THE ACTION, PROCESS, OBJECT, AND SCHEMA THEORY FOR SAMPLING

Neil J. Hatfield
Arizona State University

This paper puts forth a new theoretical perspective for students' understanding of sampling. The Action, Process, Object, and Schema Theory for Sampling serves as a potential bridge between Saldanha's and Thompson's Multiplicative Conception of Sampling and APOS Theory. This theoretical perspective provides one potential way to describe the development of a student's conception of sampling. Additionally this perspective differs from most other perspectives in that it does not focus on the sample size the student uses or the sampling method, but rather how the student understands sampling in terms of a sampling distribution.

Key words: Sampling, APOS Theory, theoretical perspective, statistics, multiplicative conception of sampling

Introduction

Statistics is one of the most ubiquitous branches of mathematics in everyday life and it is also arguably one of the least understood areas of mathematics. One important aspect of Statistics is that of statistical inference. While students are taught how to calculate a statistic and then perform an appropriate test on that statistic, often the meanings that these students have is a procedural one (Lipson, 2003). Students rarely understand that when making inferences from a statistic, they are making judgments from a sample distribution (theoretical or experimental). Sampling and sampling distributions are thus an integral part of statistical inference. The National Council of Teachers of Mathematics (NCTM, 2000) noted that students should be able to construct and use sampling distributions in regards to statistical inference. Thus, it is important to examine how a student might understand sampling.

While there are several articles on students' understandings of sampling, the work of Saldanha and Thompson (2002) was highly influential in the development of this theoretical perspective of a student's understanding of sampling. Saldanha and Thompson's multiplicative conception of sampling (MCS) serves as a basis for building a more nuanced framework for student thinking about sampling. APOS theory's applicability to a wide range of mathematical concepts including, but not limited to functions, infinity, limits, mean, standard deviation, and the Central Limit Theorem, suggests that it could provide a beneficial framework to refine MCS. In essence, this proposed theoretical perspective serves as a link between MCS and APOS and provides a useful way to characterize a student's understanding of sample with an eye towards the theoretical development of that understanding.

To layout the APOS Theory for Sampling (APOS-Sampling), a brief description of the constructs involved in APOS-Sampling, APOS theory in general, and the multiplicative conception of sampling will be presented. Following this will be a presentation of the APOS-Sampling theoretical perspective with examples. Finally, a comparison to other theoretical perspectives will be made.

Constructs and Background Theories

Given that a construct is an idea existing in the mind of an observer about how an individual understands a concept, a catalytic construct serves as an observer's model of how an individual's understanding might advance within a given framework. Within the APOS framework, there are three catalytic constructs. As described by Sfard (1992), these catalytic constructs are interiorization, condensation, and reification[1]. Interiorization may be thought of as an individual's mental "process performed on already familiar objects" (Sfard, 1992, p. 64). In addition, Sfard (1992) defines condensation as taking the mental process used in interiorization and creating "a more compact, self-contained whole" (p. 64). The final catalytic construct used here is reification; a qualitative leap in how an individual thinks. To help explain these catalytic constructs, a brief example will be given.

When APOS theory is applied to a variety of mathematical contexts, an action is thought of as a transformation on familiar objects to obtain new objects (Breidenbach, Dubinsky, Hawks, & Nichols, 1992; Clark, Kraut, Mathews, & Wimbish, 2003; Dubinsky & McDonald, 2002; Dubinsky, Weller, McDonald, & Brown, 2005a, 2005b; Mathews & Clark, 2007; Sfard, 1992). Using APOS-Function[2] as a backdrop, consider the already familiar objects of numbers to a student. As the student repeats carrying out manipulations on numbers, effectively interiorizing mental processes, he/she develops an Action conception of function. Essentially, Sfard's interiorization is the catalyst that allows for a student to move from an understanding of function where "function" is meaningless to the Action conception of function. Similarly, as the student continues to work with functions more and more, he/she begins to condense the processes involved in functions. These processes settle into what Thompson (1994) describes as "self-evaluating expressions" (p. 6). When this happens we say that the student has the Process conception of function. Finally, when the student reifies (or engages in reflective abstraction) function from a set of processes into an object akin to a number, the student has made qualitative leap in his/her understanding, reaching the Object conception of function.

A brief note should be made about the use of the word "process". "Process" has been used multiple times with multiple meanings. In particular, "process" is used within the description of interiorization as well as within the Process conception. While the meanings here are complimentary, one should not view the meanings as interchangeable. There is a subtle distinction between the two meanings that centers on the vantage point of the individual. The meaning of "process" for interiorization places the student *inside* the process, while the meaning of "process" within the Process conception places the student *outside* of the process. For additional clarity, consider the following function, $f:\mathbb{R}\to\mathbb{R}$, $f(x)=2x^2-4$. When a student is *inside* the process, he/she focuses on the individual components of the process even if he/she sees the whole process. Thus, this individual would see $f(2)$ as the following sequence: square 2, get 4; multiply 4 by 2, get 8; subtract four from 8, get 4; the result is 4. When an individual has a vantage point of *outside* the process, the individual sees the process as one thing, i.e., $f(2)$ is what you get out of f when you put 2 into f.

Within the APOS literature, schema is most often expressed as a student's "collection of actions, processes, objects, and other schemas which are linked" (Dubinsky & McDonald, 2002, p. 3). Dubinsky and McDonald (2002) go on to note that the usage of the term "schema" within APOS is consistent with Tall and Vinner's (1981) term "concept image" ("the total cognitive structure associated with a concept" (p. 152)). For the APOS-Sampling perspective, schema will have the same meaning as expressed here.

---

[1] These catalytic constructs are akin to Piaget's constructs of internalization, interiorization, and reflective abstraction, each of which can be considered a catalytic construct.

[2] The author uses APOS-(Concept) as a shorthand reference to the body of work where APOS theory is applied to a specific concept.

Saldanha and Thompson (2002) characterize a student having a MCS as the student "having conceived a sample as a quasi-proportional mini version of the sampled population, where the 'quasi-proportionality' image emerges in anticipating a bounded variety of outcomes, were one to repeat the sampling process" (p. 266). They propose the following three levels to MCS. The first level requires that an individual visualize a proportional[3] relationship between the sample and the population. The second level occurs when the individual begins to see a sampling distribution take shape as he/she looks at the same statistic across multiple samples. Finally, the third level of understanding supports the individual making inference statements about the expectation of a particular sampling outcome and the outcome's representativeness. It should be noted that the use of the word 'multiplicative' does not entail multiplication, but rather an individual keeping in mind multiple attributes of a quantity (or quantities) at the same time.

Action, Process, Object Conceptions of Sampling

Using APOS Theory and the above catalytic constructs, I will describe the generation of APOS-Sampling. As with APOS-Function (see Breidenbach et al., 1992; Sfard, 1992; Thompson, 1994), APOS-Mean (see Clark et al., 2003; