann distribution why does the shape appear to remain invariable (i.e., at equilibrium)? In other words, why does the distribution not change back into its initial state or some other state (see Ben-Naim, 2007, 2008; Gell-Mann, 1994)?

Theoretical Framework

This study aims to reveal how SPSTs understand and interpret the concepts of randomness and entropy within a sample of gas. In particular, this study wants to determine how SPSTs apply statistical reasoning to the random motion of gas molecules and how they use the random motion of molecules to explain the properties of gases. In this section, I will discuss the theoretical framework that will guide this study. The theoretical framework will identify the level of participant understanding. This theoretical framework will guide the collection of the data and will be the foundation of my interpretation of the data. Specifically, four components need to be included within the theoretical framework for this study: 1. SPST's conception of probability and statistics; 2. SPST's combinatorial reasoning; 3. SPST's statistical reasoning; and 4. SPST's conception of randomness or chance.

Students' Conceptions of Probability

Tarr and Lannin (2005) list several theoretical frameworks for probabilistic reasoning. However, none of these addressed the probabilistic knowledge of post-secondary students. Thus, I will use the theoretical framework developed by Jones, Langrall, Thornton, and Mogill (1997, 1999) and expand the framework to include any necessary pieces needed to address the participants' understanding of statistical mechanics.

Framework for Probability Thinking

Jones et al. (1997, 1999) developed, tested, and revised a framework for probability thinking (see Table 3). This framework consisted of four constructs of probability: sample space, probability of an event, probability comparisons, and conditional probability. There were four levels of thinking for each of these constructs: subjective, transitional, informal qualitative, and numerical.

According to this framework, an individual understands sample space if he or she is able to determine “the complete set of outcomes” (Jones, et al., 1997, p. 104; see also Jones, et al., 1999) in a one-stage experiment (e.g., the sequences formed from the flip of one coin or the roll of one die) or a two-stage experiment. In the first stage (subjective), individuals do not construct the complete sample space for a simple event. Children at the remaining three levels construct the complete sample space for one-stage and two-stage events. The levels were dependent on the strategy used in constructing the sample space. In the second stage (transitional), students do not use a systematic procedure for constructing the complete sample space. In the third stage (informal qualitative), children to some extent use a generative strategy to construct the complete sample space. In the final stage (numerical), students completely use a generative strategy. This is appropriate for this study since SPSTs will determine the total number of microstates for a given macrostate. The generation of each microstate requires constructing a sample space of the various configurations.

The next construct in this framework was an “understanding of the probability of an event … [and] is exhibited by the ability to identify and justify which of the two or three events are most likely or least likely to occur” (Jones, et al., 1997, p. 105). Individuals exhibiting the subjective thinking predict the probability of an event based on subjective thinking. In the transitional level, children begin to base their predictions of the probability of an event on quantitative thinking but may still use subjective thinking. In the third level (informal qualitative), students justify their choice for probability using quantitative thinking. In the last level (numerical), students assign a numerical value to the probability of an event. This is appropriate for this study since SPSTs will determine the most probable microstate of a system based on their generated microstates.

Table 3. Framework for students' probabilistic thinking (Jones, et al., 1999, p. 489).

CONSTRUCT	Level 1 Subjective	Level 2 Transitional	Level 3 Informal Quantitative	Level 4 Numerical
SAMPLE SPACE	• lists an incomplete set of outcomes for a one-stage experiment	• lists a complete set of outcomes for a one-stage experiment *and* • sometimes lists a complete set of outcomes for a two-stage experiment using limited and unsystematic strategies	• consistently lists the outcomes of a two-stage experiment using a partially generative strategy	• adopts and applies a generative strategy that enables a complete listing of the outcomes for two- and three-stage cases
PROBABILITY OF AN EVENT	• predicts most/least likely event on the basis of subjective judgments • recognizes *certain* and *impossible* events	• predicts most/least likely event on the basis of quantitative judgments but may revert to subjective judgments	• predicts most/least likely events on the basis of quantitative judgments including situations involving non-contiguous outcomes • uses numbers informally to compare probabilities • distinguishes *certain*, *impossible*, and *possible* events, and justifies choice quantitatively	• predicts most/least likely events for single-stage experiments • assigns a numerical probability to an event (either a real probability or a form of odds)
PROBABILITY COMPARISONS	• compares the probability of an event in two different sample spaces, usually on the basis of various subjective or numeric judgments • cannot distinguish "fair" probability situations from "unfair" ones	• makes probability comparisons on the basis of quantitative judgments (may not quantify correctly and may have limitations where non-contiguous events are involved) • begins to distinguish "fair" probability situations from "unfair" ones	• makes probability comparisons on the basis of consistent quantitative judgments • justifies with valid quantitative reasoning, but may have limitations when non-contiguous events are involved • distinguishes "fair" and "unfair" probability generators on the basis of valid numerical reasoning	• assigns numerical probability measures and compares events • incorporates noncontiguous and contiguous outcomes in determining probabilities • assigns equal numerical probabilities to equally likely events
CONDITIONAL PROBABILITY	• following one trial of a one-stage experiment, does not give a complete list of outcomes even though a complete list was given prior to the first trial • recognizes when *certain* and *impossible* events arise in nonreplacement situations	• recognizes that the probabilities of *some* events change in a nonreplacement situation; however, recognition is incomplete and is usually restricted to events that have previously occurred	• can determine changing probability measures in a nonreplacement situation • recognizes that the probabilities of all events change in a nonreplacement situation	• assigns numerical probabilities in replacement and nonreplacement situations • distinguishes dependent and independent events

The third construct in this framework was "probability comparisons [which] is measured by [children's] ability to determine and justify: (a) which probability situation is more likely to generate the target event in a random draw; or (b) whether two probability situations offer the same chance for the target event" (Jones, et al., 1997, p. 105). In the subjective level, students compared the probability of an event based on subjective thinking and were unable to distinguish between a "fair" and an "unfair" game. At the transitional level, students base their probability comparisons quantitatively (though the quantities they use may not be correct) and begin to differentiate between "fair" and "unfair" games. In the informal quantitative stage, children based their probability comparisons using consistent and valid quantitative thinking. In the numerical stage, students compared probabilities by assigning and comparing a numerical value for each event. Furthermore, students in this level were able to assign the same numerical value to equally likely events. This is appropriate for this study since SPSTs will compare the probabilities of the various microstates of a system. The final construct in this framework was conditional probability. Jones et al. measured conditional probability by the children's "ability to recognize when the probability of an event is and is not changed by the occurrence of another event" (1997, p. 106).

The probability framework of Jones et al. (1997; 1999) considered only the classicist (discrete) approach to probability. This framework needs to be extended to include more than two or three possible outcomes and situations with non-enumerable sample spaces. Statistical mechanics and entropy (e.g., the large number of microstates for a system) are contexts in which this extension is important.

Framework for Teacher Subject-matter Knowledge and Understanding about Probability

Kvatinsky and Even proposed a "theoretical framework for teacher subject-matter knowledge and understanding about probability" (2002, p.1). The first part of the framework dealt with "two approaches … for handling questions about probability" (Kvatinsky & Even, 2002, p.2): The objective approach, a.k.a., frequentist approach to probability, and the subjective approach, i.e., the opinion made by the individual (Kvatinsky & Even, 2002). Another aspect of the framework dealt with the teachers' "familiarity with different representations and models" (Kvatinsky & Even, 2002, p.3) of probability and the ability to make connections among these models to form a more complete understanding of probability. A third aspect was teachers' familiarity with the frequentist vs. classicist approaches to probability and the ability to decide which is more appropriate for a given situation. A fourth aspect concerns the catalog of examples that are readily available to teachers to “illustrate important ideas, principles, properties, theorems” (Kvatinsky & Even, 2002, p.3). The next aspect concerns the “various forms of knowledge and understanding [including] conceptual, procedural, instrumental, relational, formal” (Kvatinsky & Even, 2002, p.4) and conditional. The final aspect concerns teachers’ “knowledge about the nature of mathematics” (Kvatinsky & Even, 2002, p.5). It is important to differentiate the SPSTs’ use of the different approaches to probability. Furthermore, it is important to determine how the various approaches may lead to misconceptions.

Students’ Combinatorial Reasoning

Batanero, Navarro-Pelayo and Godino (1997) argued, “a key point in assessing combinatorial reasoning is identifying the students’ difficulties in solving combinatorial problems” (p. 9). The authors identified nine difficulties experienced by students when solving combinatorial problems. Some of the difficulties include the non-systematic enumeration of the sample space (e.g., using trial-and-error) and the incorrect use of the tree diagram. Students may have difficulties with errors of order (i.e., confusing ordered and non-ordered problems) and errors of repetition (i.e., confusing replacement and non-replacement problems). Students may also confuse the type of object (i.e., confusing distinguishable and non-distinguishable objects) or the type of cells (i.e., confusing distinguishable and non-distinguishable cells or subsets). Finally, there may be a misunderstanding of the type of partition needed for a problem. There are two possibilities: first, students partition a set such that the union of all subsets (i.e., partitions) is not equal to the total set or, second, the students ignore some of the partitions. This theoretical framework compliments the frameworks of Jones et al. (1997, 1999) and Metz (1997, 1998).

Students' Statistical Reasoning

Garfield's (2002) general model of statistical reasoning was used to identify the participants' levels of statistical reasoning. The model consists of five levels of reasoning. The first level is idiosyncratic reasoning. An individual in this level knows the statistical terms but does not know the definition of the term nor how to apply the term to a particular situation. The next level is verbal reasoning. This level differs from the previous level in that individuals know the definition of the term. The third level is transitional reasoning. An individual in this level is able to identify two or more statistical terms relevant to a given situation, but is unable to recognize how the terms are related. The fourth level is procedural reasoning. This level differs from the previous level in that the individual knows how the statistical terms are related but is unable to explain why. The final level is integrated process reasoning. In this level, an individual "has a complete understanding of a statistical process, coordinates the rules and behavior" (2002, n. p.).

Students' Conceptions of Randomness and Chance

Whereas the probability framework only considered the classicist approach to probability, Metz (1997, 1998) provided a framework that considered both classicist and frequentist probabilities. More importantly, Metz (1997) developed an assessment protocol to analyze students' understanding and application of chance along three dimensions: (a) cognitive – the student's conceptual constructions; (b) epistemological – the student's beliefs about the place of chance and uncertainty in the world; (c) social – how students work together in developing their understanding of chance (p. 225). Within the cognitive dimension, students can recognize "patterns without uncertainty", "uncertainty without patterns" or "randomness." In the first case, students "exaggerate the information given" (Metz, 1997, p. 227) and is analogous to

Tversky & Kahneman’s Law of Small Numbers (1971). In the second case, students “conceptualize the situation as simply unpredictable” (Metz, 1997, p. 228). In other words, students fail to recognize any patterns in the long run (i.e., the Law of Large Numbers). The last case is the recognition or understanding of randomness. Metz claimed the “application of randomness demands an integration of the uncertainty and unpredictability of a single event, with an understanding of the patterns that can emerge across repetitions of events” (Metz, 1997, p. 229). This framework is appropriate for determining the SPSTs’ understanding of randomness using the grid-based complexity questions (see discussion for interviews three through six) and their understanding of the stability of the Maxwell-Boltzmann distribution (i.e., dynamic equilibrium).

Metz (1998) suggested there are various interpretations of random situations (see Figure 10). First, students will interpret the random event either outside the deterministic/indeterministic interpretive frame or inside the deterministic/indeterministic interpretive frame. I assume that the participants in this study will have the personal experiences to interpret all events within the deterministic/indeterministic interpretive frame. Inside this frame, there are two interpretations of random situations: deterministic and indeterministic. Finally, the latter consists of interpreting a situation as random (unpredictable but having patterns in the long run) or as a situation that has no patterns and is unpredictable.

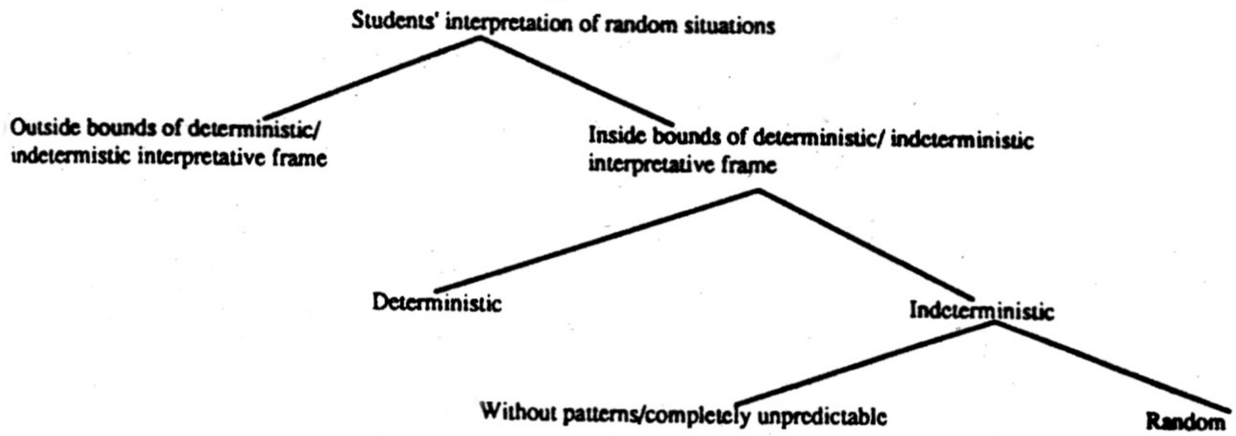


Figure 10. Conceptual framework for understanding and attribution of randomness (Metz, 1998, p. 295).

Within this framework, Metz provided a coding scheme. In other words, Metz (1998) lists several possible reasons students consider an event to be random. First, the individual may have previous experience with the random event without any additional analysis of the situation (Metz, 1998). Metz (1998) called this Data-Driven Reasoning. For example, the individual may be familiar with the randomness of selecting a card from a shuffled deck of playing cards. However, the individual does not try to determine why this event is random. Second, the underlying physical mechanism may be too complicated (i.e., there are too many variables of the system). Metz (1998) called this Indeterminate Physical Model. The tilt-box experiment (Piaget & Inhelder, 1975) was an example of an experiment that contains too many variables. Piaget and Inhelder discussed the concept of reversibility where the probability of a system returning to its original state was small. Classical physics claims that if a scientist could measure all the variables to any needed precision at any point of the experiment, then the scientist can predict the result with a certainty. Third, the individual is uninformed of the system; however, experts would be able to predict the outcomes of the system. Metz (1998) called this Internal Attribute of

Uncertainty. It is possible that a novice is unaware of an underlying physical mechanism, whereas the expert has experience with the mechanism and recognizes how the mechanism operates. Fourth, the underlying mechanism, if there is one, is unknown. In quantum mechanics, scientists do not know the underlying physical mechanism for the double slit experiment (see Christian von Baeyer, 1998; Polanyi, 1974; Stewart, 1989; von Plato, 1998). Fifth, Metz (1998) identifies probabilistic reasoning (Randomness) in which there was a "distant possibility of eventual return."

CHAPTER 3

RESEARCH METHODOLOGY

In this section, I describe the research design for this study including why this design was appropriate for this study. In addition, I describe how this research design helped address the research problem and research questions. More specifically, I discuss the detail of the research design. Furthermore, I address the issues of reliability and validity within this research design.

Overview

This qualitative study addressed SPSTs' understanding of the mathematical concept of randomness and the scientific concept of entropy. The description needed to have enough detail to establish the understanding of the participants. This study was well suited to qualitative design because the goal of describing SPSTs' knowledge requires in-depth interaction through interviews (Merriam, 1998). Furthermore, the "nature of the research questions" necessitated a qualitative design (Creswell, 1998, p. 17). In other words, in order to answer the research questions, I needed a rich, thick description (Creswell, 1998; Merriam, 1998; Stake, 1995) of what SPSTs' were thinking as they approached each task. Thus, through interviews, I collected detailed data about SPSTs' knowledge of relevant concepts in probability, statistics, combinatorics, kmt, and statistical mechanics. In addition, through analyzing the data, I attempted to explain how the participants' conceptions affected their conceptions of randomness and entropy.

Case Study

In particular, this study consisted of a collection of case studies. This study was appropriate for a case study approach because the aim was to identify SPSTs' understandings of mathematical and scientific concepts. Through a case study approach, I determined to what extent SPSTs knew the various concepts and how they made the connections among them. In other words, an in-depth probing of participants' thinking was required in order to determine how SPSTs used randomness to explain complex systems of gases and other random phenomena.

Participant Selection

Secondary mathematics and science teachers are an integral element in realization of the science standards (NSTA, 2003). In order to assist secondary students in understanding and using statistical mechanics, SPSTs need to know the relationships among these concepts and be able to provide appropriate explanations to their students (see NSTA, 2003). Thus, the participants in this study were SPSTs because teacher education programs prepare them to be future secondary science teachers.

I selected eight SPSTs in a secondary teacher education program at a large public institution in the South. I gained access through my professional connections with instructors at this institution. There were several requirements of the participants in order to qualify for the study: completion of a college-level introductory chemistry course, completion of a college-level introductory physics course, and completion of (or currently enrolled in) Modern Physical Chemistry I or Thermodynamics and Kinetic Theory. The above requirements greatly reduced the size of the population from which to draw my participants. Thus, during the participant recruitment, this requirement changed to include only introductory chemistry and introductory physics to allow the participation of preservice biology teachers. This does not seem problematic due to the reported number of biology courses that contain statistics concepts and the fact that NABT suggested that preservice secondary science teacher have exposure to probability and statistics as well as thermodynamics.

Role of Researcher

My role as researcher was limited to asking questions (Creswell, 1998; Holstein & Gubrium, 1995). I did not assist the participants in any manner during the tasks and interviews. However, I did provide suggestions by referring to earlier statements made by the participants.

Data

Merriam (1998) described a unit of data as "any meaningful (or potentially meaningful) segment of data" (p. 179). There were five sources for the data collection including transcriptions of four face-to-face interviews, survey responses, detailed field notes[22], a collection of artifacts (e.g., written work produced by the participants during the interviews), and memos to self that describe what I learned throughout the data collection and analysis (Creswell, 1998; Stake, 1995). All sources of data together facilitated a detailed description of each case (see Creswell, 1998).

Survey

The participants completed a written survey (see Table 4 and Appendix B). The survey covered background information regarding mathematics and science courses they completed on the university level and the topics these courses covered. In addition, one question asked how the participants combined kmt, scientific models, and randomness.

[22] These notes included any non-verbal communication of the participants and my reactions to participants' statements.

Interviews

I interviewed the participants individually to gain a deeper understanding of each SPST's conceptions of the various mathematical concepts underlying randomness and entropy. I carefully selected the individual interview questions so that SPSTs' understanding "can be studied in depth and [were] flexible enough to allow evidence of widely differing capabilities of the [participants]" (Goldin, 1998, p. 60).

There were various pieces involved in constructing the participants' knowledge of randomness and entropy. Each interview consisted of questions that addressed one or more of the following mathematical and scientific concepts that were cognitively appropriate for SPSTs: combinations, kmt, probability, and sample space (see Table 5). Accordingly, I met with the participants approximately once every two weeks for four semi-structured interviews lasting 45-60 minutes each (see Appendix A for tasks and sample interview protocols). Each interview started with a specific task and a list of possible questions to pose to the participants for each task. The semi-structured interviews consisted of "flexibly worded … [questions that] … allowed the researcher to respond to the … emerging worldview of the [participant]" (Merriam, 1998, p. 74; see also Creswell, 1998; Holstein & Gubrium, 1995). In other words, the semi-structured interviews allowed me not only to ask additional questions that probed or clarified participant's statements but also allowed the interview to move along a different path of questioning as the need arose (Holstein & Gubrium, 1995; Merriam, 1998). The initial questions for each task consisted of questions "at a level that all the [participants] were expected to understand … in differing ways" (Goldin, 1998, p. 60). The questions became progressively more challenging to most of the participants. I posed additional questions only when the participant had exhausted his or her investigation of the previous question or when the participant appeared to become uncomfortable with his or her knowledge of the question. During the interviews, I did not suggest the use of probability, statistics, or combinatorics; however, I encouraged their subsequent use by referring the participants back to their earlier statements. Through this manner of questioning, I was able to differentiate the differing meanings, understandings, and connections demonstrated by the participants (Creswell, 1998; Merriam, 1998; Kvatinsky & Even, 2002; Skemp, 1977). Some of the interviews took place during the Fall 2008 semester, and most interviews were conducted in the Spring 2009 semester.

Table 4. Survey of student educational background.

Question	Description	Topics
1	What college-level mathematics courses that you have taken included statistical concepts?[23]	Probability & statistics
2	What college-level mathematics education courses that you have taken included statistical concepts?	Probability & statistics
3	What college-level science courses that you have taken included statistical concepts?	Probability & statistics
4	What college-level science education courses that you have taken included statistical concepts?	Probability & statistics
5	What college-level science courses that you have taken included kinetic molecular theory, statistical mechanics and gas concepts?	kmt & randomness
6	What college-level science education courses that you have taken included kinetic molecular theory, statistical mechanics and gas concepts?	Model
7	Explain why you can smell food cooking from another room. How would you explain this to a secondary student?	Kmt, model & randomness

[23] I want to determine if students recall discussion of statistical concepts in the various courses.

The participants approached each task with whatever method(s) they deemed appropriate without any interference from me. The participants were encouraged “to talk aloud about what they [were] doing and to describe what they [were] thinking” (Goldin, 1998, p. 60). However, participants who were uncomfortable with this process were allowed to consider quietly the activities and to discuss their thoughts afterward. All participant responses were accepted without any judgment regarding the correctness of the answers. Additionally, various materials were provided to allow “for a wide variety of external representations” (Goldin, 1998, p. 60) appropriate for the given tasks including paper and pencil, various tokens of different colors, sets of dice, coins, and a calculator capable of computing combinations.

Table 5. Overall research design.

Interview	Topics	Data
1	Probability, sample space, randomness, kmt	Written & Verbal
2	Random patterns	Written & Verbal
3	Sample space, combinatorial reasoning, statistical and probabilistic reasoning, random process	Written & Verbal
Survey	Background information and kmt	Written
4	Sample space, combinatorial reasoning, statistical and probabilistic reasoning, random process and random patterns	Written & Verbal

At the beginning of the first interview, I advised the participants about their rights during the study (i.e., informed consent) including the purpose of the study, confidentiality, and their right to withdraw voluntarily from the study at any time (Creswell, 1998). The purpose of the first interview was to establish SPSTs' level of understanding of classicist probabilities (e.g., questions based on flipping coins and casting die) and sample space (i.e., questions based on casting die). Each of these questions was based on previous research (see Garfield, 2003, Jones et al., 1997, 1999; Metz, 1997, 1998; Watson, 2006). The levels of understanding of the participants were based on the theoretical frameworks by Jones et al. (1997, 1999) and Metz (1997, 1998).

In addition, in the first interview I attempted to establish SPSTs' conceptions of randomness in binary sequences (see Batanero, Green, & Serrano, 1998; Batanero & Serrano, 1999; Falk & Konold, 1994, 1997; Flores, 2006; Green, 1989; Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993). In this activity, the SPSTs ranked each set of five binary sequences from the least complex arrangement to the most complex configurations. The binary sequences were carefully constructed with various $P(A)$ values. The data collected from the binary sequence question established SPSTs' conception of randomness/complexity (see Falk, 1975; Falk & Konold, 1997; Metz, 1997, 1998).

In the second interview, I attempted to address the participants' conceptions of complexity and randomness as pattern. Randomness as pattern concerned the distribution of differently colored squares in two-dimensional space (see Attneave, 1959; Batanero & Serrano, 1999; Falk, 1975; Falk & Konold, 1994, 1997; Green, 1979, 1986; Klinger & Salingaros, 2000; Paparistodemou, Noss, & Pratt, 2002; Riveros, 1995). This two-dimensional extension to the distribution of coin flips more closely resembled the distribution of molecules in a gas (see Atkins, 1994; Ben-Naim, 2007, 2008; Eigen & Winkler, 1993; Gell-Mann, 1994; Wolfram, 2002). For all these interviews, the grids were carefully constructed with various C values. In this interview, SPSTs ranked six different series of five grids from the least complex to the most complex configurations. In the first two series, each 10×10 grid consisted of 50% white (W) cells and 50% black (B) cells. This is analogous to Piaget and Inhelder's (1975) raindrop activity and Green's (1979) snowflake task and was studied by Falk (1975) and Falk and Konold (1994, 1997). The second series was composed of the same grids as in series one, but with the B and W switched and presented in a different order. Series 3 and 4 consisted of five 10×10 grids with 84% white cells and 16% black cells. This configuration is more appropriate to the density of a gas (see Atkins, 1994; Styer, 2000). The last two series consisted of five 10×10 grids with 68% white cells, 16% black cells and 16% gray (G) cells. This configuration is analogous to the mixing of two gases. The sixth series was composed of the same grids as in series five but with each grid rotated 90º. The purpose of this series was to determine if the orientation affected the participants' perceptions of the complexity of the grids. After the participants ranked each series, they were asked for the rationale for their rankings.

Interview three consisted of determining the number of microstates for a given macrostate. This was essentially finding the number of ways to place x items into y boxes. However, the participants had to consider the total energy of the system as they placed the molecules into the energy levels. The purpose of this interview was to determine the SPSTs' combinatorial reasoning. In other words, I wanted to describe how the participants found each microstate and how they determined if they had found all possible microstates (see Metz, 1998; Jones et al., 1997, 1999). Furthermore, this interview allowed me to determine if the participants showed signs of any of the nine difficulties typically experienced by students in solving combinatorial problems (see Batanero, Godino, & Navarro-Pelayo, 1997) with small sample spaces. The microstate activity was also used to decide if the participants made the distinction between distinguishable and indistinguishable objects while they determined the number of combinations. Lastly, the participants were asked to determine the most probable microstate to determine if they treated each microstate as equally-likely and if they could relate this microstate to the dynamic equilibrium of a system.

The main topic for interview four was randomness as a process (see Ben-Naim, 2007, 2008; Eigen & Winkler, 1993). The participants were presented with an activity that models entropy using dice (see Ben-Naim, 2007, 2008; see also Styer, 2000). The materials consisted of a 4×4 board with columns labeled 1, 2, 3, 4 and rows labeled 1, 2, 3, 4 (see Appendix B); 16 two-colored tokens (i.e., red on one side and white on the other side); one green four-sided die labeled with integers 1, 2, 3, 4; one blue four-sided die labeled with integers 1, 2, 3, 4; and one six-sided die with three sides labeled with R and three sides labeled with W. Initially, the grid started with the tokens red-side up. The participant then rolled all three dice. The two tetrahedral dice determined the cell on the board, and the six-sided die determined the color of the token exposed in that cell. See Table 6 for coding schema used to analyze the data from this interview. The overall key questions for these two interviews were: What was the thing that changed at each step? How was this change achieved? Why was this change always in one direction toward equilibrium? (see Ben-Naim, 2007, p. 114; see also Ben-Naim, 2008). Furthermore, this activity allowed me to determine if the participants had trouble when the sample sizes dramatically increased. In sum, Table 7 shows which interviews address each research question.

Researcher Assumptions

The assumptions were identified through a pilot study of two recent university graduates in science and from my experience as a secondary mathematics and science teacher. The participants will demonstrate Level 4: Numerical Reasoning for the constructs of Sample Space, Probability of an Event, and Probability Comparisons (see Jones et al., 1997, 1999). More specifically, the participants will have few difficulties with combinatorial problems (see Batanero, Godino, & Navarro-Pelayo, 1997). In other words, the participants should be able to complete each task using a variety of methods (see Batanero, Godino, & Navarro-Pelayo, 1997). I assume the most common errors determining the number of combinations will be errors of omission (i.e., not listing a few possible combinations), double counting in non-ordered situations, and interpreting the replacement problem as a non-replacement problem (see Batanero, Godino, & Navarro-Pelayo, 1997). In addition, the participants will use sample space in their answers to questions involving probabilities.

On the other hand, participants will have difficulties relating the complexities of the grids to combinatorics (see Metz, 1998). In addition, they will not use the uncertainty with patterns (see Metz, 1999) definition of randomness. Furthermore, the participants will have difficulties connecting the concept of randomness (see Falk, 1975; Falk & Konold, 1994, 1997; Metz, 1998) to the concept of entropy. I assume the participants will use some form of the following terms in their definition of entropy: order, disorder, complex, spontaneity, laws of thermodynamics, gases. In addition, I assume they will use examples of gases and of spontaneous changes when discussing entropy. However, I assume they will not make any connection between entropy and sample spaces. Lastly, the participants may have difficulties with indistinguishability of atoms and molecules when determining sample space.

Data Analysis

The purpose of data analysis was to extract the emic perspective (participant's voice) within data collected (Creswell, 1998; Merriam, 1998; Stake, 1995). The data collection and data analysis proceeded simultaneously. The interviews were audio recorded for a detailed continuing analysis of all interviews. The audiotapes ensured that "everything said was preserved for analysis" (Merriam, 1998, p. 87). I transcribed each interview as soon as possible on the same day after each interview.

Table 6. Coding schema for Interview # 4 (see Metz, 1998).

Interpretation	Coding Criteria
Data-Driven Reasoning	• Predictions and explanations formed on the basis of prior outcomes the participant has observed in this situation or related task • No analysis of these data above and beyond correlations • No consideration of how these outcomes are generated
Order as the Natural State	• Natural state as ordered • Tendency of elements to return to ordered state expressed in teleological or animistic terms
Deterministic Physical Model	• Participant bases predictions on an analysis of the physics of the [physical model] • The physical model supports precise predictions • The absence of noise in the system or other source of chance enables the physical model to support precise predictions
Affordance of Inanimate Objects of Motion	• Chance attributed to assumption that the [dice], as inanimate objects, have no intentionality or internal controls
Internal Attribute of Uncertainty	• Indeterminacy stems from the participant's perception of personal ignorance of the system • Assumption that the system would be determined from the perspective of an expert
Indeterminate Physical Model	• Participant bases predictions on an analysis of the physics of the [physical model] • Decision that the physical model does not support precise predictions, due to noise in the system or imperfections of the [physical model]
Randomness	• Probabilistic reasoning concerning configuration outcomes • Some outcome more likely than others • Distant possibility of eventual return
Other Uncodable	• To be specified by coder

Qualitative analysis of data is making sense of the data (Stake, 1995). One way to make sense of data was to break it down into units and combine these units into categories or themes (Creswell, 1998; Merriam, 1998; Stake, 1995). I employed various strategies in analyzing the data. One strategy was "categorical aggregation" (Creswell, 1998, p. 153; see also Stake, 1995), which is a "collection of instances from the data" (Creswell, 1998, p. 154). For example, I located when and where the participants used the word "random." Another strategy was "direct interpretation … [in which I considered] a single instance and drew meaning from it without looking for multiple instances" (Creswell, 1998, p. 154; see also Stake, 1995). For example, I looked at each mention of the word random in context in order to determine the participant's definition of random. Direct interpretation enabled me to "[pull] the data apart and [put] them back together in more meaningful ways" (Creswell, 1998, p. 154). A third strategy was looking for patterns among the various categories (Creswell, 1998). This collapsed the categories into fewer patterns from which "naturalistic generalizations" (Creswell, 1998, p. 154) were developed. These generalizations concerned the understanding of the participants *in this research study* and were "compared and contrasted with published literature" (Creswell, 1998, p. 154).

The constant comparative method was another strategy used to construct categories (Merriam, 1998). This method of analyzing data consisted of starting with a particular episode from one of the data sources and comparing it with other episode(s). Merriam (1998) suggested, "[t]hese comparisons lead to tentative categories [or themes] that are then compared to each other and to other instances" (p. 159). The "[c]omparisons were constantly made within and between levels of conceptualization until a theory can be formulated" (Merriam, 1998, p. 159). For most interviews, I set aside one hour after each interview to annotate any notes and recorded any additional thoughts while the interview was fresh in my mind. I transcribed each interview and combined the transcription with other data (i.e., photographs and interviewer notes). After each set of interviews in the fall, I examined the data from the participants to find relevant themes within the data (Merriam, 1998). With each corresponding interview in the spring, I examined the additional data to find evidence confirming the categories and themes, looked for any additional categories and themes that appeared in the data, and searched for disconfirming evidence.

Table 7. Connection between research questions and interviews.

Research Question	Interview and Survey Questions	Corresponding Interviews
1	Random patterns	1, 2, 3, 4
2	Random process	1, 4
3	Combinatorial reasoning; randomness, complexity	1, 2, 3, 4
4	Combinatorial reasoning; random patterns; random process; statistical and probabilistic reasoning; sample space	1, 2, 3, 4

Verification of Interpretation

Throughout the collection of data, I monitored all aspects of data collection and analysis to ensure all conclusions were based on the data. There were three main sources of verifying the interpretations of qualitative research: internal validity, external validity, and reliability (Merriam, 1998). In this section, I discuss how each of these issues was addressed in this study.

Internal Validity

Internal validity refers to the credibility of the conclusions (Merriam, 1998). I used the following strategies to ensure internal validity: detailed data collection, triangulation, and member checks (Merriam, 1998). As previously mentioned the data were collected over approximately eight weeks and included multiple sources of data. This resulted in enough detailed data to support the research findings. The multiple sources of data allowed me to use triangulation to confirm the emerging findings (Stake, 1995). In other words, when I found similar themes from the different data sources, this contributed to the validity of the theme. Additionally, I used triangulation to find contradictory evidence (see Stake, 1995) to the emerging themes, which enabled me to refine the analysis. In addition, throughout the analysis process, I consistently investigated my subjectivity. In addition to keeping a fieldwork journal, I produced a journal of my thinking during the analysis of the data. These journals were "an introspective record of … [my] ideas, fears, mistakes, confusion and reactions to the experience" (Merriam, 1998, p. 110). During the ongoing analysis, I asked for the participants' feedback (i.e., member check) on my interpretations of the data (i.e., the categories and themes) to determine if they thought the results were plausible (Stake, 1995). Together, these various methods enhanced the validity of the results.

External Validity

External validity refers to generalizability of the results (Merriam, 1998). External validity is limited in qualitative research (Merriam, 1998), especially case study research. I did not attempt to make any generalizations from the data collected in this study. To enhance reader generalizability, I provided "enough description [i.e., a rich, thick description] so that readers will be able to determine how closely their situations match the research situation, and hence, whether findings can be transferred" (Merriam, 1998, p. 211).

Reliability

Reliability refers to the consistency of the data (Merriam, 1998). In other words, were the conclusions based on the data? I employed the following methods to enhance the reliability of the study: triangulation, audit trail and description of my assumptions (Merriam, 1998). The audit trail included a detailed description of "how data were collected, how categories were derived and how decisions were made throughout the [study]" (Merriam, 1998, p. 207). The audit trail enables readers to determine if the findings were based in the data (Merriam, 1998). Lastly, I provided detailed information on any assumptions made during the study, the theoretical frame in which the data collection and analysis were conducted, and the basis of the participant selection (Merriam, 1998). Employing these various methods enhanced the reliability of the results.

CHAPTER 4

FINDINGS

This chapter reports the findings of this research study. The chapter has four sections. The first section was a description of the background of each of the participants. The remaining three sections present the major findings of the study. The first major finding was a proposed measure describing the complexity of two-dimensional grids. The proposed measure was informed by the data collected in this research study and is supported by previous research (Attneave, 1959; Falk & Konold, 1994, 1997; Klinger & Salingaros, 2000). The second major finding was a description of the participants' conceptions of entropy. I connect the participants' knowledge to the theoretical framework and show how their conceptions of probability and randomness affect their conceptions of entropy. Falk and Konold (1994) suggested there was a "similarity between the concepts of randomness and complexity" (p. 9) and that "it might be possible to foster a more intuitive, yet mathematically sound, conception of randomness if it is introduced via the complexity interpretation" (p. 10). The last major finding of this research study suggests that the complexity interpretation of randomness may not necessarily foster a sound understanding of randomness.

The Participants

All the participants who volunteered for the study were post secondary students who happened to be enrolled in a secondary science education methods course at a large public university in Georgia. Five participants attended classes full-time at the main campus of the university. Three participants were non-traditional students from a secondary science education methods course at a satellite campus of the same university.

Traditional SPSTs

Four of these participants were undergraduate secondary science education majors: two in biological sciences (Laura and MK), one in chemistry (Alice), and one in earth sciences (Jackie). Maria received an undergraduate degree in cellular biology from the same university and was a graduate student in secondary biology education at the time of the study. The teaching experiences of these participants were limited to science education projects in which they taught one class under the supervision of their cooperating teacher. For example, between the second and third interviews, Alice and a classmate co-presented the periodic table to one secondary chemistry class. The participants from the main campus completed the interviews during the Spring 2009 semester.

Non-traditional SPSTs

After completing an undergraduate degree at another institution, the non-traditional SPSTs spent time working in a career other than science. Lucy was the only participant with full-time experience in secondary school prior to enrolling in the secondary science teacher education program. All three SPSTs had experience as substitute teachers in a secondary science class. None of these SPSTs had experience teaching statistical concepts, kmt, or entropy to secondary students.

Lucy received an undergraduate degree in chemistry and worked in non-chemistry related fields between graduation and entering the science education program. Lucy was the only participant with full-time experience at a secondary school where she worked as a school psychologist and taught a secondary psychology class for one year. More recently, Lucy has experience with substitute teaching on the elementary level. Lucy was currently assisting the teacher of a college preparatory chemistry class at a large public high school near the university. In addition, she has occasionally substituted for the teacher of this class.

Erica has an undergraduate degree in broadcasting. After a few years, Erica decided to pursue a medical degree so she took several semesters of a pre-medicine program. After a couple of years in the pre-medicine program, Erica decided to pursue a degree in education. Erica has experience as a substitute teacher in a secondary chemistry class.

Tree has an undergraduate degree in ocean engineering. For his undergraduate degree, he has taken more courses that include statistical mechanical concepts than any other participant. Tree also has experience as a substitute teacher in secondary mathematics and physics classes.

Two of these participants, Erica and Lucy, completed all four interviews during the second half of the Fall 2008 semester. Tree completed two interviews during the Fall 2008 semester and completed the remaining two interviews during the first month of the Spring 2009 semester.

The Case of Alice

Alice was enrolled in a secondary science education methods course with an emphasis in chemistry. Alice claimed that among the university-level courses that contained concepts of statistics were algebra and trigonometry, Calculus I, Physical Chemistry, and physics. Further, Alice stated that she was exposed to kmt in physics and chemistry courses.

When defining complex arrangement, Alice explicitly used the term random. She stated that more complex arrangements looked “more jumbled”, i.e., “more randomly placed” in the sequence or grid. Alice admitted that she did not “really know” what randomness was. When ranking the binary sequences according to complex arrangement, Alice considered length of the runs (i.e., when the same symbol repeats within the sequence) and the number of runs that the sequence contained.

Even though Alice did not use the word pattern in her definition of complex arrangement, she explicitly used the idea of pattern when she considered the complexity of the grids. Specifically, Alice claimed Series 2 – Grid III had no pattern. When ranking the binary and tertiary grids, she used a different major characteristic for each type of grid: For the 50/50 binary grids, the appearance of alternations was the key characteristic she used. However, she did identify some patterns, e.g., a figure that she recognized in Series 1 – Grid III. In ranking the 16/84 grids, Alice used the number of pairs of B cells recognized in the grids; however, she did mention alternations for some of the grids. Lastly, in the 16/16/68 tertiary grids, she used the number of plus signs that she noticed in the grids as the only characteristic that determined the complexity of the grids.

When casting dice, Alice considered all outcomes as mathematically equal, but experientially the outcome with different numbers is more likely than the outcome with the same number. Thus, she based her responses to the dice questions on her experience, which is "Data-Driven Reasoning" (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Because used her personal experiences, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999). Furthermore, by reasoning from personal experience instead of based on the sample space, she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

When determining the most probable outcome of strings of letters, finding the distribution of molecules within a set of energy levels, or calculating the number of combinations of red and white tokens, Alice exhibited Level 3 Informal Quantitative Reasoning (Jones, et. al., 1997, 1999), because she mostly only partially determined the sample space by using a generative strategy. The smaller number of objects for which she had to find the number of combinations, Alice was more consistently able to determine the combinations. However, and more importantly, Alice was not aware of the degree to which the number of combinations increases as the number of objects used increases. This will have implications for the third key finding of this research.

When discussing the movement of odors within a room, Alice suggested the diffusion was the movement of a gas throughout the room from high concentrations to lower concentrations. However, she was not able to explain the mechanism or driving force that makes diffusion work in the direction that Alice identified. Lastly, Alice suggested entropy is a measure of the disorder of things (Entropy as Disorder).

The Case of Erica

Erica was enrolled in a secondary science education methods course with an emphasis in chemistry. Erica had an undergraduate degree in broadcasting. After a couple of years in a pre-medicine program, Erica decided to pursue a degree in education. Erica claimed that no math, science, or science education class in which she enrolled contained any concepts of statistics or statistical mechanics.

When defining complex arrangement, Erica explicitly used the term random. She also stated that complex arrangements contained more alternations of symbols or colors. Erica identified randomness as not having a pattern. By not identifying any patterns in the long run, Erica did not have an understanding of randomness (see Metz, 1998). When ranking the binary sequences according to complex arrangement Erica did not use any specific characteristic of the sequences; however, she did argue that she ranked the sequences on how complex the sequences seemed to her.

Even though Erica used the word random in her definition of complex arrangement, she did not explicitly use the idea of randomness when she considered the complexity of the grids. Even though Erica identified the characteristics of a few grids, she mostly ranked on how complex the grids appeared to her. This was especially true for the 16/16/68 grids, in which she identified no specific characteristics of the grids while ranking them.

When casting dice, Erica stated the odds for the outcome with different numbers is more than the odds for the outcome with the same number. However, she was not able to identify the value for the odds for any outcome of dice. It appeared that Erica ultimately based her responses to the dice questions on her experience, which is “Data-Driven Reasoning” (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Because used her personal experiences, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999). Furthermore, by reasoning from personal experience instead of based on the sample space, she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Erica had difficulty in finding the most probable outcome of strings of letters, finding the distribution of molecules within a set of energy levels, or calculating the number of combinations of red and white tokens. In particular, she was unable to determine the number of combinations of molecules within the energy levels. Thus, Erica exhibited Level 1 Subjective Reasoning (Jones, et. al., 1997, 1999).

When discussing the movement of odors within a room, Erica suggested the diffusion was the movement of a gas throughout the room from high concentrations to lower concentrations. However, she was not able to explain the mechanism or driving force that makes diffusion work in the direction that Erica identified. Lastly, Erica suggested entropy is disorder (Entropy as Disorder) and suggested that entropy was similar to randomness.

The Case of Jackie

Jackie was enrolled in a secondary science education methods course with an emphasis in earth sciences. Jackie claimed that no science or science education class in which she enrolled contained any concepts of statistics or statistical mechanics. She thought Calculus I included some discussion of statistical concepts, but was not completely sure. She claimed not other math course contained statistics.

Jackie defined complex arrangements as more mixed and without having patterns. She defined randomness as no preparation or not based on anything. For example, Jackie suggested selecting two socks from a drawer without being able to see which socks she was selecting. However, she made no mention of how mixed the drawer of socks was which may introduce some dependence in the selection. When ranking the binary sequences according to complex arrangement Jackie did identify long runs (5 or more) of symbols and recognized alternations of symbols; however, she ultimately ranked the sequences on the "goodness of the pattern" recognized in the sequences. She was not able to explain what she meant by the goodness of the pattern.

Jackie ranked the grids consistently with her definition of complexity. She consistently considered the large groups of colors and the alternations of color. However, in the 16/16/68 grids, she identified the large groups of colors and alternating pattern in only one grid each. In addition, she identified a pattern of triangles in two grids. More importantly, Jackie considered the "goodness of the overall pattern" for the tertiary grids.

When casting dice, Jackie stated different numbers is more probable than the same numbers based purely on her experience with dice and the fact that she cannot control the outcome of the dice. Thus, Jackie was using "Data-Driven Reasoning" (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Jackie systematically determined the distribution of molecules within a set of energy levels. When determining the number of combinations of letters, she found the number of combinations of *abc*, but was not able to extend the result to the other sets of letters. Thus, Jackie exhibited Level 2 Transitional Reasoning (Jones, et. al., 1997, 1999), because she occasionally was able to determine completely the sample space using a generative strategy. The smaller number of objects for which she had to find the number of combinations, Jackie was more consistently able to determine the combinations. However, and more importantly, she was not aware of the degree to which the number of combinations increases as the number of objects used increases.

When discussing the movement of odors within a room, Jackie suggested the movement of the fumes (she did not mention the molecular aspect of the gas) allows the smell to move throughout the room. However, she was not able to explain the mechanism or driving force that makes fumes move throughout the room. Lastly, Jackie suggested entropy is heat (Entropy as Heat) and the formula, ΔS (Entropy as Formula). Jackie did not describe the relationship between heat and the random motion of the molecules.

The Case of Laura

Laura was enrolled in a secondary science education methods course with an emphasis in biological sciences. Laura listed evolutionary biology, genetics, and ecology as containing statistical concepts. Laura claimed no math, science, or science education course discussed statistical mechanics.

Laura defined complex arrangements as being more random with more variety of symbols or colors, containing no patterns, and not predictable. She defined randomness as not having any order. When ranking the binary sequences according to complex arrangement Laura did identify one very long run of nine symbols, but eventually decided she was not comfortable enough to rank any of the sequences.

For the 50/50 grids, Laura did identify the lack of patterns and lack of predictability of the cell colors which was consistent with her definition of complexity. However, she was not comfortable in ranking the grids according to complexity. In the 16/84 grids, she identified some patterns, but in the 16/16/68 grids, she was not able to identify any characteristic to distinguish the complexity of the grids.

When casting dice, Laura admitted that she did not know how to determine the answer; thus, she based her responses on her experience with dice. Thus, Laura was using “Data-Driven Reasoning” (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Laura systematically determined the number of combinations of three letters, *abc*, and then compared the result to all the other sets of three letters. However, she undercounted the result because she missed counting the combinations from one set of three letters. She also systematically determined the distribution of molecules within a set of energy levels. However, she did double count in one microstate. Thus, Laura exhibited Level 3 Informal Quantitative Reasoning (Jones, et. al., 1997, 1999), because she determined the sample space by using a generative strategy. The smaller number of objects for which she had to find the number of combinations, Laura was more consistently able to determine the combinations. However, and more importantly, she was not able to apply her generative strategy to determine the number of combinations of red and white tokens. Furthermore, she was not knowledgeable of the degree to which the number of combinations increases as the number of objects used increases.

When discussing the movement of odors within a room, Laura suggested the diffusion was the movement of a gas throughout the room from high concentrations to lower concentrations. However, she was not able to explain the mechanism or driving force that makes diffusion work in the direction that she identified. Lastly, Laura suggested entropy is a measure of the disorder of things (Entropy as Disorder), but that it also is a measure of the random motion of the molecules (Entropy as Random Motion). More importantly, she was not able to determine how the random motion caused the molecules to spread out.

The Case of Lucy

Lucy was enrolled in a secondary science education methods course with an emphasis in chemistry. She had graduated a few years earlier with a degree in chemistry. Lucy returned to college to obtain teacher certification in secondary chemistry. Lucy had previously taken physical chemistry. Lucy claimed college algebra, trigonometry, and General Chemistry I/II included statistical concepts.

Lucy defined complex arrangements in terms of difficulty in reconstructing the sequence from memory. In other words, the more difficult the sequence is to copy, the more complex the sequence is. Falk and Konold (1994) suggested this is randomness as "difficulty in encoding" (p. 36). She defined randomness as not having any pattern and no connection between items, e.g., no connection between the numbers in a list of random numbers. However, Lucy did not identify the possibility of long-term patterns within the random numbers, e.g., in a large random number table, the distribution of the numerals 0, 1, 2, …, 8, 9 will be the same. This is also true for the distribution of two numerals, three numerals, etc. (see Attneave, 1959). When ranking the binary sequences according to complex arrangement Lucy did attempt to decide which sequences would be harder to reproduce from memory. To this end, she considered patters of long runs of symbols as well as the repeating pattern of alternating symbols.

For the 50/50 grids, Lucy again used her definition of complexity to rank the grids. However, for the remaining sets of grids, she identified additional characteristics that did not necessarily affect her ability to reproduce the grid, but which did have an effect on the complexity of the grids. These characteristics included horizontal, vertical, or diagonal groups of colors; symmetry of sections of the grids; and large groups of colors.

When casting dice, Lucy knew the probability of rolling any number on one die was 1/6. She also knew from personal experience that different numbers from a set of dice are more likely than the same number occurring on all the dice. She had “trouble reconciling” the fact that any number of the same die has the same probability with her experience knowing that different number are more likely for multiple dice. Thus, she rolled a pair of dice about twenty times to verify that different numbers were more likely. Thus, Lucy was using “Data-Driven Reasoning” (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Lucy systematically determined the number of combinations of letters and she systematically determined the distribution of molecules within a set of energy levels. Thus, Lucy exhibited Level 3 Informal Quantitative Reasoning (Jones, et. al., 1997, 1999), because she determined the sample space by using a generative strategy.

When discussing the movement of odors within a room, Lucy suggested that molecules in the gaseous state move throughout the room. She made no mention of diffusion or concentrations. Furthermore, she was not able to explain the mechanism or driving force that makes move throughout the room. Lastly, Lucy suggested entropy is a measure of the tendency for things to be disordered (Entropy as Tendency for Disorder). She further suggested that it requires less energy to be ordered, so systems have the tendency to become disordered. More importantly, Lucy was not able to connect her definition for entropy to the movement of food molecules.

The Case of Maria

Maria was enrolled in a graduate level secondary science education methods course with an emphasis in biology. She had graduated a few years earlier with a degree in cellular biology. Maria returned to college to obtain teacher certification in secondary biology. Maria listed the most courses taken that had a statistical component, including: Calculus I, General Biology I/II, General Chemistry I/II, Organic Chemistry I/II, Physics I/II, genetics and biochemistry. Maria also took a statistics course as part of her Masters coursework.

Maria defined complex arrangement differently than the other participants. Her biology background led her to conclude arrangements that are more complex are more patterned. She provided the analogy of simple, one-celled organism (not very complex) versus complex organisms (more complex and more patterned due to the specialized organ systems). Thus, most of her rankings are different from the other participants; however, I am not suggesting Maria is incorrect with her definition of complex arrangement. She defined randomness as unpredictable, but her definition does not include possibility of patterns in the long-term. When ranking the binary sequences according to complex arrangement Maria identified longs runs of symbols, repetition of shorter runs, and the total number of runs in the sequence and used all three to rank the sequences. For the grids, Maria used a wide variety of characteristics to rank the grids according to complexity, including characteristics she learned from art classes.

When casting dice, Maria stated different numbers is more probable than the same numbers based purely on her experience with dice and the fact that she cannot control the outcome of the dice. Thus, Maria was using “Data-Driven Reasoning” (Metz, 1998). She did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and she exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Maria had difficulty in finding the most probable outcome of strings of letters, finding the distribution of molecules within a set of energy levels, or calculating the number of combinations of red and white tokens. In particular, she was unable to determine the number of combinations of molecules within any of these activities. Thus, Maria exhibited Level 1 Subjective Reasoning (Jones, et. al., 1997, 1999).

When discussing the movement of odors within a room, Maria claimed, “molecules … will travel via diffusion from the source throughout the space it is contained within.” However, she was not able to explain the mechanism or driving force that makes the molecules spread throughout the container. Lastly, Maria was not able to provide a definition for entropy but knew that entropy “involved energy.”

The Case of MK

MK was enrolled in a secondary science education methods course with an emphasis in biology. However, MK is taking additional physical science courses so that he will be prepared to teach physical science, if necessary. MK did not complete the questionnaire, so the data collected from him is limited.

When defining complex arrangement, MK explicitly used the term random. He stated that complex arrangements had more variation and looked random. MK defined randomness as “totally unrelated things grouped together for whatever reason.” However, MK was not able to explain how this definition of randomness fit with his definition of complex arrangement. When ranking the binary sequences according to complex arrangement, MK considered only the length of the longest run in the sequence. MK considered sequences with longer runs as less complex.

MK used the same definition when considering the complexity of the grids. For the 50/50 binary grids, he considered the appearance of many cells of the same color that are connected, i.e., long runs. He used the same characteristic to rank the complexity of the tertiary grids, but also considered the weight (i.e., the comparison of the number of each of the colors on the sides of the grids) of the colors. However, MK claimed there was not enough information (i.e., “no particular clumps”) in the 16/84 grids so he could not rank them.

When casting dice, MK stated different numbers is more probable than the same numbers based purely on his experience with dice. Thus, MK was using “Data-Driven Reasoning” (Metz, 1998). He did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, she exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and he exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

MK had some difficulty in finding the most probable outcome of strings of letters, finding the distribution of molecules within a set of energy levels, or calculating the number of combinations of red and white tokens. In particular, he occasionally over counted the number of combinations. Thus, MK exhibited Level 1 Subjective Reasoning (Jones, et. al., 1997, 1999). Lastly, MK was not able to provide a definition for entropy but knew that entropy involved $T\Delta S$ from the free energy equation.

The Case of Tree

Tree has an undergraduate degree in ocean engineering several years ago. Tree returned to college to earn a teacher certificate in physics, so he was enrolled in a secondary science education methods course. For his undergraduate degree, he had taken more courses that include statistical mechanical concepts than any other participant. Tree claimed that no math, science, or science education class in which he enrolled contained any concepts of statistics.

When defining complex arrangement, Tree explicitly used the term random. He also stated that complex arrangements had a greater variation of symbols. However, Tree was not able to explain how randomness fit into his definition of complex arrangement. When ranking the binary sequences according to complex arrangement, Tree consistently used his definition of complex arrangement by considering the number of alternations in the sequences. However, Tree used several other characteristics when considering the complexity of the grids, none of which fit with his definition of complex arrangement. He used different characteristics to rank the different classes of grids.

When casting dice, Tree claimed different numbers is more probable than the same numbers based purely on his experience with dice. Thus, Tree was using "Data-Driven Reasoning" (Metz, 1998). He did not attempt to list the solution set of all possible outcomes for the dice. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nell & Sciolis Marino, 1991). Batanero, Navarro-Pelayo and Godino (1997) claim this misconception may be due to a lack of combinatorial reasoning. Furthermore, he exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones, et. al., 1997, 1999) and he exhibited the availability misconception (Fischbein, Sainati Nell & Sciolis Marino, 1991; Fischbein & Schnarch, 1997).

Tree had difficulty in finding the most probable outcome of strings of letters, finding the distribution of molecules within a set of energy levels, or calculating the number of combinations of red and white tokens. For the strings of letters, he did not attempt to list the sample space; he thought the number of combinations was a mathematical formula that included a factorial. For the distribution of molecules within a set of energy levels, he found some microstates. However, he did not find all possible microstates nor did he find the number of combinations. It seemed that Tree had difficulty with the activity. He also had difficulty with the need to consider the molecules as distinguishable in order to count the number of combinations. Thus, Tree exhibited Level 1 Subjective Reasoning (Jones, et. al., 1997, 1999). Lastly, Tree has similar difficulties with the entropy activity. Tree defined entropy as “the measure of the order of things.” He thought the concept of entropy was primarily used for systems of gases.

Questionnaire

Seven of the eight participants (except MK) submitted completed questionnaires (Appendix A) prior to interview 4. Each of the responses was self-reported and there was no attempt to verify the participants’ recollections. First, I summarized the responses, and then used the data to support the findings of the research.

Previous Exposure to Statistical Concepts

Erica and Tree claimed that no math, science, or science education class in which they enrolled contained any concepts of statistics. However, they may not recall this due to the length of time between their undergraduate degree and their participation in this study. Of the physical science preservice teachers, Alice claimed she was exposed to statistical concepts in algebra and trigonometry, Calculus I, Physical Chemistry, and physics; Lucy asserted that college algebra, trigonometry and General Chemistry I/II included statistical concepts; and Jackie thought that calculus may have incorporated some statistical concepts but was not completely sure.

The biological science preservice teachers claimed more of their courses contained statistics. Laura listed evolutionary biology, genetics, and ecology as containing statistical concepts. Maria listed the most courses taken that have a statistical component, including: Calculus I, General Biology I/II, General Chemistry I/II, Organic Chemistry I/II, Physics I/II, genetics and biochemistry. Maria was the only participant to have reported taking a statistics course.

Previous Exposure to Kinetic Molecular Theory

Six of the seven participants (not Tree) responding to the questionnaire claimed that general chemistry exposed them to the concepts of kmt. Tree claimed that general physics and not general chemistry introduced him to kmt. Laura and Maria also claimed that the general physics courses included kmt. Furthermore, Maria stated that organic chemistry included kmt concepts. Six of the seven participants (not Tree) claimed that general chemistry exposed them to the properties of gases. Alice, Maria, and Tree claimed general physics included properties of gases.

Previous Exposure to Statistics and Statistical Mechanics

At the beginning of the first interview in the Fall 2008 semester, each non-traditional SPST self-reported previous enrollment in physical chemistry, statistical mechanics, or thermodynamics. At the beginning of the Spring 2009 semester, Alice was the only traditional SPST to report that she completed physical chemistry. In addition, Alice and Maria reported exposure to concepts of statistical mechanics in general physics. Maria further claimed that general chemistry, organic chemistry, and molecular biology included concepts of statistical mechanics. Tree took a course on fluid dynamics, which he claimed contained concepts of statistical mechanics. Jackie, and Erica claimed that no math, science, or science education classes exposed them to any concepts of statistics or statistical mechanics.

SPSTs' Perceptions of Complexity

Participants' Definition of Complex Arrangement

Prior to ranking the sequences according to complex arrangement, each participant gave a definition of complex arrangement. Alice, Erica, Laura, MK, and Maria all explicitly used the concept of "random" in their definition of complex arrangement. Although Tree did not use the idea of randomness in defining complex arrangement, Tree did use the term random while supporting his ranking of the sequences. Jackie and Lucy never made any connection between randomness and complexity. Maria used a definition of complexity based on her biology background as well as from several art classes. To Maria, it “takes organization to form a pattern which requires more complex interactions.” In other words, an image that appears to have a pattern was more complex.

Complex Arrangement of Binary Sequences

In the first interview, the participants were asked to define complex arrangement and then to rank five different binary sequences containing 11 H and 10 T (see Figure 11). The participants were told the H and T represented heads and tails, respectively, of a random coin flip. Alice, Erica, Jackie, Laura, MK, and Tree provided similar definitions of a complex arrangement for the binary sequences. For these participants, the important characteristic that determined the complexity of the string was random assortment, more variation, more jumbled or more mixed up. Most definitions of complex arrangement contained the word random (Alice, Erica, Laura, MK and Tree) and the words “jumbled”, “mixed”, “variation” or “alternating” (Alice, Erica, Jackie, Laura, MK and Tree). Lucy defined complex arrangement in terms of reconstructing the sequence from memory. In other words, the more difficult the sequence was to copy, the more complex the sequence is. Falk and Konold (1994) suggested this was randomness as "difficulty in encoding" (p. 36). Maria suggested that a complex arrangement “would be the hardest to achieve and the least complex would be the easiest to achieve or more likely” to result from actually flipping a coin.

The average participant ranking of the binary sequences concurs with the research by Falk and Konold (1994, 1997) and Falk (1975). Falk and Konold concluded that a common misconception held by students was that random sequences have more alternations. Falk and Konold used the probability of alternations, $P(A)$, to measure the number of alternations in a sequence. In this research, all the participants, except Maria, ranked the binary sequences according to the misconception identified by Falk and Konold. In other words, the least complex sequence contained the fewest alternations and the more complex sequences have more alternations.

I.
H T T T H T T H H H H H H H H H T T T T T

II.
T T H H T T T H H T H T T H H H H H T H T

III.
H T T H H H T H H T T H T H T H T H T H T

IV.
H T H T T T H H T T H H H T T T T H H H H

V.
H T T T H H H T H H T T T H T H T H H T H

Figure 11. Five sequences of 21 coin flips.

Sequence I, with $P(A) = 0.250$, was consistently ranked by all the participants (except Maria) as the least complex arrangement. Maria placed Sequence I as the most complex due to her different definition of complex arrangement. There were some additional exceptions. Whereas most participants ranked Sequence III as the most complex sequence, Maria and Jackie did not. Jackie noticed that the alternating section at the end of the sequence, THTHTH, "is more patterned than the others." Jackie ranked this as the second least complex even though her definition of complex arrangement included "more mixed up." Alice and Lucy ranked Sequence IV as the second least complex because the sequence contained more runs of length four than

Sequence II. The other participants considered Sequence II a little more complex than Alice and Lucy. Both participants considered the interplay between the longer runs and the greater number of runs, both of which make sequences less complex. More specifically, Lucy claimed that she based her ranking of the grids on "the total number of runs, and, um… the number of time, the length of the run, maybe." Erica used the same reasoning to place Sequence V as the second least complex where the other participants considered this sequence as one of the two most complex sequences.

Key Finding 1: Proposed Complexity Measure of Two-dimensional Grids

Complexity of Two Dimensional Grids

50/50 Binary Grids With 50 Black and 50 White Cells

The first two series contained five grids each with 50 black (B) and 50 white (W) cells. Series 1 is shown in Figure 12 and Series 2 is shown in Figure 13. The purpose was to compare the results of this study with the results of Falk and Konold (1994, 1997). Although the participants in this study did consider how the grids alternated in color, they did recognize other characteristics, which they used to determine the complexity of the grids. Furthermore, the data will support the use of overlapping 2×2 and 3×3 subgrids in $C_{2,3,4}$ and the need to reduce the influence of highly symmetric subgrids and diagonal subgrids on the complexity measure.

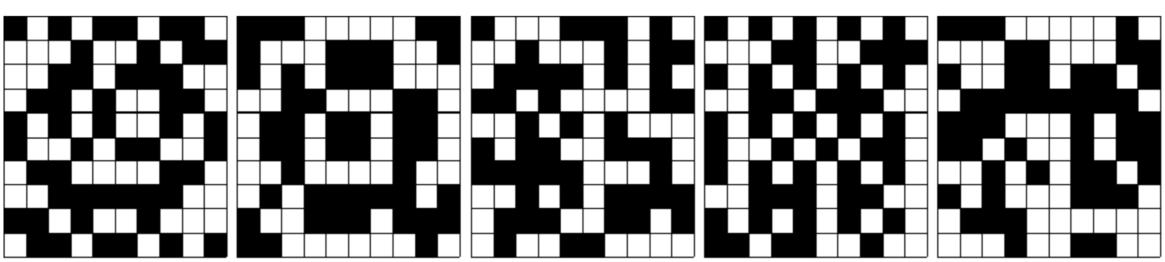


Figure 12. Grids I–V in Series 1.

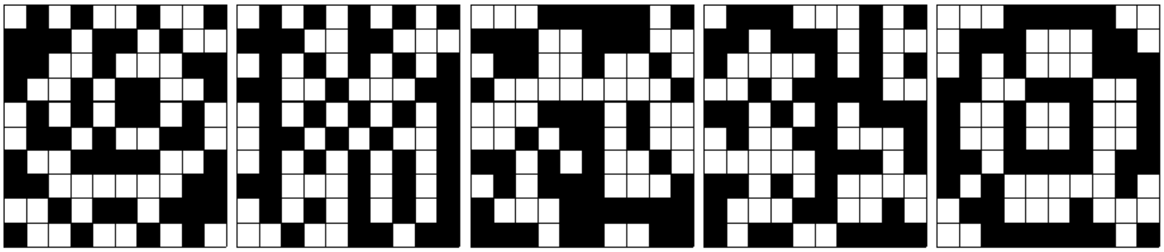


Figure 13. Grids I–V in Series 2.

Prior to ranking the grids according to complex arrangement, each participant provided their definition of complex arrangement for 10×10 binary (containing black and white cells) and tertiary (containing black, grey and white cells) grids. Most definitions of complex arrangement contained the word random (Alice, Erica, Maria, and Tree) and the word “pattern” (Erica, Jackie, Laura, Lucy, Maria, MK, and Tree). Even though Alice did not use the word pattern in her definition of complex, she explicitly used the idea of pattern when she considered the complexity of the grids. Specifically, Alice claimed Series 2 – Grid III had no pattern. Other ideas used by participants in their definition of complex included: jumbled (Alice and Jackie), “lack of sequencing” (Tree), “not really neat looking” (Jackie), “not predictable” (Laura) and “harder to make sense out of it” (Alice). Occasionally, Lucy considered the difficulty of reproducing a grid from memory. The more complex grids were more difficult to reproduce. This idea was the “difficulty in encoding” conception of randomness (Falk & Konold, 1994; Green, 1979, 1989, 1991). While ranking the complexity of the grids, the participants consistently used their definitions of complex arrangement with most participants expanding on types of patterns they noticed in the grids. For example, the participants identified symmetries; specific patterns like figures, letters, checkerboard; and other groupings of black and white cells. However, there was one important exception: None of the participants mentioned randomness when evaluating the complexity of the grids.

Proposed Complexity Measure of Two-dimensional Grids

I propose a complexity measure [identified in this document as $C_{2,3,4}$[24]] for 10×10 grids based mainly on the complexity measure suggested by Klinger and Salingaros (2000) [identified in this document as $C_{2,5,10}$]. The proposed measure has a few important differences supported by psychology research (Attneave, 1959) and the data collected in this study. I will show that the proposed complexity measure concurs with most of the conclusions of Falk and Konold (1994, 1997). However, there are two main advantages of $C_{2,3,4}$ over the probability of alternations (Falk & Konold, 1994, 1997), $P(A)$: first, participants did not solely use alternations while comparing the complexity of the grids; and second, $C_{2,3,4}$ can be used for grids other than ones with 50 black cells and 50 white cells. There are two disadvantages of $C_{2,3,4}$ over $P(A)$: first, the calculation of $C_{2,3,4}$ is much more complicated; and second, Falk and Konold (1994, 1997) used $P(A)$ to explain truly random grids as well as individuals' misconceptions of randomness whereas $C_{2,3,4}$ can only describe the complexity of grids.

There are three major differences between $C_{2,3,4}$ and $C_{2,5,10}$. First, the number of unique colored cells per subgrid, $T(n \times n)$, and the symmetry of each subgrid, $H(n \times n)$, were calculated in a similar manner for both complexity measures. However, in $C_{2,5,10}$ only disjoint subgrids were considered; whereas in $C_{2,3,4}$, all 81 overlapping 2×2 subgrids, all 64 overlapping 3×3 subgrids, and all 49 overlapping 4×4 subgrids were used to calculate the complexity measure. Attneave's (1959) higher-order measure of randomness for binary sequences provided support for the use of overlapping subgrids in $C_{2,3,4}$.

Second, $C_{2,5,10}$ places too much weight on the symmetry of highly symmetrical subgrids (see Figure 14 for some examples). If a 10×10 grid contains more than one of any of these subgrids, then the subgrid has a corresponding $H = 9$ and resulting in $H_{max} = 9$. Because only

[24] The complexity measures, $C_{i,j,k}$, use subscripts to identify the [sub]grids that are the basis of the calculations.

eight of the 512 total 3×3 grids are highly symmetrical, this leads to an over-valuation of $H_{max} - H$ and consequently, an over-valuation of $C_{2,5,10}$. Because there are only 64 total overlapping 3×3 grids, having two of the same highly symmetrical 3×3 grids in a truly random two-dimensional grid would be rare. I think there was too much weight (i.e., results in a higher value of $C_{2,5,10}$) attributed to these highly symmetrical 3×3 subgrids. The same argument can be made for the 4×4 subgrids. Thus, the proposed formula for $C_{2,3,4}$ assigns each highly symmetrical $n \times n$ grid an H-value of 1. More specifically, none of the participants explicitly mentioned symmetry on a scale (i.e., subgrid $m \times m$, where $m < 10$) other than 10×10 for the 50/50 grids. In addition, participants recognized large blocks of a single color but did not mention their symmetry. For example, the participants did not determine if the all-B $n \times n$ grids have any type of symmetry (h_1 through h_6) and did not reflect (h_8) or rotate (h_9) this grid to compare with another all-B $n \times n$ grid. Thus, the proposed formula for C assigns each $n \times n$ grid of a single color an H-value of 1.

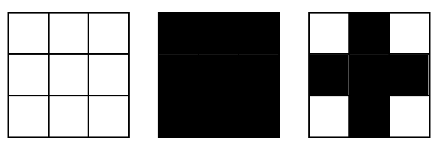


Figure 14. Three examples of 3×3 subgrids with h_1 through h_6 symmetries[25].

Third, there was a similar problem of overweight H-value (see Table 2 for the various symmetries used in the complexity measure) for the diagonal $n \times n$ patterns in the $C_{2,5,10}$. Each diagonal $n \times n$ subgrid has 180° rotational self-symmetry (h_6) and either reflectional self-symmetry over the $y = x$ line (h_3) or over the $y = -x$ line (h_4) and may have 180° rotational

[25] See Table 2 for a list of the symmetries.

symmetry with another $n \times n$ subgrid (h_9). So, in the complexity measure $C_{2,3,4}$, the calculation of H does not include the value for (h_6) for diagonal subgrids.

SPSTs' Ranking of Binary 50/50 Grids. The average results of the SPSTs' ranking of the 50/50 grids concur with the findings of Falk and Konold (1994, 1997) (see Tables 8 and 9). In general, the participants ranked from smallest $P(A)$ to largest $P(A)$. The most common characteristic of the grids identified by the participants was the appearance of alternating patterns, a.k.a., the checkerboard pattern. It is important to note however, the participants in this study considered additional characteristics of the grids. In other words, although each participant alluded to alternations as they considered each grid, they frequently suggested other characteristics that had greater impact on their determination of complexity.

Table 8. Comparison of proposed $C_{2,3,4}$ measure of complexity with $P(A)$, $C_{2,3}$, and $C_{2,5,10}$ for the binary grids in Series 1 and 2.

Series 1	I	II	III	IV	V
Series 2	I	V	IV	II	III
$P(A)$	0.622	0.461	0.533	0.656	0.483
$C_{2,5,10}$	4.547	4.496	2.733	3.863	6.325
$C_{2,3}$	2.082	2.306	2.518	1.487	2.531
$C_{2,3,4}$	2.908	3.035	3.339	2.365	3.299

Major Characteristics Used to Rank 50-50 Grids. Five participants (Alice, Jackie, Laura, Lucy, and Maria) used "checkerboard pattern" or alternations when explaining their complexity rankings. In general, for these participants, when a section of a grid looked like a checkerboard pattern, they deemed the grid more complex. These sections were highly patterned, which participants claimed affected the complexity of the grids. Even though none of the participants explicitly referred to the symmetries of the 2×2 or 3×3 subgrids, it seems reasonable that the checkerboard pattern arises from the participants' consideration of 2×2 cells with 2 black cells on either diagonal as seen in Figure 15.

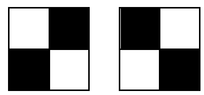


Figure 15. 2×2 diagonal cells responsible for checkerboard pattern.

Table 9. Comparison of mean participant rank with measures of the complexity of the grids in Series 1 and 2.

Series 1	I	II	III	IV	V
Mean rank	3	1	4	5	2
Median rank	3	1	4	5	2
Series 2	I	V	IV	II	III
Mean rank	4	1	3	5	2
Median rank	4	1	5	2.5	2.5
$P(A)$ Valid	2	3	4	1	5
$C_{2,5,10}$	4	3	1	2	5
$C_{2,3}$	2	3	4	1	5
$C_{2,3,4}$	2	3	5	1	4

Several of these 2×2 diagonal subgrids results in the checkerboard pattern (see Figure 16).

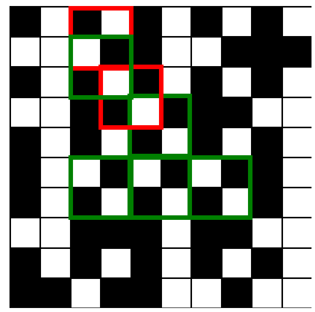


Figure 16. Identification of the 2x2 diagonal subgrids responsible for the checkerboard pattern in Series 1 – Grid IV.

Several of the participants (Erica, Jackie, Laura, and Tree) recognized small letter-shapes as shown in Figure 17. The participants considered the letter shapes as "islands" of order within the grids resulting in less complexity. Laura recognized plus signs in both Series 1 and 2. More importantly, she specifically identified overlapping plus signs as shown in Figure 18.

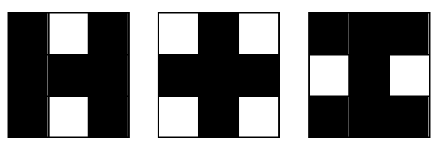


Figure 17. Letter-shapes recognized by participants in Series 2 – Grid II.

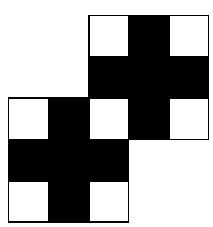


Figure 18. Two overlapping plus signs pattern from Series 1 – Grid IV.

Finally, several participants (Alice, Erica, Laura, Maria, and Tree) recognized an apparent symmetry on the 10×10 scale in Series 1 and 2 – Grid I. This symmetry was apparent because it was not a perfect symmetry; only certain pieces (i.e., subgrids) actually had some type of reflectional or rotational symmetry (see Figure 19). Even though $C_{2,3,4}$ does not include a component to measure apparent symmetry, a complexity measure needs to account for the symmetries on the 2×2, 3×3 or 4×4 scale.

Lastly, some of the participants considered either large groups (i.e., more rectangular shapes where the width of the rectangle was at least two cells) of B cells (Jackie, MK) or long strings (i.e., where the width was one or two cells) of B cells (Lucy, MK) or W cells (Maria, MK). As the number of connected B or W cells increases, the number of possible alternations decreases (see Falk & Konold, 1994, 1997). Although recognized by some participants, rectangular-shaped single color subgrids did not contribute any value to the complexity measure.

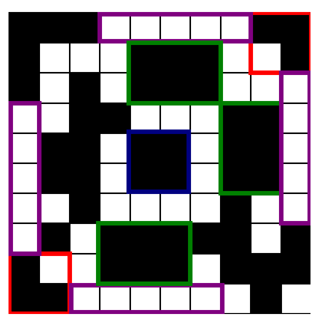


Figure 19. The sections of Series 1 – Grid II that exhibit symmetry with other sections of the grid.

Thus, the data collected and analyzed in this study provided empirical support that the participants took into consideration patterns identified on a small scale (2×2 or 3×3) as well as the overall grid (10×10). However, unless the 10×10 grid actually is self-symmetrical, then only the subgrids were important. The symmetries on the 2×2 or 3×3 scale can describe any symmetry on the 10×10 scale. Thus, any measure of the complexity of 10×10 grids must consider the 2×2 or 3×3 subgrids.

Comparison of Complexity and Randomness Measures for 50/50 Grids. Most importantly, the ranking of the 50/50 grids from least complex to the most complex using the complexity measure, $C_{2,3,4}$, matches the ranking of the grids using Falk and Konold's probability of alternations, $P(A)$. Falk and Konold (1994, 1997) compared $P(A)$ to several other measures of randomness and concluded that all the measures have a high correlation for binary grids with 50 B and 50 W cells. However, $C_{2,3,4}$ takes into consideration the scale(s) on which the majority of the participants rated each grid (see Tables 10 and 11). Actually, most participants (excluding Alice, Lucy, and MK) considered the patterns contained in the 2×2 and 3×3 subgrids for the 50/50 grids. The addition of the 4×4 subgrids was important for the grids in Series 3 and 4.

Outliers to Consider. Although the proposed complexity measure, $C_{2,3,4}$, and $P(A)$ both describe the complexity and randomness, respectively, of binary 50/50 grids, $C_{2,3,4}$, does not describe the participants' perception of complexity. Falk and Konold (1994, 1997) argued that $P(A)$ can describe participants' misconceptions of randomness; however, this solely relies on the alternations of a grid. However, the participants in this study considered more than just the number of alternations that affected their perception of the complexity of a grid.

Table 10. Comparison of participants' rank with measures of the complexity of the grids in Series 1.

Series 1	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.622	0.461	0.533	0.656	0.483
$C_{2,5,10}$	4.547	4.496	2.733	3.863	6.325
$C_{2,3,4}$	2.908	3.035	3.339	2.365	3.299
Alice	2	1	4	5	3
Erica	5	2	3	4	1
Jackie	4	1	3	5	2
Laura	1	NR	NR	2	NR
Lucy	4	1	3	5	2
Maria	4	5	2	1	3
MK	3	1	4	5	2
Tree	3	1	4	5	2

First, most of Maria's rankings were opposite of the other participants. This was due to her definition of complex arrangement. Maria argued that "more complex" equals "more ordered." She claimed her reasoning was a direct result of her biology background where she learned that organisms that were more complex were more highly ordered. Having taken some art classes, Maria frequently considered the "weights" of the colors. In other words, Maria considered if the colors were evenly distributed throughout the grid or if one color was more grouped to one side or corner of the grid. Maria considered evenly distributed grids as more complex.

Second, Erica was an outlier for both Series 1 and 2 because she ranked the grids significantly different from the other participants. However, Erica's supporting statements for the basis of her rankings were no different from the statements made by the other participants. Erica's definition of complex arrangement and definition of randomness were similar to Laura's definitions. The only difference was that Erica had some difficulty in verbalizing her reasons. There were only two statements that Erica previously made that may provide insight. First, when Erica ranked the binary series, she stated that she guessed the complexity rankings "from just looking at it." It seems that she completed the ranking of the grids in a similar manner. Second, Erica argued that in a sequence of ten fair coin tosses, she would expect "more T than not T." She also made the same argument for the number of H in ten fair coin tosses. Further questions did not provide any additional insight into Erica's reasoning.

Third, Tree's Series 2 ranking was significantly different from the other participants, though the ranking was similar to Erica's ranking of Series 2. Likewise, there was only one statement made by Tree that may offer some insight. Similar to Erica, Tree argued that in a sequence of ten fair coin tosses, he would expect "fewer T." He also made the same argument for fewer H in a sequence of ten fair coin tosses. Additional questions posed to Tree did not provide any additional insight.

Table 11. Comparison of participants' rank with measures of the complexity of the grids in Series 2.

Series 2	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.622	0.656	0.483	0.533	0.461
$C_{2,5,10}$	4.547	3.863	6.325	2.733	4.496
$C_{2,3,4}$	2.908	2.365	3.299	3.339	3.035
Alice	2	5	4	3	1
Erica	5	4	3	2	1
Jackie	4	5	2	3	1
Laura	1	2	NR	NR	NR
Lucy	4	5	2	3	1
Maria	4	1	2	3	5
MK	3	5	2	4	1
Tree	5	2	3	4	1

Fourth, Alice and Maria recognized a man in Series 1 – Grid III or Series 2 – Grid IV. Maria stated the figure "looks like Mario – the old school one" (see Figure 20). MK saw a "monster" (see Figure 21) in Series 2 – Grid I. In each of these cases, Alice, Maria, and MK used these figures in their determination of the complexity (as among the least complex in the series) of the entire 10×10 grid based on the subgrid that contained the figure. For example, MK stated, "I would go with this one as more ordered because I can kind of see, what looks like a little monster in between two white lines."

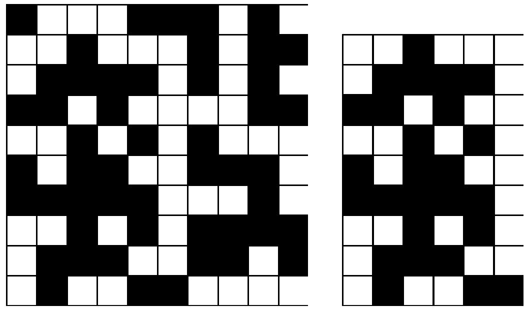


Figure 20. Series 1 – Grid III with the "old school Mario" section isolated on the right.

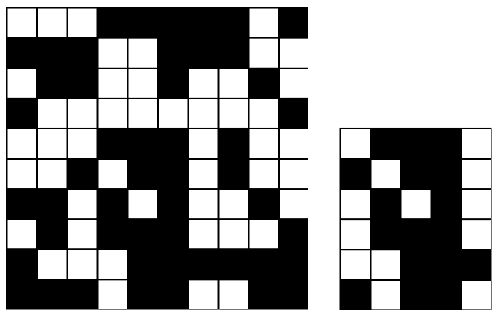


Figure 21. Series 2 – Grid II with the "monster" section isolated on the right.

In summary, the data shows that on average the participants' ranking of the complexities of the grids was consistent with the conclusions of Falk and Konold (1994, 1997). It was important to show that $C_{2,3,4}$ was consistent with $P(A)$, because Falk and Konold have shown the reliability of P(A) as a measure of randomness of two-dimensional grids. However, not only did the participants considered the alternations when contemplating the complexity, they also looked at eleven other characteristics (see Table 12). Some of these characteristics supported the use of the same basic structure of $C_{2,5,10}$ (Klinger & Salingaros, 2000). Ultimately, the work of Attneave (1959) encouraged the use of overlapping grids in $C_{2,3,4}$. Lastly, based on the data, there were some additional modifications to decrease the effect of highly symmetrical subgrids.

16/84 Binary Grids With 16 Black and 84 White Cells

By considering 16/84 binary grids, I will provide evidence for the need to include 4×4 subgrids in $C_{2,3,4}$. In addition, I will show there was no evidence that 5×5 subgrids were necessary for a complexity measure. In fact, 5×5 subgrids did not contribute much to the complexity measure (see Attneave, 1959). Lastly, the data will show that some of the participants (i.e., Laura, Lucy and Maria) recognized rotational and reflectional symmetries among the subgrids $n \times n$, $n < 5$. This provided support for the use of the $H(n \times n)$ as part of the calculation of $C_{2,3,4}$. Lastly, I will show that $C_{2,3,4}$ was a more consistent measure of the grids than $P(A)$ (see Tables 13 and 14).

Ranking grids with only 16 B cells with respect to complex arrangement was more difficult for all but one participant (Jackie). MK and Laura were not comfortable in ranking the grids because they did not think there was enough information contained in the grids to distinguish the grids based on complex arrangement. For example, MK argued, “there are no particular clumps in any of the grids.” Alice, Erica, Jackie, Lucy and Maria considered the number of pairs (i.e., two B cells next to each other horizontally or vertically) in their explanation of complex arrangement. However, when asked how many pairs to expect from randomly generated grids containing 16 B cells, all participants suggested 1-3 pairs. Lucy stated “random leads to fewer pairs.” Whereas Jackie argued that, “the most random is the middle complexity.” Thus, it seemed that complex arrangement did not have anything to do with randomness even though each participant mentioned randomness in his or her definition of complex arrangement. It may be that the title of the study caused them to include the word random in their definition.

Table 12. Coding schema used to analyze participants' conceptions of complexity.

Characteristics	Coding	Participants' use of Characteristics		
		Grids I and II	Grids II and IV	Grids V and VI
Symmetry over $y=x$	SYM-YX	E, La		
Symmetry over y-axis	SYM-AX	A, E		
Symmetry–general	SYM-Gen	La, T	Ma, Lu	T
Symmetry–rotational	SYM-Rot		La	
Crosses/Plus signs	PLS	La, E, J		Ma, A
Letters	ABC	T	La	
Triangles one color	TRI			Ja
Figures	FIG	Ma, A, MK, Lu, T		
Diagonals of B, G and/or W cells	DIA			Ma, MK, J, Lu, T
Diagonal pattern of B	DIA-B		Lu	
Diagonal of W only	DIA-W		Ma	
Pairs of cells	PRS		Ma, A, J, Lu, E	MK, E, T
Single cells	SGL		J, Lu	J, Lu
Strings of B cells or W cells	STR	MK, J, Lu		
Strings of W cells	STR-W		T	
Large blocks of B cells or W cells	BLK	MK, J		
Large blocks W cells	BLK-W		E	MK, Lu, T
Checkerboard pattern	CBP	Ma, A, J, Lu, T	A, Lu	
Scattered	SCT		Ma, Lu, La, T	
Vertical/Horizontal appearance of patterns	VHP	Lu	J, MK, Lu	Lu
Weight	WGT	Ma		MK, J
Pattern – generic	PAT			T
No pattern	NOP	La, A		

Table 13. Comparison of participants' rank with measures of the complexity of the grids in Series 3.

Series 3	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.300	0.272	0.344	0.322	0.278
$C_{2,5,10}$	5.846	5.654	6.230	3.380	5.743
$C_{2,3,4}$	2.434	2.304	2.383	2.496	2.370
Alice	3	5	1	2	4
Erica	3	1	5	2	4
Jackie	2	5	1	3	4
Laura	NR	NR	NR	NR	NR
Lucy	5	4	1	2	3
Maria	3	1	5	4	2
MK	NR	NR	NR	NR	NR
Tree	4	2	5	1	3

According to Styer (2000), the random generation of these grids would result in 4-5 pairs. Most of the participants considered the relative locations of the B cells; Erica, Jackie, Maria, and Tree did identify large groups of W cells contained in some of the grids (notably, Series 3 – Grids II and IV). There were some discussions of checkerboard patterns by Alice and Lucy in Series 3 – Grid III. For example, Alice stated, "[Series 3 – Grid III] looks kinda like a checkerboard pattern to me" and ranked the grid as most complex. It is important to note that the checkerboard pattern in Series 3 and 4 were not the same as the checkerboard pattern of Series 1 and 2.

Several participants recognized various possible symmetries. Laura noticed rotational similarity between the top-half and bottom-half of Series 3 – Grid I. Laura also noticed rotational symmetry between two 4×4 subgrids in Series 4 – Grid I. Maria noticed reflectional symmetry [over the $y = x$ line] between two 3×3 subgrids containing one B pair and a single B cell in Series 4 – Grid I. Lucy noticed rotational symmetry between two 4×4 subgrids containing two B pairs in Series 3 – Grid II.

Table 14. Comparison of participants' rank with measures of the complexity of the grids in Series 4.

Series 4	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.283	0.311	0.322	0.311	0.306
$C_{2,5,10}$	5.678	6.101	5.891	5.789	5.962
$C_{2,3,4}$	2.165	2.551	2.653	2.197	2.227
Alice	3	4	5	1	2
Erica	3	1	2	5	4
Jackie	3	5	4	1	2
Laura	NR	NR	NR	NR	NR
Lucy	2	1	4	5	3
Maria	4	2	3	5	1
MK	NR	NR	NR	NR	NR
Tree	3	2	4	5	1

Even though most participants recognized some sort of pattern in Series 4 – Grid IV, none of them identified the repeating pattern of the 5×5 subgrid. In fact, the computation of $C_{2,3,4}$ changed very little when including $n \times n$, $n \geq 5$ subgrids (see Attneave, 1959). After the researcher pointed out the pattern, none of the participants considered changing their re-ranking of Series 4. Most of the participants argued that the characteristics that they noticed in the grids determined their rankings. In other words, because they did not really see the pattern, then their ranking should not change. Lucy said, the ranking "had to do with just looking at it and how does it appear versus really trying to understand the arrangement." However, Erica was unsure whether she would change the ranking but eventually decided to leave the original ranking. In order to re-rank Series 4, Erica argued that she "would have to go back and re-evaluate [the series] all over again." Furthermore, Erica was not sure if she should apply the initial characteristics that she used or to determine some other set of characteristics in order to re-rank the series. After Tree was shown the pattern, he claimed that if the pattern repeated itself more often, he would have noticed it. In addition, only Tree considered the all-W 5×5 subgrid significant in the upper-left corner in the ranking of Series 3 – Grid IV because he considered large sections of W as important.

The complexity measure $C_{2,3,4}$ appears more reliable for Series 3 and 4 than $P(A)$ for several reasons (see Tables 13 and 14). First, $P(A)$ can easily change depending on how many cells were along the edge of the 10×10 grid. Consider Series 4 – Grids II and IV (see Table 14). Both have the same value for $P(A)$, but Grid IV has no pairs and Grid II has four pairs. The grids were too visually different to have the same value for $P(A)$. The $C_{2,3,4}$ values for these two grids more closely agrees with the expected number of 4-5 pairs (see Styer, 2000). Second, Series 4 – Grid IV was highly patterned because it contains the same pattern on each of the four disjoint

5×5 subgrids. However, $P(A)$ does not distinguish Series 4 – Grid IV as highly patterned (i.e., smaller $P(A)$). Third, the $P(A)$ values for Series 3 – Grid IV and Series 4 – Grid III were identical because each grid has no B cells along the edges and each grid has three pairs. However, P(A) does not differentiate the all-W 5×5 subgrid in Series 3 – Grid IV from Series 4 – Grid III. The complexity measure $C_{2,3,4}$ appears more consistent with $C_{2,5,10}$ for Series 3 and 4. However, both complexity measures were more reliable than for the other types of grids.

Table 15. Comparison of proposed $C_{2,3,4}$ measure of complexity with $P(A)$ and $C_{2,5,10}$ measures of the binary grids in Series 3.

Series 3	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.300	0.272	0.344	0.322	0.278
No. of pairs	2	5	1	3	4
Mean rank	4	2.5	2.5	1	5
Median rank	3	3	3	1	5
$C_{2,5,10}$	5.846	5.654	6.230	3.380	5.743
$C_{2,3,4}$	2.434	2.304	2.383	2.496	2.370

Comparison of Ranking 16/84 Binary Grids. There was no previous research with which to compare these results. However, Styer (2000) argued there should be approximately 4-5 pairs (i.e., two adjacent black cells horizontally or vertically) in a randomly constructed grid. However, the participants predicted fewer pairs for randomly constructed grids. Styer (2000) claims this was a common misconception.

Table 16. Comparison of proposed $C_{2,3,4}$ measure of complexity with $P(A)$ and $C_{2,5,10}$ measures of the binary grids in Series 4.

Series 4	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.283	0.311	0.322	0.311	0.306
No. of pairs	2	4	3	0	1
Mean rank	3	2	4.5	4.5	1
Median rank	3	1.5	4	5	1.5
$C_{2,5,10}$	5.678	6.101	5.891	5.789	5.962
$C_{2,3,4}$	2.165	2.551	2.653	2.197	2.227

The most interesting result from Series 3 (Figure 22) and 4 (Figure 23) is the wide variety of characteristics that the participants used to rank the 16/84 grids compared to the characteristics used by the participants in the 50/50 grids (see Table 12). An additional difficulty may be the close values of all the various measures of randomness or complexity for these grids. This may be the reason Laura and MK could not find any characteristics with which to distinguish the grids from one another in Series 3 and 4. The consideration of large blocks of W cells supports the inclusion of 4×4 subgrids in the proposed complexity measure. However, because no participant recognized the 5×5 repeating pattern in Series 4 – Grid IV, there was no need to include subgrids of 5×5 or larger in $C_{2,3,4}$.

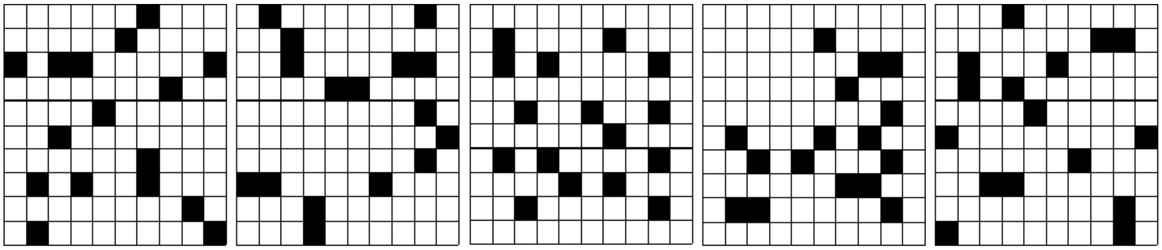


Figure 22. Grids I–V in Series 3.

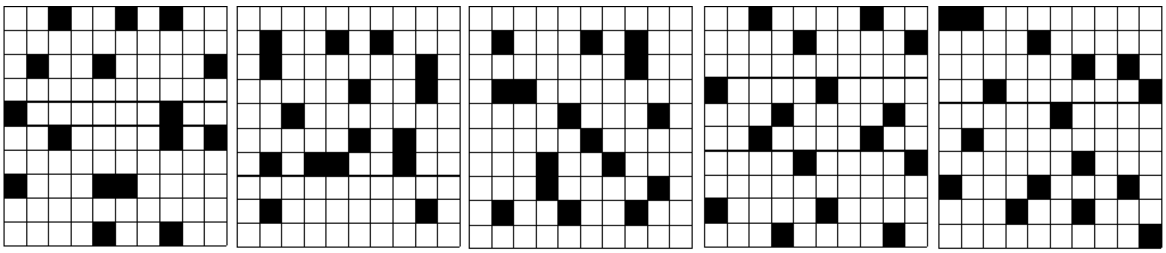


Figure 23. Grids I–V in Series 4.

Alice and Jackie seemed to be the most consistent over both Series 3 and 4. These two participants mainly ranked the grids based on the number of pairs. Alice additionally mentioned that a checkerboard pattern was another key characteristic of these grids. In other words, the more pairs in the grids, the less complex the grid appears. The minor differences between their rankings are: switched Series 3 – Grid I and Series 3 – Grid IV as the second and third least complex in Series 3; switched Series 4 – Grid II and Series 4 – Grid III as the top two most complex in Series 4. In both cases, Alice ranked the grids with more pairs as having more complexity. Although Alice did talk about the perception of the checkerboard pattern during the actual ranking of these series, she claimed at the end of the interview that the checkerboard pattern was "the essential characteristic" for these grids. However, she essentially ranked these grids according to the number of pairs that she recognized – the more complex arrangement had more pairs. This seemed to contradict her definition of complexity as more jumbled. In these grids, fewer pairs results in more singles and, possibly, a checkerboard-like pattern. Here was a brief exchange between Alice and the researcher:

> R: Let me go back to your definition of complex arrangement. You said in the more complex arrangement, [the cells] would be more jumbled. So, you're saying [Series 4 – Grid IV] is least complex?
>
> A: (quietly) Yeah.
>
> R: Even though some may say that Grid IV is more jumbled?
>
> A: Um, it looks kinda like a checkerboard pattern to me. It seems like it's going in the direction [i.e., tracing with her finger up and to the left]. So, it's making something (laughs).

In this exchange, Alice was not just using any single characteristic, but the interplay of several characteristics, including the appearance of a checkerboard pattern, the number of pairs and the appearance of directional pattern of the black cells.

Erica, Maria, and Tree ranked Series 3 and 4 similarly to each other (see Tables 13 and 14). However, their rankings were not as consistent as those of Alice and Jackie. All three participants ranked Series 3 – Grid II as the least complex. Both Erica and Tree considered large blocks of W cells in this grid. These large blocks of W make the grid more patterned and appear less complex. Maria noted the number of pairs in Series 3 – Grid II. It seemed reasonable to suggest that more B pairs results in the appearance of larger blocks of W. However, Jackie also recognized the large W groups but ranked this grid as the most complex. Lucy was the only participant who suggested the appearance of some symmetry in Series 3 – Grid II; however, she ranked this grid as the second most complex because each of the other grids had characteristics that made them appear less complex than Series 3 – Grid II. For example, she considered Series 3 – Grid III to be alternating and thus, patterned and Series 3 – Grid V had more pairs.

Erica, Maria, and Tree ranked Series 3 – Grid III as the most complex grid in the series. Tree again noticed the large blocks of W cells in this grid. Even so, Tree ranked Grid III as most complex whereas he ranked Grid II with its large blocks of W cells as least complex. There does not appear to be any statement that Tree made that explains this possible discrepancy. Additional questioning of Tree did not reconcile this discrepancy. Both Maria and Erica noted the scattered appearance of the B cells in the grid, which makes this grid more complex to them. They compared the scattered appearance to the checkerboard pattern even though neither participant recognized that it was not a true checkerboard pattern.

Erica, Maria, and Tree ranked Series 4 – Grid II among the least complex grids in the series. Maria noticed this grid had the most B pairs. As she argued in Series 3, the larger number of B pairs makes the series appear less complex. Erica and Tree made no specific comments regarding this grid, so it was difficult to determine what characteristics of this grid they considered.

Erica, Maria, and Tree ranked Series 4 – Grid IV as the most complex grid in the series. There was no common argument made for this grid. Maria considered a diagonal group of W cells. In addition, Maria stated that the grid appeared symmetrical. Tree noticed there was at least one B cell in each column and row. Erica made no explicit comment about any characteristic of this grid. In addition, Lucy noted that this grid appeared more spread out.

Erica and Maria ranked Series 4 – Grid III as among the least complex grids in the series. Lucy and Tree ranked this grid as among the most complex grid. Unfortunately, there was no specific discussion of this grid to determine what characteristics of this grid that the participants considered the most important.

The remaining grids (Series 3 – Grids I, IV, V and Series 4 – Grid I) were the most consistently ranked grids among all the participants. Unfortunately, nothing the participants said provided any insight into this consistency. Consider Series 3 – Grid IV. Laura noticed T-shapes and upside-down T-shapes formed by W cells. Lucy noticed more single B cells while Maria noticed the pairs of B cells.

Outliers to Consider. There were no real outliers among the participants' rankings for these grids composed of 16 B cells and 84 W cells. The participants' rankings of these grids were more diverse than their rankings for the 50/50 grids. Again, this may be due to the close values of $C_{2,3,4}$. One possible problem was the values of $C_{2,3,4}$ for Series 4 – Grids II and Series 4 – Grid III were larger and distinct from the other grids; however, on average the participants did not rank these grids as the two with the most complex arrangement. Even so, the relatively small spread of values may explain how easy it was for a participant to confuse their perception of the complexity of the two grids. This may be supported by the fact that it took the participants longer to rank Series 3 and 4 than the first two series.

In summary, $C_{2,3,4}$ was a better description of the complexity of binary grids. $C_{2,3,4}$ was consistent with established $P(A)$ (Falk and Konold; 1994, 1997); however, $C_{2,3,4}$ provided a more detailed description of the complexity of grids. Second, the data collected and analyzed in this research study also provided the support for the necessary changes from $C_{2,5,10}$. In addition, it seemed that $C_{2,3,4}$ can be used to describe binary grids with any number of B and W cells.

16/16/68 Tertiary Grids With 16 Black, 16 Gray, and 68 White Cells

Falk and Konold (1994, 1997) did not consider tertiary grids in their research. The definition of complexity by Klinger and Salingaros (2000) seemed more appropriate to use with tertiary grids because they developed $C_{2,5,10}$ for use with a grid containing several different symbols. In this research, the participants ranked the complexity of tertiary grids containing three different colored cells (black, white, and gray) for the purpose of determining the viability of $C_{2,3,4}$ for tertiary grids.

The data show that $C_{2,3,4}$ was a more viable description of the complexity of tertiary grids than either $P(A)$ or $C_{2,5,10}$ for three reasons. First, the data show the importance of including the information contained in 3×3 subgrids in a complexity measure. Second, the data show the importance of using overlapping subgrids in a complexity measure. Finally, the data show that rotating the entire grid affects the perception of complexity.

There were many characteristics used by the participants to categorize the complex arrangement of the tertiary grids. The participants provided more diverse explanations for the rankings of the grids in Series 5 and 6 than in either of the previous two sets of series. This resulted in diverse rankings of the grids by the participants. More importantly, in these two series, the identified characteristics of the grids affected the placement of the grids in the list of increasing complexity. Two characteristics seen in the previous four series were diagonals (i.e., alternations) and “plus signs.” As shown in Series 1 and 2, the 2×2 subgrids were the basis of the alternations (see Figure 24, see also Figures 15 and 16). Jackie, Laura, Lucy, Maria, and MK identified diagonals or alternations as a key characteristic for ranking Series 5 and 6 with respect to complex arrangement.

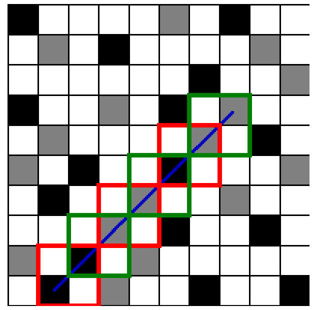


Figure 24. One diagonal and the corresponding 2×2 diagonal subgrids that are alternating.

Alice considered the plus signs the “essential characteristic” by which to rank the tertiary grids. As Table 12 shows, Alice claimed to use only the total number of plus signs to rank the grids. Alice defined a plus sign as

> A: It seems like I notice [plus signs] more when there’s, like, one color in the middle and, like, two [cells] of one color or two of another color, … or maybe, like, three of one color but not three [of one color] in a row.

As seen in Figure 25, the plus signs have a W cell in the center. Although her definition allows another color in the center, most of the plus signs to which she specifically pointed had a W cell in the center and at least three of the four sides of the plus signs needed to be either B or G. The reason that she mainly identified plus signs with a W cell in the center may be due to the large proportion of W cells in the grid. Finally, the various plus signs were allowed to overlap. Figure 25 illustrates examples of plus signs that Alice identified in Series 5 – Grid II. As seen in Figure 25, the plus signs occupy 3×3 subgrids.

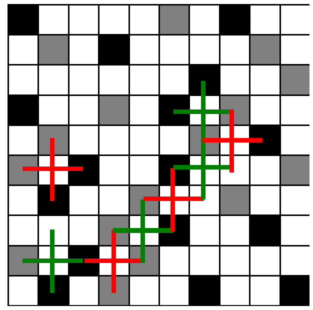


Figure 25. The plus signs identified by Alice in Series 5 – Grid II.

There were other subgrid patterns identified by some of the participants. For example, Jackie identified “pyramids,” i.e., three B cells in a 3×3 subgrid connected on a diagonal forming a triangle (see Figure 26).

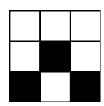


Figure 26. The “pyramid” shape identified by Jackie in Series 5 – Grid V.

The W spaces were an important consideration to Jackie, Lucy and MK. MK considered the W cells within a background of colored (i.e., not distinguishing between B and G) cells when considering the complexity of the grids. Maria considered the colored cells together within a background of W cells without distinguishing between the colors. Erica, MK and Tree considered pairs of the same color (i.e., two B cells or two G cells) to be important. Erica also considered triplets (i.e., three cells connected horizontally and/or vertically of either B or G color) (see Figure 27). Note that in the second figure the triplet consisted of the B cell with the two G cells in the middle row.

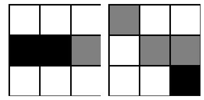


Figure 27. Horizontally- and vertically-connected triplets identified by Erica.

Comparison of Ranking 16/16/84 Tertiary Grids. Series 5 (see Figure 28) and Series 6 (see Figure 29) were the same series of grids but each grid was rotated 90° and listed in a different order. Four of the participants were consistent in their ranking of the two series (see Tables 17, 18, and 19). MK was the only participant to rank the two series in exactly the same order. Maria and Tree switched the two grids they had listed as the least complex. Alice switched her second and third grids in the lists. The remaining three participants had only two grids consistently ranked between the two series. Erica had the same two grids listed as the most complex in both series, whereas Jackie had the same two grids listed as the least complex in both series. Lucy was consistent with the least complex grid and the second most complex grid in both series.

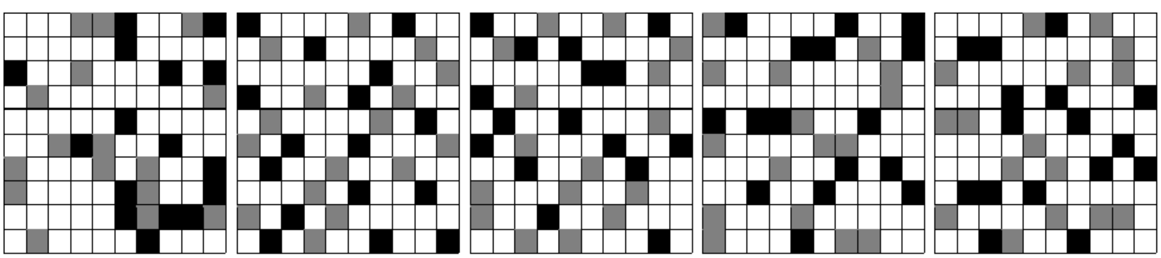


Figure 28. Grids I–V in Series 5.

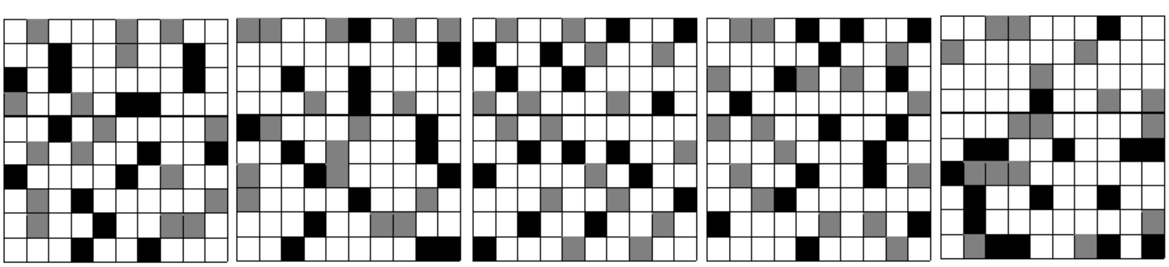


Figure 29. Grids I–V in Series 6.

Table 17. Comparison of proposed $C_{2,3,4}$ measure of complexity with $P(A)$ and $C_{2,5,10}$ measures of the binary grids in Series 5 and 6.

Series 5	Grid I	Grid II	Grid III	Grid IV	Grid V
Mean rank	2	5	4	1	3
Median rank	1.5	5	3.5	1.5	3.5
Series 6	Grid V	Grid III	Grid IV	Grid II	Grid I
Mean rank	1.5	5	4	3	1.5
Median rank	1	5	3.5	3.5	2
$P(A)$	0.461	0.633	0.594	0.511	0.572
$C_{2,5,10}$	12.802	4.844	8.274	12.455	4.194
$C_{2,3,4}$	5.209	5.270	5.613	5.847	5.538

Erica, MK, and Tree were the most consistent among the participants for ranking Series 5 and 6 (see Tables 18 & 19). However, the participants did not seem to use similar characteristics of the cells. Erica mostly considered pairs and triplets of colored cells. Tree did appear to use pairs; however, he claimed that the most important characteristic to rank these grids was symmetry, saying, “If there is symmetry, then there is a pattern and it is not complex.” The next important characteristic for Tree was randomness, i.e., the alternations of the colors. MK considered the blocks of W cells on a colored background. Maria ranked Series 6 similar to Erica, MK, and Tree. Maria identified diagonals of B, G and W cells and the number of plus signs.

Alice and Jackie ranked Series 6 in a similar manner (see Tables 18 and 19). Again, Alice and Jackie did not seem to use similar characteristics of the cells. Jackie considered single cells of B, single cells of G and pyramids formed by a single color. In addition, Jackie identified the large group of colored cells in Grid V. The only characteristic used by Alice was the number of plus signs. Alice claimed, “In the last two series I was looking more at the plus signs formed mostly by the G and B, might be the W but if there is too much of any of those colors, then not a plus sign.”

There were some specific common elements identified by a few participants in some of the grids (see Table 12). For example, in Series 5 – Grids I and IV, MK and Lucy identified large groups of W cells, which was the reason they ranked this grid among the least complex of the grids. In Series 5 – Grid II, Laura, Maria, and MK noticed diagonals of single cells. These three participants all ranked this grid as the most complex of the series. In Series 6 – Grid III, Jackie, Lucy, MK, and Tree all noticed the single cells (i.e., no pairs). However, Jackie and Lucy ranked this grid as the least complex grid, whereas MK and Tree ranked this grid as the most complex.

Of the three types of grids used in the study, the grids with 50 black and 50 white cells were the most consistently ranked grids among all the participants. More specifically, the standard deviation of rankings for the 50/50 grids range from 0 to 1.272 whereas the standard deviation of the rankings for Series 3 and 4 range from 0.548 to 2.191 and for Series 5 and 6 from 0.756 to 1.988. Furthermore, on average, the participants’ ranking for the grids in Series 2 was consistent with their rankings for Series 1 – the only difference between the series was that the B cells were switched with the W cells. More specifically, three participants (Jackie, Lucy, and MK) ranked both series the same in terms of complex arrangement and two participants (Alice and Maria) only switched the third most complex with the fourth most complex grids. Erica moved the least complex grid in Series 1 to the third most complex in Series 2 with the rest of the order remaining the same. Tree ranked only two grids consistently between the two series, his least complex grid and his fourth most complex grid.

Table 18. Comparison of participants' ranks with $P(A)$, $C_{2,3,4}$, and $C_{2,5,10}$ measures of the binary grids in Series 5.

Series 5	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.461	0.633	0.594	0.511	0.572
$C_{2,5,10}$	12.802	4.844	8.274	12.455	4.194
$C_{2,3,4}$	5.209	5.270	5.613	5.847	5.538
Alice	5	1	2	4	3
Erica	2	5	4	1	3
Jackie	4	1	3	5	2
Laura	NR	NR	NR	NR	NR
Lucy	1	3	5	2	4
Maria	4	5	1	3	2
MK	1	5	3	2	4
Tree	2	5	4	1	3

Table 19. Comparison of participants' ranks with $P(A)$, $C_{2,3,4}$, and $C_{2,5,10}$ measures of the binary grids in Series 6.

Series 6	Grid I	Grid II	Grid III	Grid IV	Grid V
$P(A)$	0.572	0.511	0.633	0.594	0.461
$C_{2,5,10}$	4.194	12.455	4.844	8.274	12.802
$C_{2,3,4}$	5.538	5.847	5.270	5.613	5.209
Alice	2	4	1	3	5
Erica	2	3	5	4	1
Jackie	2	3	1	4	5
Laura	NR	NR	NR	NR	NR
Lucy	4	5	2	3	1
Maria	1	3	5	2	4
MK	4	2	5	3	1
Tree	3	2	5	4	1

In summary, the participants used a wide variety of characteristics to rank the 16/16/68 grids. The small spread of $C_{2,3,4}$ values can explain the variation in the participants' complexity rankings. The participants recognized patterns on 2×2 and 3×3 overlapping subgrids, and none of the participants recognized any patterns on the 5×5 and larger subgrids.

Note also that the range of values of $C_{2,5,10}$ for the various types of grids significantly overlap. However, on average the values of $C_{2,3,4}$ for the 16/84 grids were the lowest and for the 16/16/68 grids were the highest. The number of different arrangements for 16/16/68 grids was significantly larger than the number of different arrangements for 16/84 grids. Thus, there was a positive correlation between the number of different arrangements of a grid and the $C_{2,3,4}$ complexity measure of a grid. Because entropy is a measure of the number of microstates for a system, there was a positive correlation between the entropy of a system and the proposed complexity measure, $C_{2,3,4}$.

Measure for Complexity. It seemed that the modified complexity formula, $C_{2,3,4}$, provided a good measure of the perceived complexity of binary and tertiary grids. I have shown that $C_{2,3,4}$ was in close agreement with Falk and Konold's probability of alternations when 50% of the cells are B and the other 50% are W. The $P(A)$ was the current standard for the perception of randomness for 50/50 grids since it has been extensively studied and compared to various other measures (see Falk & Konold, 1994, 1997). I used the complexity measure of Klinger and Salingaros, $C_{2,5,10}$, to develop the formula for $C_{2,3,4}$ through careful consideration of the participants' use of the symmetric properties of the 2×2, 3×3 and 4×4 subgrids. Only in a single case did a participant consider a 5×5 subgrid or larger and this was for one specific grid. Thus, the new complexity measure, $C_{2,3,4}$ was more appropriate than the complexity measure developed by Klinger and Salingaros, $C_{2,5,10}$. Furthermore, $C_{2,3,4}$ describes the complexity of 16/84 grids and 16/16/68 grids.

Key Finding 2: SPSTs Exhibit Subjective or Transitional Conceptions of Probability and Sample Space and Do Not Make Connections between Probability and Sample Space with Randomness

Underlying randomness were three constructs of probability that are important for an understanding of entropy: the probability of an event and sample space (Metz, 1998) as well as the equiprobability of all outcomes (Ben-Naim, 2007, 2008). I will show that the participants demonstrate Subjective (Level 1), Transitional (Level 2) and Informal Quantitative (Level 3) understanding of probability and sample space, but not Numerical (Level 4) understanding

(Jones et al., 1999). Furthermore, I will show that the participants interpret random situations as being completely unpredictable and without exhibiting patterns. According to Metz (1998), this was not a complete understanding of randomness. Metz (1997) argued that an understanding of randomness "demands an integration of the uncertainty and unpredictability of a single event, with an understanding of the patterns that can emerge across repetitions of events" (p. 229). I will show that the participants do not make the connection between uncertainty and long-term patterns. In fact, the participants reason experientially and deterministically when discussing random events. In particular, I will describe the following: first, the participants' conceptions of randomness including their definitions of randomness in specific situations and their use of experiential notions to explain random situations; second, how the participants did not consistently make the connection between sample space and randomness; and three, how the participants did not have a complete conception of randomness (Metz, 1997, 1998).

SPSTs' Conceptions of Randomness

At the beginning of the first interview, each of the participants provided a definition of randomness and an example of a random event. Most of the definitions provided by the participants show that their understanding of randomness was within a deterministic/indeterministic interpretive frame (Metz, 1998). More specifically, the examples suggest the participants were interpreting their examples within the "without patterns/completely unpredictable" dimension of the randomness framework. The exceptions were the examples of random events given by Alice and Erica, which have some deterministic aspects. None of the participants provided either a definition of randomness or an example of randomness that showed an understanding within the randomness dimension as described by Metz (1998). In other words, none of the examples integrated uncertainty with patterns (see Metz, 1998; Moore, 2001; Pratt, 2005).

Erica, Laura and Lucy defined randomness as “no order” or “no pattern” (see Langrall & Mooney, 2000; Metz, 1998; Pratt, 2005). Erica suggested that students moving through the commons area after school was an example of randomness with no pattern. Erica described the example as:

> E: When the bell rings at the end of class, and all the kids are going through the commons area, to me, that’s randomness.
>
> R: Are they [the students] going somewhere with a purpose?
>
> E: They should be going somewhere with a purpose but they are being released from a classroom setting and they are dispersing and they are supposed to be going home but they are just kind of, like, going through this area … in no form or fashion, they are just going through this area.

It seemed that Erica’s interpretation of the event was “Data-Driven Reasoning” (see Metz, 1998) and randomness as “no causality” (see Le Coutre et al., 2006). It seemed that Erica was describing the movement of the students after class as random (i.e., no order) because prior to the bell the students sat in their seats in an ordered manner. Furthermore, within Erica’s explanation, she does not attempt to determine how the interactions among the students determine the movement. In other words, she did not analyze the situation other than basing the movement on prior experience. In addition to random as “no pattern,” Lucy provided a second definition of random as “no connection.” Lucy gave the following example:

> L: Like if you are doing random numbers you know there is like no pattern that you are going to see in a series of numbers … it's not like every third number or something like that there is no connection …

In other words, in a list of random numbers, any number in the list has no connection (i.e., does not determine) the next number. In fact, in a truly random list, all possible groups of numbers of any length are in the list (Attneave, 1959; Metz, 1997; Moore, 2001; Pratt, 2005). Lucy's interpretation of the event was "Data-Driven Reasoning." Lucy based her example on her previous experience and does not consider any long-term pattern or how random numbers were generated in a list. Laura did not provide an example of randomness.

MK shared Lucy's definition of randomness as "no connection." MK explained:

> MK: … Let's see … random. Umm, everything I think of as random, I kind of put it in an arbitrary way not as random. Umm, I'm not sure that is a fair distinction or not …
>
> R: Ok.
>
> MK: Umm, something random, I guess, would be a, somebody happened to leave something on the floor in the exact spot where I happen to be walking and I just happen to be not looking and hit it and tripped … that is what I think a random situation is.

This was an example of randomness as unpredictability (see Langrall & Mooney, 2005; Metz, 1998; Piaget & Inhelder, 1975; Pratt, 2005). MK's interpretation of the event was "Data-Driven Reasoning." He makes no determination of how the outcome was generated other than it was unpredictable. MK mentions "arbitrary randomness" while describing the random event. MK claimed that arbitrary randomness occurs when someone selected something; however, this was not truly random because there was something, perhaps implicitly, in an individual's thinking that affected their selection.

Jackie said randomness was caused by “no preparation” and was “not based on anything.” For Jackie, selecting two socks from a drawer in the dark was an example of randomness. This was different from MK’s arbitrary randomness since Jackie set down prerequisites (e.g., too dark to see socks and selecting the sock quickly) to ensure that there was nothing about the socks that causes her to select them. This definition was randomness as unpredictability (see Langrall & Mooney, 2005; Metz, 1998; Piaget & Inhelder, 1975; Pratt, 2005) because the outcome cannot be predicted. It was also possible that this definition of randomness comes from Jackie’s personal experiences (see Piaget & Inhelder, 1975; Batanero, Henry, & Parzysz, 2005). Thus, her interpretation of the event was “Data-Driven Reasoning.”

Maria explicitly suggested randomness was something that was unpredictable like the location of an electron. Maria argued that even though we know in general where the electron may be, we cannot know its exact location. This was randomness due to incomplete data (see Gal, 2004; Metz, 1998; Polkinghorne, 1984) regarding the location of the electron. Thus, her interpretation of the event was “Indeterminate Physical Model.”

Tree suggested there were two different ideas when talking about randomness. First, a random situation “can be controlled” whereas something that was chaotic is “out of control” due to noise in the system. Tree did not provide a clear example of a random event; however, he was using randomness (i.e., chaos) as unsteerability (see Langrall & Mooney, 2005; Pratt, 2005). The interpretation was “Deterministic Physical Model” for the random situation and “Indeterministic Physical Model” for the chaotic situation.

Finally, Alice admitted that she did not “really know” what randomness was. The example that Alice provided was weather because “it kinda does what it wants.” Thus, her interpretation of the event was "Data-Driven Reasoning" because she bases her interpretation of the event on previous experiences without an analysis of the how the outcomes were determined.

In summary, most of the participants provided and interpreted a random event based on prior experiences with the situation. Metz (1998) claimed the randomness interpretation includes probabilistic reasoning concerning the sample space; however, none of the participants used the concept of probability in their discussion of the random event. In all the examples, the participants did not make the connection between probability and sample space with the random situation. The participants mostly reasoned from their own prior experiences. In addition, in most of the examples, the participants were using the idea that no causality implies randomness (see le Coutre et al., 2006).

SPSTs' Connections between Sample Space and Randomness

SPSTs' Conceptions of Probability

In this section, I will show how the participants' experiences confound their responses to various probability problems. The data show that the participants do not utilize the sample space to determine a probability. Even after a sample space activity and the researcher revisiting the probability problem, most participants still did not utilize the sample space to calculate probability. Most participants were in Level 1 (Subjective Reasoning) or Level 2 (Transitional Reasoning). This will have implications for the third key finding regarding SPSTs' conceptions of entropy.

Casting *n* Dice at Same Time. When casting *n* dice at the same time, there were two distinct answers provided by the participants: first, mathematically all outcomes were equal, but experientially the outcome with different numbers was more likely; second, mathematically all outcomes were not equal, but experientially the outcome with the same number (5, 5) and the outcome with different numbers (5, 6) were equally-likely. Seven of the participants (Alice, Erica, Jackie, Laura, Lucy, MK, and Tree) fell into the first category. They each had difficulty in reconciling why their everyday experience conflicted with their expected mathematical answer. Eventually, they based their response on their experience. In other words, most participants initially argued that all outcomes were equally likely (e.g., 5, 5 and 5, 6 on two dice simultaneously cast) mathematically, but that their experience told them otherwise (that different numbers on each die were more likely to occur). This was “Data-Driven Reasoning” where the prediction is based participants’ experience. None of the participants attempted to list the solution set of all possible outcomes for two or three dice. Instead, the participants argued using the probability of casting any number on a single die. Fischbein and Schnarch (1997) defined this misconception as the confusion of compound and simple events (see also Fischbein, Sainati Nello, & Sciolis Marino, 1991). For example, Lucy stated that the probability of any number to come up was 1/6 and since they (the different die) were separate events, the probability of casting a 5, 6 and 5, 5 were equally likely. Batanero, Navarro-Pelayo and Godino (1997) claimed this misconception may be due to a lack of combinatorial reasoning. Lucy seemed to have the most difficulty reconciling the two answers. After rolling a pair of dice numerous times, she decided that the experiment agrees with her experience. Lucy added, “What do dice know about randomness?” This last interpretation was the “Affordance of Inanimate Objects of Motion” in which “chance is attributed to the assumption that the [dice], as inanimate objects, have no intentionality or internal controls” (Metz, 1998, p. 304).

Maria argued that because all outcomes were equally likely, then the 5, 5 and 5, 6 were equally likely. Maria never mentioned experience as conflicting with the mathematical conclusion. Erica argued that the outcome with the different numbers was more likely “because of the odds." However, she did not know how to calculate the odds for either the two or three dice question. So, I assume that she was reasoning more from experience than mathematically.

Because Alice, Erica, Jackie, Laura, Lucy, MK, and Tree all argued using their common experiences, they exhibited Level 1 Subjective thinking for the constructs Probability of an Event and Probability Comparisons (Jones et al., 1997, 1999). Furthermore, by reasoning from personal experience instead of based on the sample space these participants exhibited the availability misconception (Fischbein, Sainati Nello, & Sciolis Marino, 1991; Fischbein & Schnarch, 1997). I claim that even though Erica recognized the odds of casting a 5 and 6 was greater than the odds of casting two 5s, she showed evidence of Level 1 Subjective thinking because she did not know the odds nor how to calculate the odds. Maria also appeared to show signs of Level 1 Subjective thinking for these constructs despite using quantitative judgments.

Casting One Die Three Times. Again, the participants did not attempt to consider all the possible outcomes. The difference here lies in fact that the order of the outcomes of the die was important. Erica, Laura, and Lucy immediately and explicitly identified the importance of the order of the outcomes on the die. The other participants implicitly understood the difference in the order as identified after some questioning from the researcher. Most participants (Alice, Jackie, Lucy, Maria, and MK) restated their same arguments for the ordered situation as in the previous unordered situation. In other words, although the outcomes of casting all the dice at once (unordered) versus casting a single die multiple times (ordered) were different, the participants did not change their reasoning between the two questions. So, these participants exhibited Level 1 Subjective Reasoning for casting dice.

Lucy and Erica claimed that any pattern of three dice (even 5, 5, 5) was possible if there were enough attempts. This showed that Lucy and Erica have some understanding of the Law of

Large Numbers. Erica and Tree used quantitative judgments to calculate the probabilities and to compare the probabilities. However, they based their quantitative judgments on "more than" 50% (Erica) and "less than" 50% (Tree). Thus, because Erica and Tree argued quantitatively (even though both quantified incorrectly), they appeared to demonstrate Level 2 Transitional Thinking in the constructs of Probability of an Event and Probability Comparisons.

SPSTs' Combinatorial Reasoning

Three activities attempted to address the participants' understanding of combinations. The first was determining the number of three-letter strings that can be formed from a list of five letters. In interview 3, the participants needed to find all possible microstates and the most likely microstate within a given macrostate. Finally, in interview 4, the participants were asked how many ways x red tokens and $16 - x$ white tokens can be distributed among 16 squares on a 4×4 board.

Every participant knew there was some formula that could determine the number of combinations; however, none of them was able to remember the formula in any of the cases. The strategies used by the participants to generate the list of items were similar in most cases. In a few cases, the researcher provided a hint to start with a list (Erica and Tree) or encouraged the use of the tokens (MK and Tree).

In this section, I will show that most of the participants (except Erica) did not completely generate the sample space. The participants were able to generate partially the sample space, with most using some sort of strategy. Thus, most participants (Alice, Jackie, Laura, Lucy, Maria, and MK) exhibited Level 3 Informal Quantitative Reasoning. This will have implications for the third key finding of this research.

Strings of Ordered Letters. Some participants immediately realized they could list all possibilities (Jackie, Laura, Lucy, Maria, and MK). Laura and Jackie initially found the number of combinations formed by three different letters. Once Laura found that there were six combinations of the letters *a*, *b*, *c*, she applied the same reasoning to the other eleven three-letter strings without having to list any additional strings. This led to an incomplete 66 combinations total. Once the researcher reminded Laura that strings could contain the same letter more than once, she considered the additional strings in which a letter appeared twice and concluded with the 120 combinations total. She did not consider the five strings in which all the letters were the same, e.g., *aaa*. Jackie started in a similar manner, but was not able to extend the six combinations of *a*, *b*, *c* to the rest of the letters. Jackie stated, “you have to do something with the number six and maybe the number 5, because there is five [letters].”

Erica wrote out a short incomplete list without the appearance of a strategy to generate all the combinations: *cdc*, *bce*, *cbe*, *abc*, *acb*, *bca*, *bac*. Erica abandoned this list and started a new one: *abc*, *bcd*, *cde*, *eab*. In the middle of writing the list, she guessed 5^3, but was not convinced – she thought there were more than 125 combinations of strings of three letters formed from five different letters. Erica lamented, “It has to be more than 125 because there were so many ways you can use the letters twice or three times. I want to write more than 125 of three sequences [i.e., three letters in a sequence]. There is a formula. I wish I knew.” MK after a brief moment stated, “there are five possibilities for each slot, so $5^3 = 125$.” Lucy listed all strings that begin with *a* starting with *abc*, and then listed all strings that begin with *b* starting with *bcd* and ending with *bda*. Lucy continued through the list so that she has listed all strings starting with *a*, *b*, *c*, *d*, and *e*. Alice starts the list with *abc*, *acd*, and *ade*. Alice did not complete the list but concluded that the formula needed to calculate the total number of strings involves a factorial. Neither Alice nor Lucy included strings that contained any letter more than once, which Batanero, Navarro-Pelayo, & Godino (1997) called an error of repetition. Maria started with a systematic list: *abc*, *abd*, *abe*, *acb*… and extended the list by changing the third letter. Once all strings starting with the letter *a* were exhausted she would change the first letter to *b* and continue the process. Once the researcher reminded Maria that strings could contain the same letter more than once, Maria restarted the list with the same process: *aaa*, *aab*, *aac* … Tree, without starting any list of possible strings, guessed the number of combinations was $3!$ then changes the answer to $5!$ after answering the next question on unordered strings.

Strings of Unordered Letters. Most of the participants said they would start from the list of ordered strings they listed for the previous question and cross out strings that counted more than once. For example, the two unordered strings, *abc* and *cba* were the same string because they had the same letters. The participant would delete one of these strings from the list so as not to count the string more than once. Tree was an exception. Tree initially thought the ordered strings had $3!$ different combinations. However, Tree knew that the unordered list of strings would have fewer combinations. So, he concluded that there must be $3!$ unordered strings and $5!$ ordered strings. MK was the other exception. MK argued that any of the five letters can be first, and then there were four letters that were available for the second position and three for the last position. MK concluded there were 60 combinations, which was an over-count of the total number of unordered strings.

Alice, Jackie, Laura, Lucy, Maria, and MK demonstrated Level 3 Informal Quantitative Thinking for Sample Space in this framework. Each of these participants only partially generates the sample space using a generative strategy. Erica exhibited Level 2 Transitional thinking because she did not use a systematic strategy to generate the sample space and only generated a partial sample space. The framework was not valid for Erica and Tree because they did not attempt to generate the sample space.

After the participants answered the question on finding the number of combinations of three-letter strings, they revisited the casting of three dice questions. Four of the participants (Jackie, Laura, Maria, and MK) stated that they would not change their answer to this question. For example, Maria claimed that she "did not know how [the sample space formed by three letters selected from five letters and the dice toss] are connected." Alice and Lucy did see the connection between the sample space question and the probability question. They provided new answers to 2a and 2b (5, 6 and 5, 3, 6, respectively). However, Alice still did not distinguish the order of the die throws in question 2c. Also, Lucy was confused about whether order was important for question 2c, so provided an answer if order was important (all were equally-likely) and an answer if order was not important (5, 3, 6). Erica and Tree did not revisit this question because they did not generate enough of a list of strings to which they could refer.

None of the participants made any connection between the unordered outcomes of the dice question and the list of strings generated in question 5. In other words, this may be the reason why the participants did not use sample space to find probabilities when discussing the more probable outcome for rolling two or three dice in Interview 1. Additionally, all the participants exhibited the error of order (Batanero, Navarro-Pelayo, & Godino, 1997) by not recognizing the difference between three dice cast at the same time and one die cast three times.

<u>SPSTs' Spatial Representation of Molecules in Integral Energy Levels.</u> For this activity, each energy level, E_n had n units of energy. For simplicity, each system was labeled as $N = x$, $E = y$ which refers to a system with x molecules and a total amount of energy of y units. After reviewing a couple of small systems to make sure the participants understood the activity, the

participants were asked to find all possible microstates for the following systems: a. $N = 3$, $E = 3$; b. $N = 3$, $E = 5$; c. $N = 5$, $E = 3$ or $N = 4$, $E = 3$. A large bag of colored tokens was on the table available to the participants.

Lucy did not use any tokens but was able to list all possible microstates, to list the number of combinations for each microstate, and to determine the most likely microstate(s). Lucy systematically listed the microstates by starting with one molecule in the highest energy level and the rest in E_0. Then Lucy moved the molecule in the highest energy level to the next lowest and moved one molecule from E_0 to E_1. She continued this process until she found all possible microstates. Most of the other participants (Alice, Erica, Jackie, Laura, Maria, and MK) approached the question in a similar manner, but Laura listed them in a more haphazard manner. Maria and MK missed one microstate in the system $N = 3$, $E = 5$; Erica missed one microstate in the system $N = 4$, $E = 3$, and Tree missed at least one microstate in each system.

Erica and Maria were not able to determine the number of combinations for each microstate and thus, were unable to determine definitively the most likely microstate. Erica did guess which microstate was most likely, but not consistently. For example, Erica claimed that the microstate $E_0 = 2$, $E_2 = 1$ (three total combinations) was more likely than the microstate $E_0 = 1$, $E_1 = 1$, $E_2 = 1$ (six total combinations) in the system $N = 3$, $E = 3$. However, Erica also concluded that the microstate $E_0 = 2$, $E_5 = 1$ was less likely than the microstate $E_0 = 1$, $E_2 = 1$, $E_3 = 1$ in the system $N = 3$, $E = 5$.

Several participants did not correctly count all possible combinations. Laura double counted the number of combinations in the microstate $E_0 = 3$, $E_1 = 1$, $E_2 = 1$ in the system $N = 5$, $E = 3$. MK was unable to count the number of combinations in the microstate $E_0 = 2$, $E_1 = 3$ in the system $N = 5$, $E = 3$ but did estimate there were less than 20 combinations and guessed there were about 8 combinations. Alice systematically used the tokens to count the number of combinations in the microstate $E_0 = 2$, $E_1 = 3$ in the system $N = 5$, $E = 3$ concluding there were 9 combinations while missing the remaining combination. In addition, Alice over counted by two the number of combinations in the microstate $E_0 = 2$, $E_1 = 1$, $E_2 = 1$ in the system $N = 4$, $E = 3$.

Erica, Maria, and Tree did not attempt to count the combinations of any of the microstates. For the participants who did count the number of combinations, there were a few errors of omission (Batanero, Navarro-Pelayo, & Godino, 1997) with two specific types of microstates: for example, the microstates $E_0 = 3$, $E_1 = 1$, $E_2 = 1$ and $E_0 = 2$, $E_1 = 3$ in the macrostate $N = 5$, $E = 3$. These two microstates required more than a simple movement of a single token on the worksheet. The participants needed a more detailed strategy to determine how to rearrange the tokens systematically. For example, counting the number of combinations of the microstate $E_0 = 3$, $E_1 = 1$, $E_2 = 1$ was more challenging to the participants than finding the number of combinations of the microstate $E_0 = 1$, $E_1 = 1$, $E_2 = 1$ in the macrostate $N = 3$, $E = 3$. In general, if there were at least three occupied energy levels and at least two molecules in any of these energy levels, then finding the number of combinations becomes a more difficult problem to solve (see Batanero, Navarro-Pelayo, & Godino, 1997).

Jackie and Lucy showed evidence of Level 4 Numerical thinking in the Framework for Probabilistic Thinking for Sample Space (Jones et al., 1997, 1999). Both Jackie and Lucy used a systematic strategy to generate the complete sample space by finding all possible microstates and count the combinations of the microstate. All the other participants (Except Tree) demonstrated Level 3 Informal Quantitative thinking in this framework. Either the participants missed identifying a microstate or only partially generated the various different forms of the microstate using a generative strategy. The framework was not valid for Tree because his answers showed that he did not fully understand the question.

In summary, the data showed that most participants were able to construct the sample space for very small systems. Two of the participants (Tree and Laura) had difficulty with distinguishable vs. indistinguishable objects when counting the sample space.

SPSTs' Incomplete Knowledge of Randomness

In summary, none of the participants exhibited knowledge of randomness as uncertainty with patterns (see Metz, 1998) in the long run. Three participants (Erica, Laura, and Lucy) defined randomness as no order, a definition that seemed to contradict the actual definition of randomness. Furthermore, the participants used various definitions of randomness, each influenced by their personal experiences. This caused most of the participants to interpret random situations through Data-Driven Reasoning and not interpreting the situation as random – "probabilistic reasoning concerning configuration of outcomes" (Metz, 1998, p. 304).

Key Finding 3: SPSTs Have an Incomplete or Instrumental Knowledge of Entropy

Ben-Naim (20077, 2008) claimed that basic computations of probability and a basic understanding of sample space were all that was needed for understanding entropy. So, in this section, the data showed that no participant had a relational understanding of entropy. In other words, they did not make the connection between entropy and probability and sample space in order to explain why the smell of cooking must necessarily move from one side of the room to another, the second law of thermodynamics, or dynamic equilibrium of systems. Instead, the participants provide incomplete explanations based on their previous experiences.

Participants' Conceptions of Entropy

The questionnaire included a question about why you can smell food cooking from another room. Four (Alice, Erica, Lucy and Maria) of the seven participants (MK did not complete the questionnaire) explicitly referenced the gas particles that composed the smell of the food. Even so, Tree identified the gaseous nature of the food smell, and Jackie mentioned the fumes from the food and the molecules of the air. Laura did not refer to the particulate nature of any substance. This does not mean that Laura does not have a conception of the particulate nature of gases, but she did not use the properties of the gaseous particles to explain the smell.

Four of the participants (Alice, Erica, Laura, and Maria) explicitly used the word diffusion in their explanation as to why we can smell food cooking from another room. Two more participants (Jackie and Lucy) mentioned the movement of the food smell. Erica and Laura defined diffusion as the movement from the most concentrated to the least concentrated. Lucy and Maria described diffusion as the spreading of the molecules to fill the container. Alice and Jackie suggested the diffusion was the movement of a gas throughout the room. The explanation provided by Tree was that the "mixture of different gases equalize with time." At the beginning of the next interview, each participant answered why the direction of the movement of the food smell must necessarily be in the particular direction they suggested. All of the participants (except Tree) suggested that is how diffusion works. However, none of the participants was able to explain why diffusion works in the manner they suggested. This suggests the participants have an instrumental understanding (see Skemp, 1977) of diffusion and did not make the connection between randomness, probability, sample space, and entropy.

At the beginning of interview 1, each participant provided a definition of entropy. There were five different definitions used by the participants. First, Erica and Laura claimed that entropy is disorder resulting from the random motion of molecules (Entropy as Random Motion). Erica and Laura mentioned only the random directions of the molecular motion and not the variation in molecular speeds. Second, Alice and Tree suggested entropy is a measure of the disorder of things (Entropy as Disorder). Entropy as Disorder measures the current state of a system. Third, Lucy and MK define entropy as the tendency to be disordered (Entropy as Tendency for Disorder). In other words, molecules tend to move away from order to disorder or chaos. Fourth, Jackie mentioned that entropy had something to do with ΔS (Entropy as a Formula) but was not able to provide any additional information about ΔS. It was possible that Jackie learned entropy in a procedural manner and thus remembers entropy as a procedure or formula (see Gal, 2004). Lastly, Maria admits that she does not know the definition of entropy. However, Maria did suggest that entropy had something to do with energy (Entropy as a Form of Energy).

This was an extremely important result. This research identified five different definitions of entropy held by SPSTs, none of which included the concept of microstates. Even though several participants included the particulate nature of matter, all participants were unable to connect entropy to the diffusion of gases, the Maxwell-Boltzmann distribution, or the apparent equilibrium of matter. The variety of entropy definitions leads to different methods they would use to present the concept of entropy to secondary students.

The last key idea in this study was the participants' understanding and perception of entropy. The formal Boltzmann definition of entropy is $S = k \ln(W)$, where W is the number of combinations of a system. Thus, one essential idea to understanding entropy is determining the number of combinations for a system. It is important to know that the number of combinations greatly increase as n becomes large. Actually, the calculations are prohibitively large; however, it will suffice if the SPSTs grasped the magnitude of the number of combinations.

In the fourth interview, the participants considered Ben-Naim's model of entropy (2007, 2008). The model used in the interview consisted of a 4×4 board covered with 16 bi-colored tokens (red on one side and white on the other side) with the white side showing. The participants rolled two four-sided numbered dice and one six-sided bi-colored die (three red sides and three white sides). The tetrahedral dice determined the cell in the 4×4 board and the cubic die determined the color of the token. The participants were asked to predict the outcome of casting the three dice. They also rolled the dice and answered questions regarding the outcomes. The participants made their predictions and answered questions for approximately ten rolls.

There were several key ideas discussed within this activity. One key idea was the number of combinations formed by x white tokens and $(16 - x)$ red tokens. The participants were familiar with determining the number of combinations for $x = 0, 1, 15, 16$. When there were two tokens of the same color, the participants were unable to find the number of combinations. For example, Jackie and Tree suggested there were more than 16 combinations, Maria concluded there were more than 24 combinations, Laura stated there were 32 combinations, and Alice suggested 44 combinations. However, no participants were clear about how large the number of distributions actually is. For example, Alice concluded the number of distributions of eight red and eight white tokens was "a very large number" and suggested that this number was in the thousands. The actual number of distributions was 12 870. Thus, the participants did not seem to understand how quickly the number of combinations, $\binom{n}{k}$, becomes large as n gets large (see Batanero, Navarro-Pelayo, & Godino, 1997). The number of combinations is extremely large, $>>10^{300}$, for systems with 1000 molecules (Ben-Naim, 2007). Recognition of the extremely large sample space for a molecular system is one of the key ideas needed for conceptual understanding of entropy and dynamic equilibrium (Ben-Naim, 2007, 2008). The participants were unable to recall the formula they know to exist that will easily calculate the answer. In addition, the participants did not use any sort of systematic method to find the solution. All of the participants exhibited Level 1 reasoning (Subjective Reasoning) for Sample Space.

Most of the participants qualitatively recognized that the maximum number of combinations occurred when there were equal numbers of red and white tokens. The exceptions were Tree who claimed about four tokens of either color will have the maximum number of combinations and Erica who guessed the maximum number of combination occurs when there were two or three tokens of either color.

All the participants (except Erica, Lucy, and Tree) were asked to predict the number of each color that would be on the board after n steps (i.e., tosses). In general, each participant claimed there would be approximately $n/2$ tokens of one color after n steps until the board showed the maximum of eight tokens of a single color. All the participants argued that it was more likely for the color that appears most on the grid to change than the other color to change because of a higher probability. No participant argued that the color with the smaller amount on the grid would not change, but that it was not as likely to change. In other words, when the grid had three red tokens and thirteen white tokens, the participants suggested that with the next toss of the dice, one of the red tokens have a smaller chance of changing to white than either a white token changing to red or any token remaining the same color. However, the participants were not able to quantify this probability for any step but the first. Thus, each participant appears to exhibit Level 3 Informal Quantitative thinking for the Probability Comparisons in the Probability Framework (Jones et al., 1997, 1999).

While completing the entropy activity, each participant shook the dice. When asked why they shake the dice, the most common reply was “that’s what you do with dice.” When prompted further, a few participants provided additional reasons: “Shaking provides variability” (Laura), “dice are more fair when tossing high” (MK) and “because the randomness [in the toss] provides new rules for variation in the results” (Maria). Each of the participants exhibited Level 2 (Transitional Thinking) or Level 3 reasoning (Informal Quantitative Reasoning) for Probability Comparisons.

Once the participants completed several steps of the activity, they determined the system with the highest and lowest entropy in four different series. In the first two series, the participants ranked the systems consistently (see Tables 20 and 21). In the last two series, the rankings of the participants diverged (see Tables 22 and 23). Alice ranked Interview 4, Series 1 – Grid II and Interview 4, Series 4 – Grids I and II differently than the other participants. The main reason was that she used the same characteristics that she used while ranking the grids during Interview 2. In other words, in Interview 4, Series 1 and 2, Alice looked for the number of pairs of black cells and in Interview 4, Series 3 and 4, she looked at the number of plus signs that appeared in the grids.

Erica and Tree also considered the patterns in each of the grids. Both recognized pairs of colors and groups of colors. For example, Erica ranked Interview 4, Series 2 – Grid III as the least complex because there was a string of seven B cells connected near the bottom right of the grid. Tree identified Interview 4, Series 4 – Grid I as the least complex because there was a group of four B cells and a string of five B cells. Erica and Tree claimed these groupings and pairs showed evidence of more order and less entropy.

In Interview 4, Series 2 – Grid III, Lucy actually counted the number of cells and chose the grid with the ratio of B:W closest to 50:50 as having the maximum entropy. She used the entropy activity as her explanation for this choice. Lucy was the only participant who explicitly connected the entropy of the systems to the activity. This does not mean to say that this was the only correct method of evaluating the entropy of systems. As Styer (2000) has shown, the spread of a system was not a sufficient method to describe the entropy of that system.

In summary, the participants demonstrated incomplete or instrumental knowledge of entropy. The data has shown that for very small systems of molecules (i.e., 4-5 or less), most of the participants were able to construct the sample space for a system of energy levels. However, the participants were not able to calculate or estimate the number of combinations, e.g., $\binom{16}{5}$, in the entropy activity. Any approximation provided by the participants underestimated the actual value.

The participants did not make the connection between sample space and randomness as determined by the data collected during the interview activities involving dice. More specifically, they did not make the connection between sample space and rolling dice. In particular, the initial explanation given by most participants for shaking two dice was because "that's what you do with dice." Their previous experience with dice overshadowed any other explanation until the researcher probed further. Furthermore, the participants did not make connections between sample space and diffusion as seen by their responses to the question involving the explanation of the spread of the smell of cooking food.

Table 20. Comparison of proposed $C_{2,3,4}$ measure of complexity with combinations, $P(A)$, and $C_{2,5,10}$ measures of the binary grids in Interview 4 – Series 1.

Series 1	I	II	III
B	45	50	37
W	55	50	63
nCr	6.15×10^{28}	1.01×10^{29}	3.42×10^{27}
$P(A)$	0.494	0.533	0.400
$C_{2,5,10}$	6.844	3.035	3.339
$C_{2,3,4}$	3.336	3.224	3.230

Table 21. Comparison of proposed $C_{2,3,4}$ measure of complexity with combinations, $P(A)$, and $C_{2,5,10}$ measures of the binary grids in Interview 4 – Series 2.

Series 2	I	II	III
B	17	9	26
W	83	91	74
nCr	6.65×10^{18}	1.90×10^{12}	7.00×10^{23}
$P(A)$	0.300	0.189	0.383
$C_{2,5,10}$	5.687	5.071	6.673
$C_{2,3,4}$	2.652	1.898	2.792

Table 22. Comparison of proposed $C_{2,3,4}$ measure of complexity with combinations, $P(A)$, and $C_{2,5,10}$ measures of the binary grids in Interview 4 – Series 3.

Series 3	I	II	III
B	19	16	13
G	17	17	17
W	64	67	70
nCr	1.70×10^{37}	3.44×10^{35}	3.52×10^{33}
$P(A)$	0.506	0.500	0.506
$C_{2,5,10}$	13.389	12.858	13.104
$C_{2,3,4}$	5.587	5.810	5.590

Table 23. Comparison of proposed $C_{2,3,4}$ measure of complexity with combinations, $P(A)$, and $C_{2,5,10}$ measures of the binary grids in Interview 4 – Series 4.

Series 4	I	II	III
B	25	18	16
G	18	10	15
W	57	72	69
nCr	2.32×10^{40}	6.56×10^{31}	1.99×10^{34}
$P(A)$	0.589	0.422	0.506
$C_{2,5,10}$	8.800	12.167	12.578
$C_{2,3,4}$	6.675	4.071	4.208

CHAPTER 5

SUMMARY AND CONCLUSIONS

In this chapter, I provide a brief summary of the study and its conclusions. I follow this with a discussion of the implications of the study for post-secondary students and, in particular, secondary preservice science teacher education. Lastly, I offer a discussion of the limitations of the study and recommendations for future research.

Summary

Research suggests that students of all ages have difficulties in moving from a macroscopic worldview to a microscopic worldview. My own experience teaching secondary chemistry has confirmed these findings. Furthermore, several university physics professors proposed changes to post-secondary physics to help students in first-year university physics classes make the connection between the molecular properties of substances, sample space, randomness, and entropy.

The purpose of this research was to describe how post-secondary students understand randomness and how they used probabilistic and statistical reasoning to generate explanations of the properties of gases (i.e., complex systems), especially entropy. In particular, this study focused on SPSTs' conceptions of the pattern aspect of randomness using two-dimensional grids and their conceptions of randomness as a process.

The research questions that guided the study were

1. How do SPSTs interpret randomness in two-dimensions (i.e., randomness as a pattern)? In other words, how do SPSTs determine whether a figure showing the molecules of one or two gases exhibits randomness? More specifically, the data from the SPSTs' interpretations of randomness will support the development of a measure describing the complexity of two-dimensional representations of molecules.
2. How do SPSTs make sense of entropy (i.e., randomness as a process)?
3. How do SPSTs make the connection between sample space (i.e., combinations), randomness, and entropy?
4. What are the salient features of the complex physical phenomena that SPST in chemistry/physics use to determine the definition of randomness needed to conceptualize random phenomena?

The SPSTs' conceptions of probability and sample space were analyzed using the Probabilistic Thinking Framework (Jones, Langrall, Thornton, & Mogill, 1997, 1999). This framework differentiated four levels of individuals' conceptions in four areas: probability, probability comparisons, sample space, and conditional probability. Each of these four areas is important for a conceptual understanding of entropy (see Ben-Naim, 2007, 2008). In addition, SPSTs' conceptions of randomness were analyzed using the Conceptual Framework for Understanding and Attribution of Randomness (Metz, 1998). I used this randomness framework in two important ways: first, the framework included coding for the participants' interpretations of random events; second, the framework groups the interpretations of random events into deterministic, unpredictable with no patterns, and unpredictable with patterns over the long run. The latter was the only interpretation in which a participant recognizes randomness.

A major finding of this study was a proposed measure of the complexity, $C_{2,3,4}$, of two-dimensional binary and tertiary grids. This measure was based on the data collected during this study and was based on Attneave (1959) and on the complexity measure proposed by Klinger and Salingaros (2000). It extends previous research conducted by Falk (1975), Falk and Konold (1994, 1997) and Green (1989). Specifically, the research has identified 23 different characteristics the participants used to determine the degree of the complexity of the two-dimensional grids. Most of these characteristics were described by the proposed complexity measure; otherwise, they were described as potential outliers.

Another key finding was the level of probabilistic thinking exhibited by the participants. There were several different activities, which determined the SPSTs ability to construct the sample space. For this construct, most of the SPSTs were able to create the sample space; however, for even a moderately large sample space (e.g., the number of combinations of two red and fourteen white tokens on a 4×4 grid), they were unable to completely describe the sample space. The SPSTs consistently underestimated the number of combinations of tokens as the number of red tokens increased. Lastly, some SPSTs were either unable to provide a reason for the largest number of combinations occurring with eight red and eight white tokens or they chose another arrangement of red and white tokens as having the most combinations.

There were several different definitions of complexity and randomness held by the participants. Most of the interpretations of random events were based on the prior experiences of the individuals. Furthermore, the interpretations of random events were classified as either deterministic or as unpredictable without patterns. In other words, none of the participants interpreted a random event as unpredictable with patterns.

The last major finding was the incomplete knowledge of entropy held by SPSTs. There were several misconceptions regarding entropy held by the SPSTs. First, the participants who claimed to know the second law of thermodynamics also stated that entropy would continue to increase without end without any discussion of open or closed systems. Second, the participants were unable to correctly identify the number of combinations for relatively large n and k (e.g., $\binom{16}{3}$). Third, and more importantly, the participants were unable to describe how large the number of combinations becomes as n or k become larger (e.g., the participants significantly underestimated the value of $\binom{16}{8}$). Finally, none of the participants showed they understood how probability, sample space, randomness, and entropy were connected.

Conclusions

Proposed New Complexity Measure

One major conclusion was the proposed complexity measure, $C_{2,3,4}$, for 10×10 grids. This measure was based mainly on the complexity measure suggested by Klinger and Salingaros (2000). $C_{2,3,4}$ has been shown to be consistent with most of the conclusions of Falk and Konold's probability of alternations, $P(A)$ (1994, 1997) for binary grids with a 1:1 ratio of colored cells. However, this research has identified probability of alternations to be one of 23 distinct characteristics of two-dimensional grids that SPSTs used to determine the complexity of binary and tertiary grids. In addition, $C_{2,3,4}$ provides an improved description of tertiary grids over either $P(A)$ or $C_{2,5,10}$ because it was based on the data collected in the study and was also based on Attneave's work with binary series (1959). It seemed that the modified complexity formula, $C_{2,3,4}$, provided a good measure of the perceived complexity of binary and tertiary grids.

SPSTs' Exhibit Subjective or Transitional Conceptions of Probability and Sample Space

The participants demonstrated Subjective (Level 1), Transitional (Level 2) and Informal Quantitative (Level 3) understanding of probability and sample space but not Numerical (Level 4) understanding (Jones, et al., 1999). Also, the participants interpreted random situations as being completely unpredictable and without exhibiting patterns and did not consider long-term patterns of truly random events (see Metz, 1997, 1998). Furthermore, the participants reasoned experientially and deterministically when discussing random events and did not make any connection between sample space and randomness. Most of the participants interpreted random situations through Data-Driven Reasoning and not interpreting the situation as random – "probabilistic reasoning concerning configuration of outcomes" (Metz, 1998, p. 304). Thus, the participants did not have a complete conception of randomness (see Metz, 1997, 1998).

Post-secondary Students Have an Incomplete Understanding of Entropy

The participants' subjective views of randomness affect their understanding of entropy. Ben-Naim (20077, 2008) claimed that only basic computations of probability and basic understanding of sample space were all that was needed for understanding entropy. However, the participants did not make the necessary connections among entropy and probability and sample space. For example, SPSTs were unable to calculate sample sizes for relatively small sample spaces. This adversely affected their knowledge of the sample space for larger molecular systems in which they significantly underestimated sample size. Thus, the participants provided incomplete explanations, based on their previous experiences, of events that are better explained by entropy.

Implications for Science Education

One of the findings of this study is that post-secondary science students do not make the connections among the concepts of probability, randomness, and entropy. This suggests that post-secondary physical science professors need to develop strategies to present the concept of entropy to undergraduates in a comprehensive manner. In other words, they need to emphasize the particulate nature of matter, call attention to the large number of microstates for a system, compare the probability of the various microstates, and address the random movement of the system through the various microstates. More importantly, they need to help post-secondary students understand how these concepts relate to the entropy of a system. Another implication of this study SPSTs need to be better prepared in statistical concepts, especially in probability of an event, probability comparisons, and calculating size of sample spaces (i.e., combinations).

Need to Carefully Represent Molecular Systems

SPSTs suggested 23 different characteristics that affect their perceptions of the complexity of the grids. Any two-dimensional representation of a molecular system should take into account these 23 characteristics and a measure of the complexity ($C_{2,3,4}$) of the grid. It is important to construct the molecular representation so that there were no unnoticed characteristics that may be perceived by students. So, when teacher educators, post-secondary physical science professors and textbooks refer to a molecular system, there needs to be some careful consideration of the non-random patterns that may be unwittingly included in the representation.

Need for Discussion of Statistics in Science Education

Secondary teacher education programs in science need to better prepare its students to teach science concepts that have a statistical basis. NSTA (2003) argued that secondary science teachers have a supporting competency in statistics. Lin, Cheng and Lawrenz (2000) argued that teachers should have knowledge of how the particulate nature of matter underlies the macroproperties of substances. However, a conceptual understanding of the particulate nature of matter, especially its connection to entropy and the second law of thermodynamics, requires a few statistical concepts including the probability of an event, probability comparisons, and sample space (see Ben-Naim, 2007, 2008). Secondary mathematics and science teachers may be unprepared or under prepared for teaching statistics (see Pfannkuch & Wild, 2004; Stohl, 2005). Teachers may have poor understanding of probability or hold some misconceptions about probability concepts (Jolliffe, 2005; see also Stohl, 2005). The data from this research suggests that SPSTs may be unprepared or underprepared for teaching entropy, the second law of thermodynamics, dynamic equilibrium, or kinetic theory. These concepts need to be included somewhere in the preservice science education curriculum. There are various places in the curriculum where these concepts are appropriately introduced: physical chemistry or thermodynamics courses.

Continued Effort to Make Statistical Mechanics and Entropy Accessible to First-Year Physical Science Students

Veal suggested “one of the most difficult aspects of ‘learning to teach’ is making the transition from personal beliefs about content to thinking about how to organize and represent the content … in ways that will facilitate student understanding” (2004, p. 329). The data collected during this research has shown that the SPSTs did not have a relational understanding of entropy and its probabilistic underpinnings.

There have been suggestions to make statistical mechanics accessible to first-year university level physics students (see Fuchs, 1987; Lee, 2001; Moore & Schroeder, 1997; Reif, 1999; Smith, 1993). All SPSTs are required to take introductory physics as part of their course of study. So, it is not unreasonable to assist them to understand statistical mechanics "if we apply a proper method" (Lee, 2001, p. 68). Wilensky has concluded that secondary students are capable of understanding statistical mechanics (1999a; see also Wilensky, Hazzard, & Froemke, 1999).

“Both mathematics and statistics educators recommend instruction that is grounded in concrete physical activities to help students develop an understanding of abstract concepts and reasoning” (delMas, 2004, pp. 91-2; see also Ben-Naim, 2007, 2008; Styer, 2000); however, the data from this research showed that the SPSTs may still not make the connection between the activity and what the activity is modeling. Lee claimed, “it is not so difficult to teach the statistical foundations of thermal physics in introductory college physics courses if *we know how to teach statistical concepts*” (2001, p. 75, emphasis mine). So, there is a need to guide the SPSTs with appropriate methods for teaching the concept. This includes knowledge of how to teach counting procedures and in finding probability from a sample space. Even though all the participants knew there was some formula, none of the participants recalled the formula for finding the number of combinations, $\frac{T!}{(T-N)!}$. Knowing this formula may help the participants in recognizing how large the number of combinations becomes as T and N increase (see Batanero, Navarro-Pelayo, & Godino, 1997). There is a danger in approaching the statistical basis of entropy procedurally (see Gal, 2004). The presentation may be possible (see Wilensky, 1999a), especially if the suggestions of Wilensky (1999a) and Ben-Naim (2007, 2008) are valid.

Known Limitations of the Study

The first limitation was due to the sample. The sample was small and not representative of preservice secondary science teachers. Although all the participants provided detailed information on what characteristics they used determine the complexity of the grids, there are several questions that arise from this research. For example, the participants in this study were selected from four distinct groups of SPSTs: non-traditional SPSTs who are returning to school to prepare to be secondary teachers after graduating with a degree in physical science several years prior to the study, traditional SPSTs who recently enrolled in statistical mechanics but who had little or no other exposure to statistics, SPSTs in biology who claimed to have had more exposure to statistics but who did not previously enroll in thermodynamics or statistical mechanics, and two participants who did not fit any of these groups. The wide range of exposure to statistics, or lack of exposure to statistics, did not provide any insight into differences among the participants.

The second limitation concerns the type of grids used. The research used only three types of grids all with the same 10×10 size. 10×10 grids were used so that the data collected in this study could be compared to data collected in previous research (see Falk, 1975; Falk & Konold, 1994, 1997; Green, 1989). However, molecular systems have been represented using smaller grids, e.g., an approximate 6×6 grid (see Whitten, et al. 2004), and using larger grids, e.g., 40×40 grids (see Atkins, 1994).

Further Research

Study to Determine Viability of $C_{2,3,4}$

First, because this was a small scale study, there is a need to verify the appropriateness of $C_{2,3,4}$. There have been many previous studies on individuals' conceptions of randomness (Attneave, 1959; Falk & Konold, 1994, 1997; Feldman, 2004). A larger study involving more participants may provide additional insight into how the characteristics that participants used to determine the complexity of grids affect the ranking of the grids. Future research is required to determine if SPSTs with more statistical knowledge will describe the complexity of the grids differently than the participants in this study.

In addition, there is a need to consider if $C_{2,3,4}$ is appropriate for other sized grids. Not all molecular systems can be represented by 10×10 grid. By studying larger grids, then patterns formed by the smaller subgrids may be studied in greater detail (see Attneave, 1959). For example, in a 40×40 grid (see Atkins, 1994) there are 1444 overlapping 3×3 subgrids and there are 512 different 3×3 subgrids that can be used. So, a second and third order measure (see Attneave, 1959) of complexity can then be computed for the grids and compared to SPSTs conceptions of the complexity. This is not possible for the smaller 10×10 grids. A future study involving larger grids may be able to describe possible effects any overall pattern of the 3×3 subgrids might have on the participants' conceptions of complexity. Also, Tree claimed that if the 5×5 pattern in Series 4 – Grid IV was repeated more than four times, then he would have noticed the repetition. A study involving larger grids may be able to determine a possible upper limit of subgrid size to recognizable repeating patterns. Furthermore, Maria noted several times that her background in art caused her to consider the black, white and gray cells differently depending on the background color of the paper. A future study may consider any possible effects of the background color on the perception of complexity of the grids.

Furthermore, a future study needs to determine the validity and reliability of the proposed complexity measure. More specifically, future research needs to identify any possible difference between the measures of the complexity of a two-dimensional grid as opposed to an individual's perception of the complexity of the grid. Falk and Konold (1994, 1998) have identified two values of their randomness measure, $P(A)$. According to Falk and Konold, a truly random two-dimensional grid had $P(A)$ near 0.5, whereas a commonly held misconception is that $P(A) > 0.5$, i.e., more alternations of the symbols, corresponds to a more random system.

Expand Study to Include a Variety of Participant Backgrounds

Further research is necessary to determine how experienced or expert individuals make the connections among probability, randomness, and entropy. Potential participants include physical science professors, post-secondary physical science majors, and experienced secondary science teachers. The data from a future research study that includes these suggested participants may provide insight into how to present the various concepts so that more post-secondary students to facilitate develop relational understanding. This type of study will further verify the viability of $C_{2,3,4}$ and will enable the construction of a theoretical framework that can specifically address individuals' understanding of entropy. In other words, research needs to determine how post-secondary science teachers develop the sufficient knowledge base of entropy and statistical mechanics.

Study to the Viability of the Complexity Framework to Understand Randomness

Lastly, a future study needs to clarify any possible relationship between randomness, complexity and complex arrangement. Falk and Konold (1997) suggested that randomness may be easier to understand through the complexity framework. However, the data from the current study was more ambiguous concerning the SPSTs' relational or conceptual understanding of a random event and their concept of complexity. A future study needs to determine the correlation, if any, between complexity and randomness for SPSTs.

Concluding Remarks

This research describes the conceptions of randomness, complexity, and entropy of eight pre-service secondary science teachers. The detailed explanations they provided demonstrated they brought their previous experiences to each of the questions and activities. There was evidence that the participants held differing views of complexity and randomness. Furthermore, the participants did not make the connections between probability, sample space and entropy. If we expect the pre-service teachers to competently help secondary students conceptually understand the concept of entropy, then science educators must better prepare them to make the necessary connections themselves and, more importantly, to prepare them to appropriately teach entropy to secondary students.

REFERENCES

Albert, J. (2006). Interpreting probabilities and teaching the subjective viewpoint. In G. F. Burrill, & P. C. Elliott (Eds.), *Thinking and Reasoning with Data and Chance* (pp. 417-433). Reston, VA: NCTM.

American Chemical Society. (2003, June). *Exploring the molecular vision* (Conference Report by Society Committee on Education). Washington, D.C.: Author.

Atkins, P. (1994). *The second law*. New York: W. H. Freeman and Company.

Atkins, P. (1998). *Physical chemistry* (Sixth edition). New York: W. H. Freeman and Company.

Attneave, F. (1959). *Applications of information theory to psychology: A summary of Basic Concepts, Methods, and Results*. New York: Henry Holt.

Baierlein, R. (1994). Entropy and the second law: A pedagogical alternative. *American Journal of Physics, 62*, 15-26.

Bailar, Jr., J. C., Moeller, T., Kleinberg, J., Guss, C. O., Castellion, M. E., & Metz, C. (1978). *Chemistry.* New York: Academic Press.

Batanero, C., Godino, J. D., & Navarro-Pelayo, V. (1997). Combinatorial reasoning and its assessment. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 239-252). Amsterdam: IOS Press.

Batanero, C., Green, D. R., & Serrano, L. R. (1998). Randomness, its meanings and educational implications. *International Journal of Mathematics Education, Science and Technology, 29*, 113-123.

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 15-37). New York: Springer.

Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32, 181-199.

Batanero, C., & Sanchez, E. (2005). What is the nature of high school students' conceptions and misconceptions about probability? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 241-266). New York: Springer.

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary school students. *Journal for Research in Mathematics Education, 30*, 558-567.

Bauman, R. P. (1995). Probability and disorder. *The Physics Teacher, 33*, 420.

Ben-Naim, A. (2007). *Entropy demystified: The second law reduced to plain common sense*. Hackensack, NJ: World Scientific.

Ben-Naim, A. (2008). *A farewell to entropy: Statistical thermodynamics based on information*. Hackensack, NJ: World Scientific.

Benson, D., Wittrock, M., & Baur, M. (1993). Students' preconceptions of the nature of gases. *Journal of Research in Science Teaching*, 30(6), 587-597.

Ben-Zvi, D., & Garfield, J. (2004a). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 397-409). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D., & Garfield, J. (2004b). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 3-15). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Bhushan, N., & Rosenfeld, S. (1995). Metaphorical models in chemistry. *Journal of Chemical Education, 72*, 578-582.

Brown, R. (1827). *A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies*. [unpublished manuscript] [Available: http://sciweb.nybg.org/science2/pdfs/dws/Brownian.pdf].

Bruning, R. H., Schraw, G. J., Norby, M. M., & Ronning, R. R. (2004). *Cognitive psychology and instruction (4th ed.)*. Upper Saddle River, NJ: Pearson Education Inc.

California State Board of Education. (1998a). *Mathematics Content Standards*. Retrieved September 28, 2006 from http://www.cde.ca.gov/be/st/ss/mthgrades8-12.asp.

California State Board of Education. (1998b). *Science Content Standards: Chemistry*. Retrieved September 28, 2006 from http://www.cde.ca.gov/be/st/ss/scchemistry.asp.

California State Board of Education. (1998c). *Science Content Standards: Physics*. Retrieved September 28, 2006 from http://www.cde.ca.gov/be/st/ss/scphysics.asp.

Canada, D. (2004). *Elementary preservice teachers' conceptions of variation*. Unpublished doctoral dissertation.

Canada, D. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal, 5(1)*, 36-64.

Chabay, R. W., & Sherwood, B. A. (1999). Bringing atoms into first-year physics. American *Journal of Physics*, 67, 1045-1049.

Christian von Baeyer, H. (1998). *Warmth disperses and time passes: The history of heat.* New York: Random House.

Clough, E. E., & Driver, R. (1985). Secondary students' conceptions of the conduction of heat: Bringing together scientific and personal views. *Physics Education, 20*, 176-182.

Cornely-Moss, K. (1995). Exam question exchange: Kinetic theory of gases. *Journal of Chemical Education, 72*, 715-716.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among the five traditions*. Thousand Oaks, CA: Sage Publications.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 79-95). Dordrecht, the Netherlands: Kluwer Academic Publishers.

delMas, R., Garfield, J., & Chance, B. (1999) A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*, http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm.

Dickerson, R. E. (1969). *Molecular thermodynamics*. New York: W. A. Benjamin, Inc.

Eigen, M., & Winkler, R. (1993). *Laws of the game: How the principles of nature govern chance*. Princeton, NJ: Princeton University Press.

English, L. (2005). Combinatorics and the development of children's combinatorial reasoning. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 121-141). New York: Springer.

Falk, R. (1975). *Perception of randomness*. Unpublished doctoral dissertation, Hebrew University, Jerusalem, Israel.

Falk, R., Falk, R., & Levin, I. (1980). A potential for learning probability in young children. *Educational Studies in Mathematics, 11*, 181-204.

Falk, R., & Konold, C. (1994). Random means hard to digest. *Focus on Learning Problems in Mathematics, 16*(1), 2-12.

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*, 301-318.

Falk, R., & Konold, C. (1998). Subjective randomness. In S. Kotz, C. B. Read, & D. L. Banks (Eds.), *Encyclopedia of statistical sciences* (pp. 653-659). New York: J. Wiley & Sons.

Feldman, J. (2004). How surprising is a simple pattern? Quantifying "eureka!" *Cognition, 93(3)*, 199-224.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal of Research in Mathematics Education, 28*, 96-105.

Fischbein, E., Sainati Nello, M., & Sciolis Marino, M. (1991). Factors affecting probabilistic judgments in children and adolescents. *Educational Studies in Mathematics, 22*, 523-549.

Flores, A. (2006). Using graphing calculators to redress beliefs in the "law of small numbers." In G. F. Burrill & P. C. Elliott (Eds.), *Thinking and Reasoning with Data and Chance* (pp. 291-304). Reston, VA: NCTM.

Fuchs, H. U. (1987). Entropy in the teaching of introductory thermodynamics. *American Journal of Physics, 55*, 215-219.

Gabel, D. L., Samuel, K. V., & Hunn, D. (1987). Understanding the particulate nature of matter. *Journal of Chemical Education, 64*, 695-697.

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 47-78). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 38-63). New York: Springer.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10* [Online]. Available: http://www.amstat.org/publications/jse/v10n3/garfield.html.

Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal* 2(1), 22-38.

Gell-Mann, M. (1994). *The quark and the jaguar*. New York: Henry Holt & Co.

Georgia Department of Education. (2005). *Quality Core Curriculum Standards and Resources*. Retrieved September 28, 2006 from http://www.glc.k12.ga.us/qcc/homepg.asp.

Gilbert, J. K., & Swift, D. J. (1985). Towards a Lakatosian analysis of the Piagetian and alternative conceptions research programs. *Science Education, 69*, 681-696.

Goldin, G. A. (1998). Observing mathematical problem solving through task-based interviews. In A. R. Teppo (Ed.), *Qualitative research methods in mathematics education* (pp. 40-62). Reston, VA: National Council of Teachers of Mathematics.

Green, D. R. (1979). The chance and probability concepts project. *Teaching Statistics, 1*, 66-71.

Green, D. R. (1986). Principles of Learning Probability and Statistics; Children's Understanding of Randomness: Report of a Survey of 1600 Children Aged 7-11 Years. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 766-783). Victoria, B.C.: University of Victoria.

Green, D. R. (1989). School pupils' understanding of randomness. In R. Morris (Ed.), *Studies in mathematics education: The teaching of statistics* (pp. 27-39). Paris: UNESCO.

Green, D. R. (1991). A longitudinal study of pupils' probability concepts. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics: School and general issues* (Vol. 1, pp. 320-328). Voorburg, The Netherlands: International Statistical Institute.

Greer, B., & Mukhopadhyay, S. (2005). Teaching and learning the mathematization of uncertainty: Historical, cultural and political contexts. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 297-324). New York: Springer.

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, 29, 611-628.

Harrison, A. G., & Treagust, D. F. (1996). Secondary students' mental models of atoms and molecules: Implications for teaching chemistry. *Science Education, 80*, 509-534.

Hawkins, A. S., & Kapadia, R. (1984). Children's conception of probability – A psychological and pedagogical review. *Educational Studies in Mathematics, 15*, 349-377.

Hill, J. W. (1992). *Chemistry for changing times (6th ed.).* New York: Macmillan.

Holstein, J. A., & Gubrium, J. F. (1995). *The active interview.* Thousand Oaks, CA: Sage Publications.

Holtzclaw, Jr., H. F., Robinson, W. R., & Odom, J. D. (1991). *General chemistry with qualitative analysis (9th ed.)*. Lexington, MA: D. C. Heath.

Horvath, J., & Lehrer, R. (1998). A model-based perspective on the development of children's understanding of chance and uncertainty. In S. P. LaJoie (Ed.), *Reflections on statistics: Agendas for learning, teaching, and assessment in K-12* (pp. 121-148). Mahwah, NJ: Lawrence Erlbaum.

Illinois State Board of Education. (1997a). *Illinois Learning Standards (ILS): Science*. Retrieved September 28, 2006 from http://www.isbe.state.il.us/ils/science/standards.htm.

Illinois State Board of Education. (1997b). *Illinois Learning Standards (ILS): Mathematics*. Retrieved September 28, 2006 from http://www.isbe.state.il.us/ils/ math/standards.htm.

Jacobson, M., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences, 15*, pp. 11-34.

Jolliffe, F. (2005). Assessing probabilistic thinking and reasoning. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 325-344). New York: Springer.

Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of development in statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 97-117). Dordrecht, the Netherlands: Kluwer Academic Publishers.

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

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*, 487-519.

Justi, R., & Gilbert, J. K. (1999). History and philosophy of science through models: The case of chemical kinetics. *Science and Education, 8*, 287-307.

Kahnemann, D., & Tversky, A. (1982). Belief in the law of small numbers. In D. Kahnemann, P. Slovic, & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 23-31). New York: Cambridge University Press.

Kaplan, M., & Kaplan, E. (2006). *Chance are …: Adventures in probability*. New York: Viking Penguin.

Kauzmann, W. (1966). *Kinetic theory of gases*. Reading MA: W. A. Benjamin, Inc.

Kauzmann, W. (1967). *Thermodynamics and statistics: With applications to gases*. New York: W. A. Benjamin, Inc.

Kelly, I. W., & Bany, B. (1984). Probability: Developing student intuition with 20 × 20 array. *School Science and Mathematics, 84*, 598-604.

Kendall, J. S., & Marzano, R. J. (2004). *Content knowledge: A compendium of standards and benchmarks for K-12 education*. Aurora, CO: Mid-continent Research for Education and Learning. Online database: http://www.mcrel.org/standards-benchmarks/.

Klinger, A., & Salingaros, N. A. (2000). A pattern measure. *Environment and Planning B: Planning and Design, 27*, 537-547.

Konold, C. (1991). Understanding students' beliefs about probability. In E. von Glasersfeld (Ed.), *Radical Constructivism in Mathematics Education*, (pp. 139-156). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Konold, C., & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 169-199). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students reasoning about probability. *Journal for Research in Mathematics Education, 24*, 392-414.

Kvatinsky, T., & Even, R. (2002). Framework for teacher knowledge and understanding of probability. *Proceedings of the Sixth International Conference on the Teaching of Statistics* [On-line]. Available: http://www.stat.auckland.ac.nz/~iase/publications/1/6a4_kvat.pdf.

Lambert, F. L. (2002). Disorder – A cracked crutch for supporting entropy discussions. *Journal of Chemical Education, 79*, 187-192.

Lambert, F. L. (2002). Entropy is simple qualitatively. *Journal of Chemical Education, 79*, 1241-1245.

Langrall, C. W., & Mooney, E. S. (2005). Characteristics of elementary school students' probabilistic thinking. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 95-119). New York: Springer.

Le Coutre, M., Rovira, K., Le Coutre, B., & Poitevineau, J. (2006). People's intuitions about randomness and probability: An empirical study. *Statistics Education Research Journal, 5*, 20-35.

Lee, K. (2001). How to teach statistical thermal physics in an introductory physics course. *American Journal of Physics, 69*, 68-75.

Lee, O., Eichinger, D. C., Anderson, C. W., Berkheimer, G. D., & Blakeslee, T. D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching, 30*, 249-270.

Lesk, A. M. (1974). Do the particles of an ideal gas collide? *Journal of Chemical Education, 51*, 141.

Lin, H., Cheng, H., & Lawrenz, F. (2000). The assessment of students and teachers' understanding of gas laws. *Journal of Chemical Education, 77*, 235-238.

Linn, M. C., & Songer, N. B. (1991). Teaching thermodynamics to middle school students: What are the appropriate cognitive demands? *Journal of Research in Science Teaching, 28*, 885-918.

McDowell, A. C. (1999). A simple derivation of the Maxwell-Boltzmann distribution. . *Journal of Chemical Education, 76*, 1393-1394.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Metz, K. E. (1997). Dimensions in the assessment of students' understanding and application of chance. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 223-238). Amsterdam: IOS Press.

Metz, K. E. (1998). Emergent understanding and attribution of randomness: Comparative analysis of the reasoning of primary grade children and undergraduates. *Cognition and Instruction, 16*, 285-365.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

Moore, D. S. (2001). *Statistics: Concepts and controversies*. New York: W. H. Freeman and Company.

Moore, D. S. (2004). *The basic practice of statistics (3rd Ed.)*. New York: W. H. Freeman and Company.

Moore, T. A., & Schroeder, D. V. (1997). A different approach to introducing statistical mechanics. *American Journal of Physics, 65*, 26-36.

Munisamy, S., & Doraisamy, L. (1998). Levels of understanding of probability concepts among secondary school pupils. *International Journal of Mathematical Education in Science and Technology, 29*, 39-45.

Nakhleh, M. B. (1993). Are or students conceptual thinkers or algorithmic problem solvers? *Journal of Chemical Education, 70*, 52-55.

Nakhleh, M. B., & Mitchell, R. C. (1993). Conceptual learning versus problem solving. *Journal of Chemical Education, 70*, 190-192.

Nakhleh, M. B., & Samarapungavan, A. (1999). Elementary school children's beliefs about matter. *Journal of Research in Science Teaching, 36*, 777-805.

National Association of Biology Teachers [NABT]. (2004). NABT position paper: Biology teaching preparation standards for middle and secondary teachers. Retrieved on 10/5/2009, from http://securewebcc.com/nabt/files/pdf/bio teaching prep standards.pdf.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council (NRC). (1996). *National Science Education Standards*. National Academy Press: Washington, DC.

National Science Teachers Association. (2003). *Standards for Science Teacher Preparation*. Retrieved September 28, 2006 from http://www.nsta.org/main/pdfs/NSTAstandards2003.pdf.

New Jersey Department of Education. (1996a). *Core Curriculum Content Standards for Science*. Retrieved September 28, 2006 from http://www.state.nj.us/njded/cccs/s5_science.htm.

New Jersey Department of Education. (1996b). *Mathematics Curriculum Framework*. Retrieved September 28, 2006 from http://www.state.nj.us/njded/cccs/s4_math.htm.

New York State Education Department. (2005). *Mathematics Core Curriculum*. Retrieved September 28, 2006 from http://www.emsc.nysed.gov/ciai/mst/mathstandards/revised3.htm.

New York State Education Department. (n.d.a). *Chemistry Core Curriculum*. Retrieved September 28, 2006 from http://www.emsc.nysed.gov/ciai/mst/pub/chemist.pdf.

Noh, T., & Scharmann, L. C. (1997), Instructional influence of a molecular-level pictorial presentation of matter on students' conceptions and problem-solving ability. *Journal of Research in Science Teaching, 34*, 199-217.

Novick, S., & Nussbaum, J. (1981). Pupil's understanding of the particulate nature of matter: A cross-age study. *Science Education, 65*, 187-196.

Nussbaum, J. (1978). Junior high school pupils' understanding of the particulate nature of matter: An interview study. *Science Education, 62*, 273-281.

Nussbaum, J. (1985). The particulate nature of matter in the gaseous phase. In R. Driver, E. Guesne, and A. Tiberghien (Eds.), *Children's Ideas in Science,* (pp. 124-144). Philadelphia: Open University Press.

Odgers, B. M. (2003). Teachers' beliefs about the nature of science and science education in relation to recently introduced constructivist syllabuses in secondary schools in Queensland,

Australia. *Paper presented at the Hawaii International Conference on Education*. Retrieved 10/5/2009, from http://www.hiceducation.org/edu_proceedings/Barbara M. Odgers.pdf.

Paparistodemou, E., Noss, R., & Pratt, D. (2002). Exploring in sample space:

Developing young children's knowledge of randomness. *Proceedings of the Sixth International Conference on the Teaching of Statistics* [On-line]. Available: http://icots6.haifa.ac.il/PAPERS/2A3_PAPA.PDF.

Pathria, R. K. (2001/1972). *Statistical mechanics (2nd Ed.)*. Boston, MA: Butterworth-Heinemann.

Pauli, W. (1973). *Statistical mechanics*. Mineola, NY: Dover.

Pauling, L. (1970). *General Chemistry*. New York: Dover Publications, Inc.

Paulos, J. A. (1995). *A mathematician reads the newspaper.* New York: Anchor Books.

Pennsylvania Department of Education. (2002a). *Academic Standards for Mathematics*. Retrieved September 28, 2006 from http://www.pde.state.pa.us/k12/lib/k12/MATH.pdf.

Pennsylvania Department of Education. (2002b). *Academic Standards for Science and Technology*. Retrieved September 28, 2006 from http://www.pde.state.pa.us/k12/lib/k12/scitech.pdf.

Pfannkuch, M. (2005). Probability and statistical inference: How can teachers enable learners to make the connection? In Jones, G. A. (Ed.), *Exploring Probability in School: Challenges for Teaching and Learning* (pp. 267-294). New York: Springer.

Pfannkuch, M., & Brown, C. M. (1996). Building on and challenging students' intuitions about probability: Can we improve undergraduate learning? *Journal of Statistics Education*, [On-line]. Available: http://www.amstat.org/publications/jse/v4n1/pfannkuch.html.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (pp. 17-46). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* London: Routledge & Kegan Paul.

Poincare, H. (2004/1908). *Science and method*. Mineola, NY: Dover Publications, Inc.

Polanyi, M. (1974). *Personal knowledge: Towards a post-critical philosophy*. Chicago: The University of Chicago Press.

Polkinghorne, J. C. (1984). *The Quantum World*. Princeton, NJ: Princeton University Press.

Pratt, D. (2005). How do teachers foster students' understanding of probability? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 171-189). New York: Springer.

Prentis, J. J., & Zainiev, T. (1999). The art of statistical mechanics: Looking at microscopic spectra and seeing macroscopic phenomena. *American Journal of Physics, 67*, 1123-1128.

Reif, F. (1999). Thermal physics in the introductory physics course: Why and how to teach it from a unified perspective. *American Journal of Physics, 67*, 1051-1062.

Resnick, M., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *Journal of the Learning Sciences 7*(2): 153–171.

Rhodes, G. (1992). Does a one-molecule gas obey Boyle's law? *Journal of Chemical Education, 69*, 16-18.

Riveros, H. G. (1995). Probability and disorder. *The Physics Teacher, 33*, 420.

Rubel, L. H. (2006). Students' probabilistc thinking revealed: The case of coin tosses. In G. F. Burrill & P. C. Elliott (Eds.), *Thinking and Reasoning with Data and Chance* (pp. 49-59). Reston, VA: NCTM.

Russell, D. K. (1996). The Boltzmann distribution. *Journal of Chemical Education, 73*, 299-300.

Schaeffer, R. L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004). *Activity-based statistics*. Emeryville, CA: Key Curriculum Press.

Schamp, H. W. (1990). Model misunderstandings: Teach the limitations. *Science Teacher, 57*(9), 16-19.

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

Skemp, R. R. (1977). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20–26.

Smith, N. O. (1982). *Elementary statistical thermodynamics: A problem approach*. New York: Plenum Publishing Corp.

Smith, M. A. (1993). The nature of distribution functions for colliding systems. *Journal of Chemical Education, 70*, 218-223.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stewart, I. (1989). *Does God play dice? The mathematics of chaos*. Cambridge, MA: Basil Blackwell.

Stieff, M., & Wilensky, U. (2003). Connected chemistry – incorporating interactive simulations into the chemistry curriculum. *Journal of Science Education and Technology, 12*, 285-302.

Stohl, H. (2005). Probability in teacher education and development. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 345-366). New York: Springer.

Styer, D. F. (2000). Insight into entropy. *American Journal of Physics, 68*, 1090-1096.

Taber, K. S. (1998). An alternative conceptual framework from chemistry education. *International Journal of Science Education, 20*, 597-608.

Taleb, N. N. (2005). *Fooled by randomness*. New York: Random House.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. New York: Random House.

Tarr, J. E., & Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? In Jones, G. A. (Ed.), *Exploring Probability in School: Challenges for Teaching and Learning* (pp. 215-238). New York: Springer.

Tarr, J. E., Stohl Lee, H., & Rider, R. L. (2006). When data and chance collide: Drawing inferences from empirical data. In G. F. Burrill & P. C. Elliott (Eds.), *Thinking and Reasoning with Data and Chance* (pp. 139-149). Reston, VA: NCTM.

Terry, C. (1995). Thermodynamics and entropy. *The Physics Teacher, 33*, 328.

Texas Education Agency. (1998). *Texas Essential Knowledge and Skills for Science*. Retrieved September 28, 2006 from http://www.tea.state.tx.us/rules/tac/chapter112/index.html.

Texas Education Agency. (2006). *Texas Essential Knowledge and Skills for Mathematics*. Retrieved September 28, 2006 from http://www.tea.state.tx.us/rules/tac/chapter111/index.html.

Thomsen, V. (1998). Thermodynamics concepts inventory? *The Physics Teacher, 36*, 454.

Tinnesand, M. J. (2003/1997). The unifying concepts. In ACS, *Chemistry in the national science education standards: A reader and resource manual for high school teachers*. Washington, D.C.: Author.

Tversky, A., & Kahnemann, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105-110.

Tversky, A., & Kahnemann, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Veal, W. R. (2004). Beliefs and knowledge in chemistry teacher development. *International Journal of Science Education, 26*, 329-351.

Venn, J. (2006/1866). *The logic of chance*. Mineola, NY: Dover Publications, Inc.

von Plato, J. (1994). *Creating modern probability: Its mathematics, physics and philosophy in historical perspective*. New York: Cambridge University Press.

Ware, S. A. (2003/1997). Chemistry in the national science education standards. In ACS, *Chemistry in the national science education standards: A reader and resource manual for high school teachers*. Washington, D.C.: Author.

Watkins, A. E., Schaeffer, R. L., & Cobb, G. W. (2004). *Statistics in action: Understanding a world of data*. Emeryville, CA: Key Curriculum Press.

Watson, J. (2005). The probabilistic reasoning of middle school students. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 145-169). New York: Springer.

Watson, J. M. (2006). Assessing the development of important concepts in statistics and probability. In G. F. Burrill & P. C. Elliott (Eds.), *Thinking and Reasoning with Data and Chance* (pp. 61-75). Reston, VA: NCTM.

Whitten, K. W., Davis, R. E., Peck, M. L., & Stanley, G. G. (1998). *General Chemistry (6th ed.)*. Belmont, CA: Brooks/Cole.

Whitten, K. W., Davis, R. E., Peck, M. L., & Stanley, G. G. (2004). *General Chemistry (7th ed.)*. Belmont, CA: Brooks/Cole.

Widom, B. (2002). *Statistical mechanics: A concise introduction for chemists*. New York: Cambridge University Press.

Wilbraham, A. C., Staley, D. D., Matta, M. S., & Waterman, D. L. (2004). *Chemistry*. Upper Saddle River, NJ: Pearson Prentice Hall.

Wilensky, U. (1993). *Connected mathematics: Building concrete relationships with mathematical knowledge*. Unpublished doctoral dissertation, Massachusetts Institute of Technology. [On-line]. Available: http://ccl.northwestern.edu/papers/download/Wilensky-thesis.pdf.

Wilensky, U. (1995). *Learning probability through building computational models*. Paper presented at the Nineteenth International Conference on the Psychology of Mathematics Education. Recife, Brazil. [On-line]. Available: http://ccl.northwestern.edu/papers/pme19/.

Wilensky, U. (1999a). GasLab – an Extensible Modeling Toolkit for Exploring Micro- and Macro-Views of Gases. In N. Roberts, W. Feurzeig and B. Hunter (Eds.), *Computer Modeling and Simulation in Science Education*. Berlin: Springer Verlag.

Wilensky, U. (1999b). *NetLogo*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. [Available: http://ccl.northwestern.edu/netlogo].

Wilensky, U. (2001). *Modeling nature's emergent patterns with multi-agent languages*. Linz, Austria: Eurologo 2001. [Available: http://ccl.northwestern.edu/papers/MEE/].

Wilensky, U., Hazzard, E., & Froemke, R. (1999). *GasLab - an extensible modeling toolkit for exploring statistical mechanics*. Sofia, Bulgaria: Proceedings of the Seventh Annual European Logo Conference EUROLOGO 99 [Online]. Available: http://www.ccl.sesp.northwestern.edu/cm/papers/eurologo/.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education, 8*, 3-19.

Wilson, D. B. (1981). Kinetic atom. *American Journal of Physics, 49*, 217-222.

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media , Inc.

Zumdahl, S. S. (1989). *Chemistry (2nd ed.).* Lexington, MA: D. C. Heath.

APPENDIX A

INTERVIEW PROTOCOL

Interview # 1: SPSTs' Conception of Probability

1a. A fair coin is tossed 10 times. Which sequence of coin tosses which is more likely? Explain.

i. a sequence with 2 T   ii. a sequence with 8 T

1b. A fair coin is tossed 10 times. Which sequence of coin tosses which is more likely? Explain.

i. a sequence with 3 H   ii. a sequence with 4 H

2a. Suppose two fair dice are simultaneously cast. Which outcome is more likely to occur?

i. 5, 6   ii. 5, 5

2b. Suppose three fair dice are simultaneously cast. Which outcome is more likely to occur?

i. 5, 3, 6   ii. 5, 5, 5   iii. 5, 5, 3

2c. Suppose one fair die is cast three times. Which outcome is more likely to occur?

i. 5, 3, 6   ii. 5, 5, 5   iii. 5, 5, 3

3. Given the letters *a*, *b*, *c*, *d* and *e*, If the letters may be used more than once

i. How many different strings of three letters can you form if the order of the letters is considered?

ii. How many different strings of three letters can you form if the order of the letters is *not* considered?

4.

i. What does the word "random" mean to you?

ii. Give an example of something that happens in "a "random" way (Canada, 2004; Watson, 2006).

5. What is "entropy?" Provide an example to describe entropy to secondary students.

6. Rank the following sequences from the least complex arrangement to the most complex arrangement:

I.
H T T T H T T H H H H H H H H H T T T T T

II.
T T H H T T T H H T H T T H H H H H T H T

III.
H T T H H H T H H T T H T H T H T H T H T

IV.
H T H T T T H H T T H H H T T T T H H H H

V.
H T T T H H H T H H T T T H T H T H H T H

<u>Some Follow-up Questions for Interview # 1</u>

What criteria did you use to rank the sequences? I asked follow-up questions to further probe into their reasoning. For example:

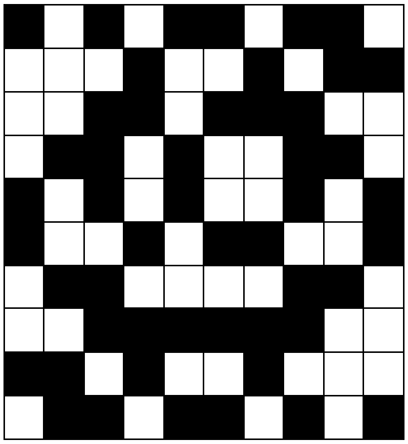
- Why did you place sequence ____ before/after sequence ____?
- What features of the sequences did you use for the exercise?
- Which features were the most important when looking at the complex arrangement?
- Did you consider [insert feature here]? Why or why not?

Interview # 2: SPSTs' Conceptions of Complexity

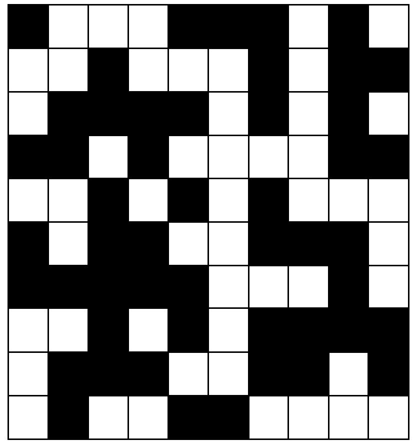
Series 1

Rank the following grids from the least complex arrangement to the most complex arrangement:

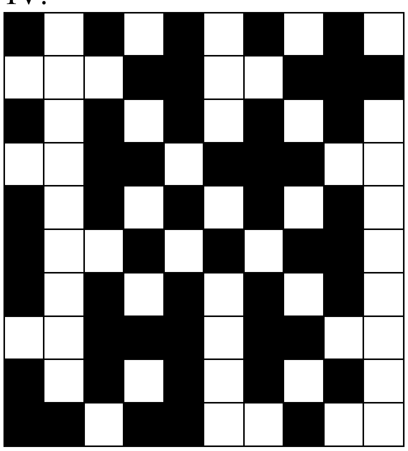
I.



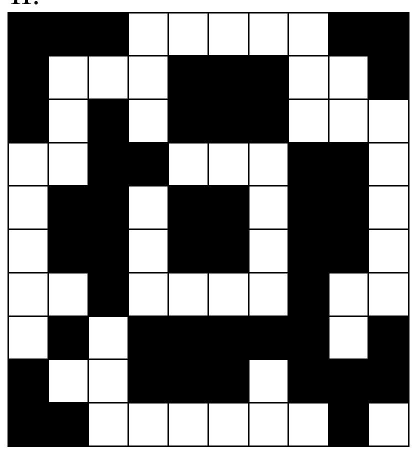
III.



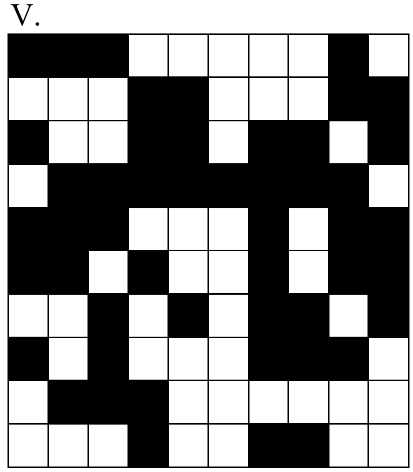
IV.



II.



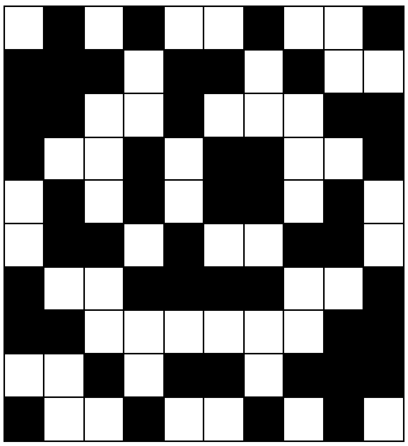
V.



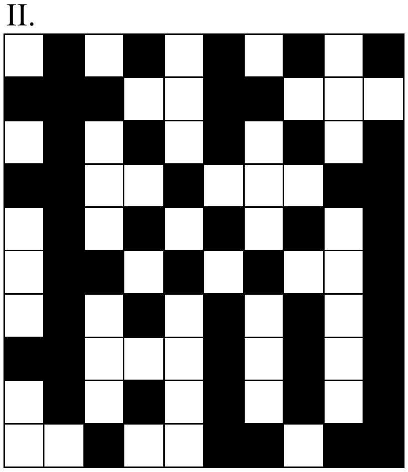
Series 2

Rank the following grids from the least complex arrangement to the most complex arrangement:

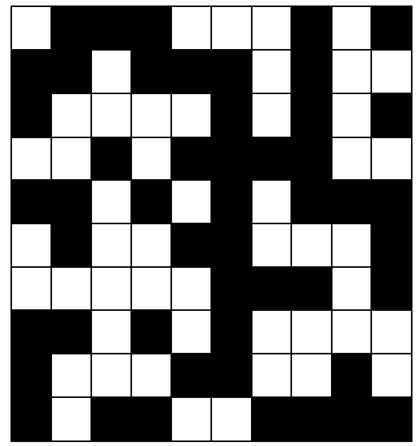
I.



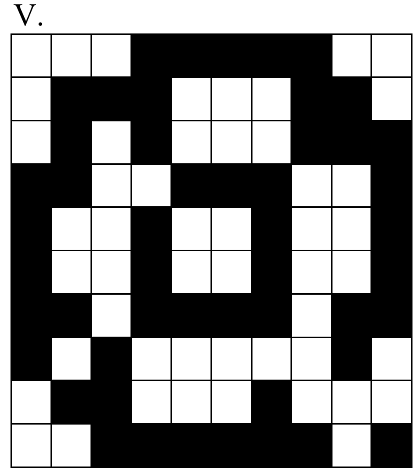
II.



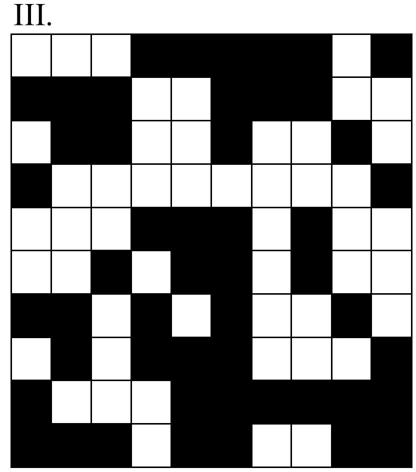
IV.



V.



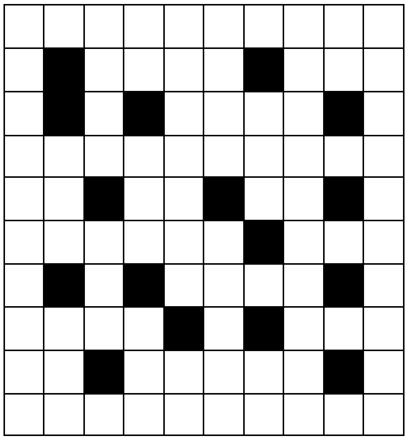
III.



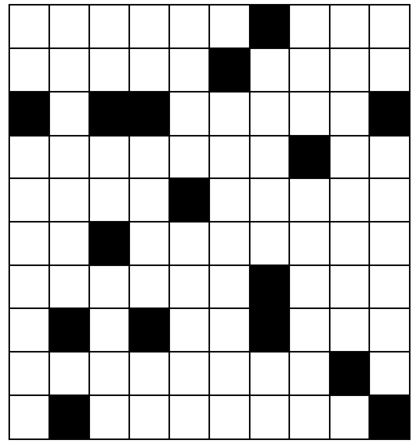
Series 3

Rank the following grids from the least complex arrangement to the most complex arrangement:

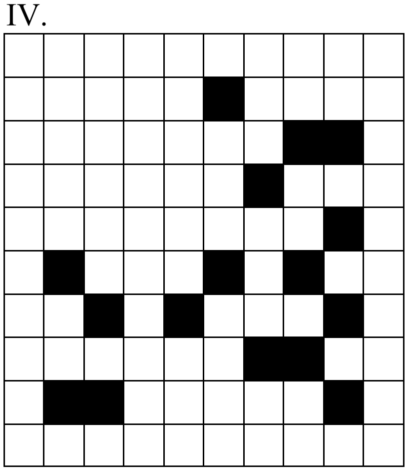
III.



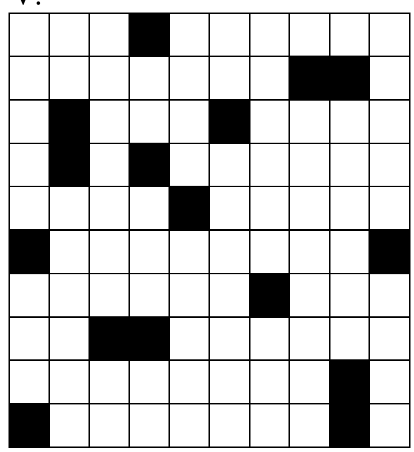
I.



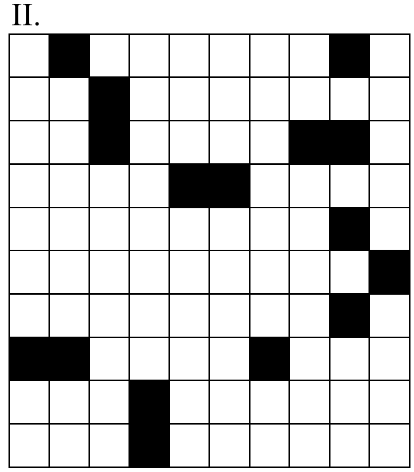
IV.



V.



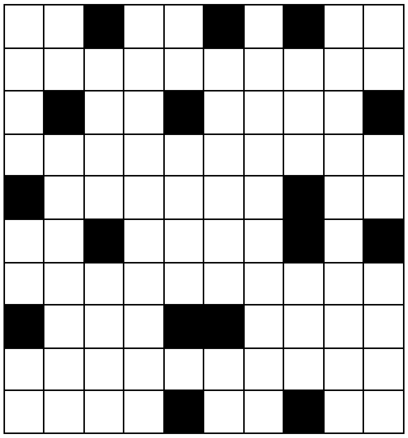
II.



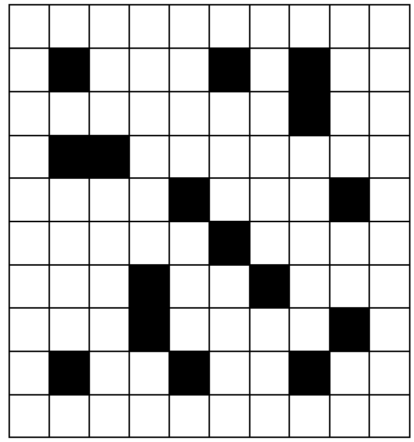
Series 4

Rank the following grids from the least complex arrangement to the most complex arrangement:

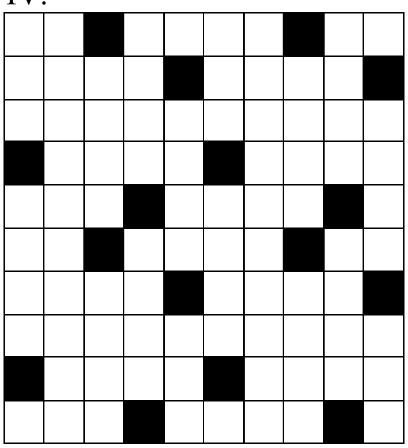
I.



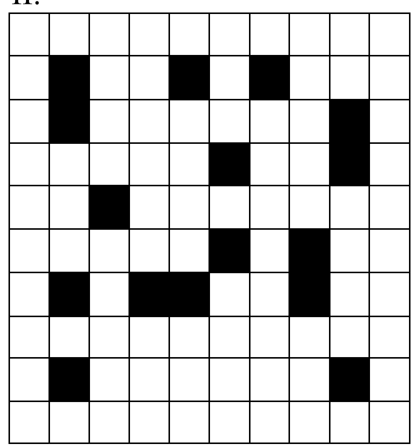
III.



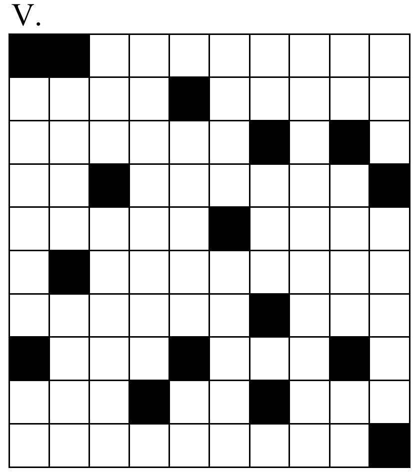
IV.



II.



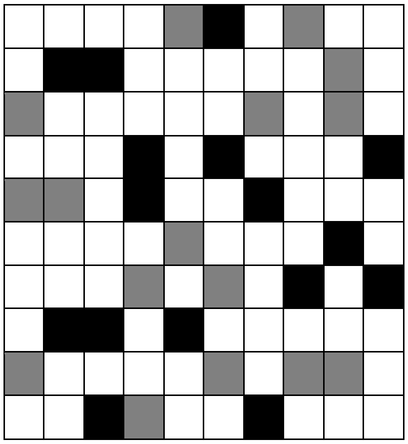
V.



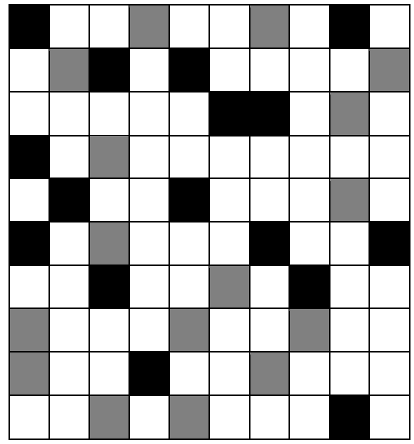
<u>Series 5</u>

Rank the following grids from the least complex arrangement to the most complex arrangement:

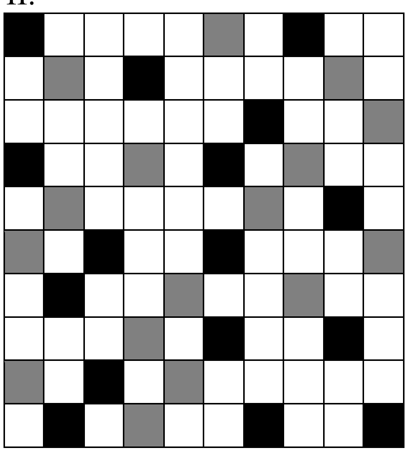
V.



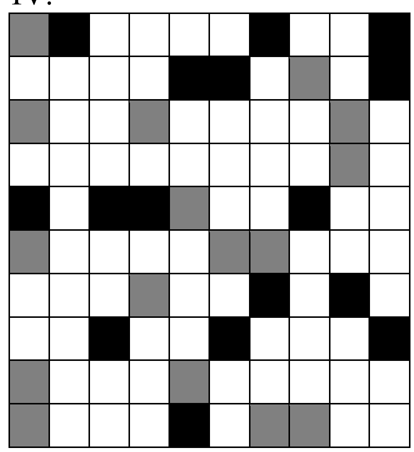
III.



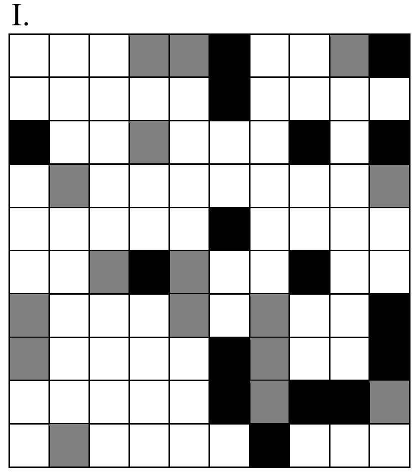
II.



IV.



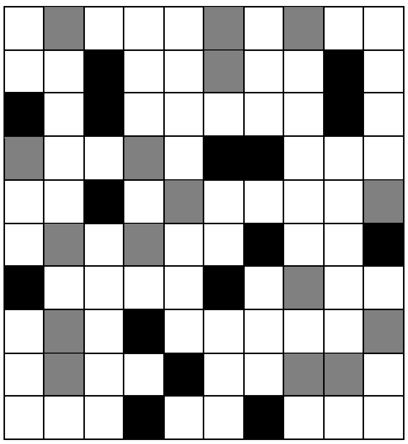
I.



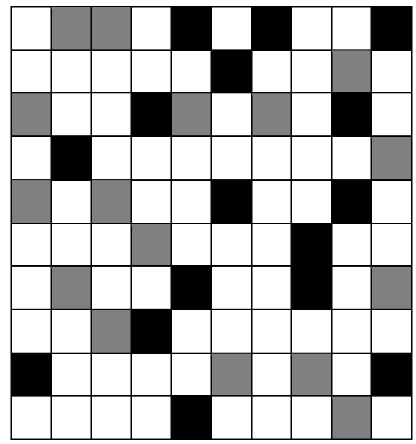
Series 6

Rank the following grids from the least complex arrangement to the most complex arrangement:

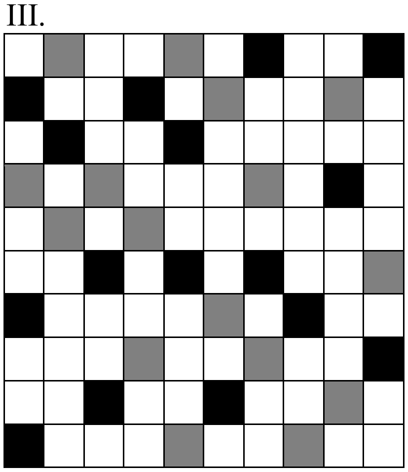
I.



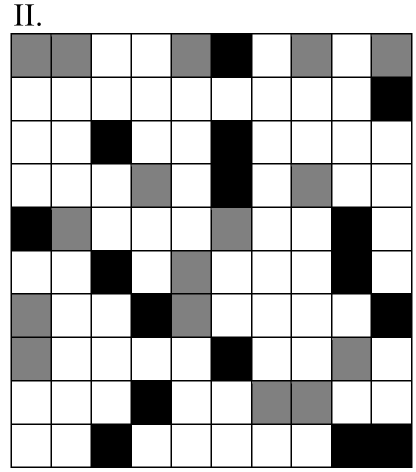
IV.



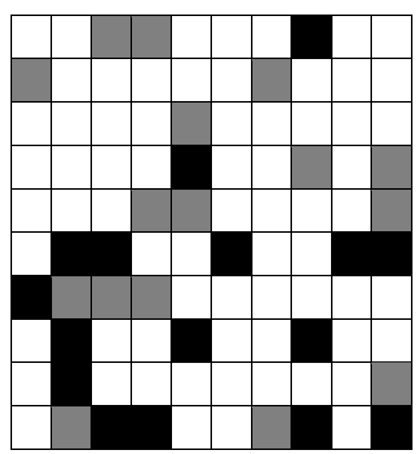
III.



II.



V.



Some Follow-up Questions for Interview # 2

What criteria did you use to rank the grids? I will ask follow-up questions to further probe into their reasoning. For example:

- Why did you place grid ____ before/after grid ____?
- What features of the grid did you use for the exercise?
- Which features were the most important when looking at the complex arrangement?
- Did you consider [insert feature here]? Why or why not?

Interview # 3

Students' Combinatorial Reasoning, Sample Space and the MB-Distribution

Students were asked to determine the number of possible discrete distributions (i.e., combinations) given a number of particles (N), a set of integral energy levels and the total energy of the system (E). Students were provided a worksheet (see Worksheet # 1) and tokens to assist them.

Some examples of systems:

$N = 2$	$N = 3$	$N = 3$	$N = 5$	$N = 4$
$E = 2$	$E = 3$	$E = 5$	$E = 3$	$E = 3$

After the determination of the number of distributions of the first system, I asked each participant the following questions:

- For each system, what was the most probable distribution? In other words, what microstate was most likely? Explain.
- What is the meaning of the most probable distribution? Why is the most probable distribution important?

After all the initial calculations (i.e., the systems with small numbers of particles), I will ask students the following questions:

- Generalize your results, in terms of
    - most probable distribution
    - descriptions of the various distributions
- How would you determine the number of distributions for a larger system?
- How does this relate to randomness?
- Predict the most probable distribution. Explain.

Worksheet # 1

Number of particles: _______________

Total Energy of system: ________________

Energy Level	Energy Units	molecules/atoms
E_5	5	______________________________________
E_4	4	______________________________________
E_3	3	______________________________________
E_2	2	______________________________________
E_1	1	______________________________________
E_0	0	______________________________________

Interview # 4

Entropy as a Process Activity

Prior to starting the game, I asked them the following question: What do you expect will happen after 1 [or 2, 3, 5, 20, 50, 100] steps in the game? Explain. The participants were asked why they chose that particular outcome over the other possible outcomes asking them to support their claim with the concepts of probability. If the participant uses the concept of combinations, then they will be encouraged to calculate the number of combinations of the board (a TI calculator was available to them[26]).

Other questions that I posed included (depending on their responses): Why did the board tend to contain more red tokens after the first few steps in the game? How many white tokens would you expect? What was the probability of any specific configuration? What was the probability that the board has 1, 2, 5, 10, 30 white tokens?

Questions that extended the game model included: Suppose the board started with the left side filled with red tokens and the right side filled with white tokens. What would happen after 1, 2, 3, 5, 20, 50, 100 steps? What would happen if I went sequentially along the cells and only used the six-sided die to choose the color? Would you expect similar or different results? Why?

How does the game relate to your conception of entropy?

What would happen if the six-sided die had four sides labeled B and two sides labeled G?

[26] I will assist the participant with the location of specific functions on the calculator only after the participant identifies the function.

Board Game for Entropy as a Process Activity

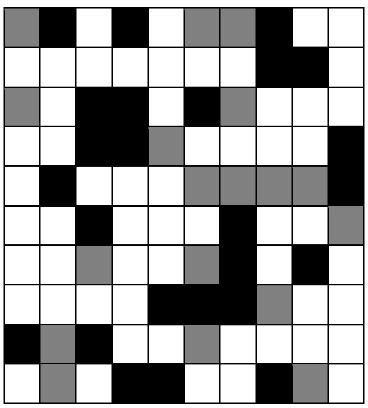
	1	2	3	4
1				
2				
3				
4				

Randomness as a Pattern Activity

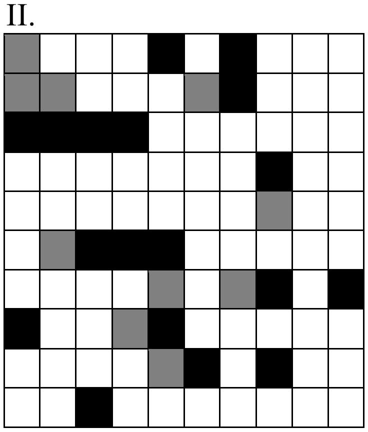
Series I

The following grids are 'snapshots' of *different* molecular systems (the black and gray cells represent different atoms/molecules). Which system(s) appears to have the most entropy? Which system(s) appears to have the least entropy? Explain.

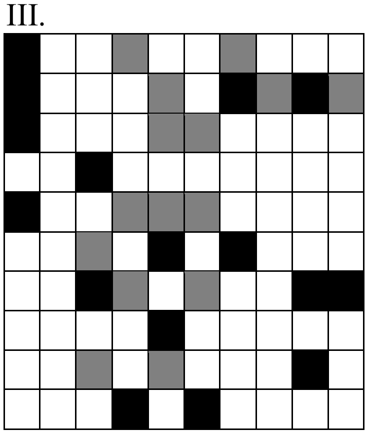
I.



II.



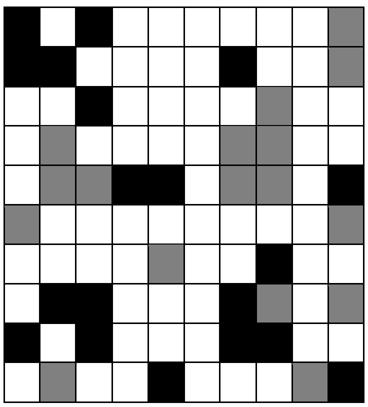
III.



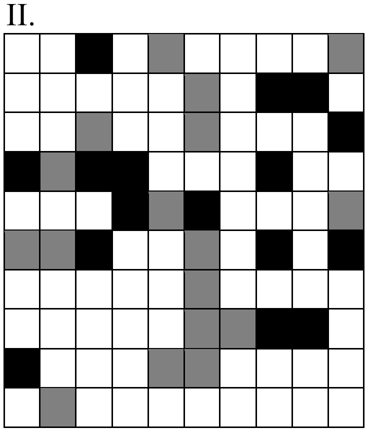
Series II

The following grids are 'snapshots' of *different* molecular systems (the black and gray cells represent different atoms/molecules). Which system(s) appears to have the most entropy? Which system(s) appears to have the least entropy? Explain.

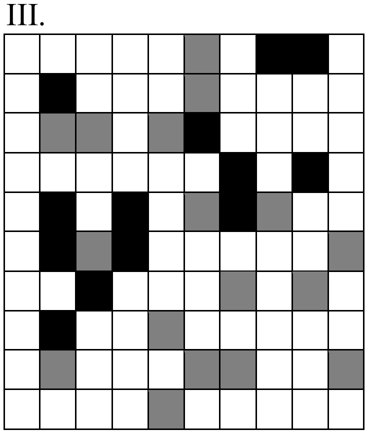
I.



II.



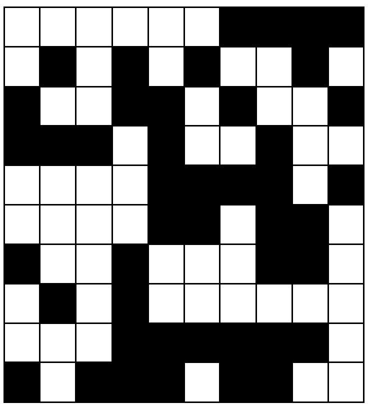
III.



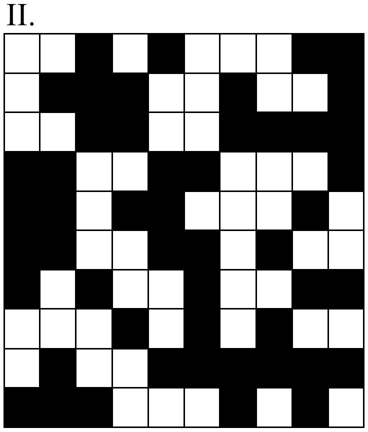
Series III

The following grids are ‘snapshots’ of *different* molecular systems (the black cells represent different atoms/molecules). Which system(s) appears to have the most entropy? Which system(s) appears to have the least entropy? Explain.

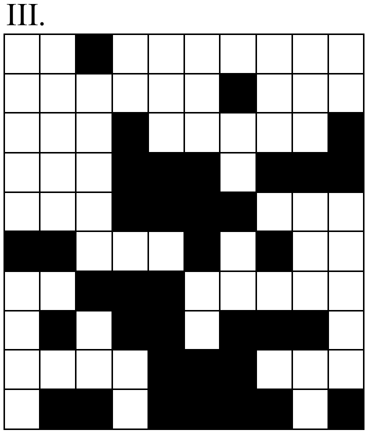
I.



II.



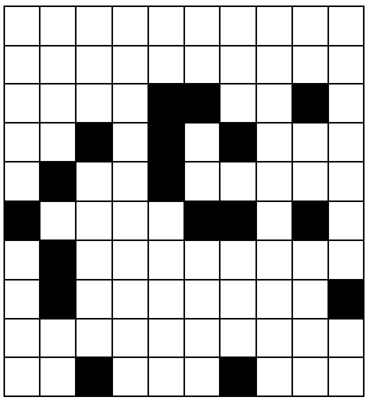
III.



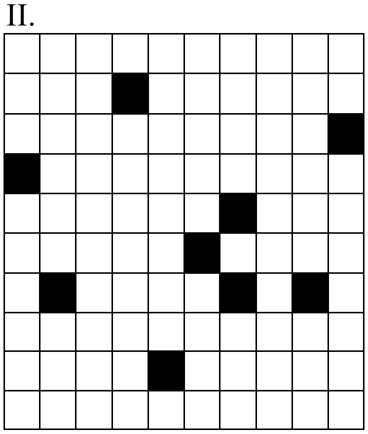
Series IV

The following grids are ‘snapshots’ of *different* molecular systems (the black cells represent different atoms/molecules). Which system(s) appears to have the most entropy (see Styer, 2000)? Which system(s) appears to have the least entropy? Explain.

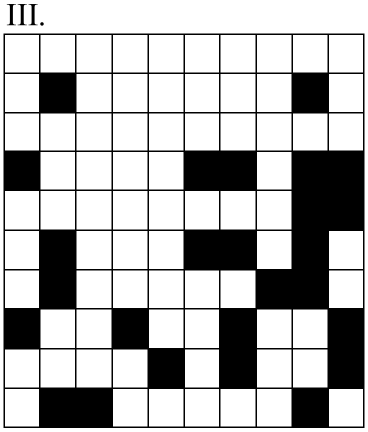
I.



II.



III.



APPENDIX B

QUESTIONNAIRE

1. What college-level mathematics courses that you have taken included statistical concepts?

2. What college-level mathematics education courses that you have taken included statistical concepts?

3. What college-level science courses that you have taken included statistical concepts?

4. What college-level science education courses that you have taken included statistical concepts?

5. Provide some detail about what was discussed and how the concepts were presented to the class.

a. What college-level science courses that you have taken included kinetic molecular theory?

b. What college-level science courses that you have taken included statistical mechanics?

c. What college-level science courses that you have taken included gas concepts?

6. Provide some detail about what was discussed and how the concepts were presented to the class.

a. What college-level science education courses that you have taken included kinetic molecular theory?

b. What college-level science education courses that you have taken included statistical mechanics?

c. What college-level science education courses that you have taken included gas concepts?

7. Explain why you can smell food cooking from another room.

8. Why do you want to be a secondary mathematics or physics teacher?

9. Why is mathematics important in science? Why is mathematics important in science education?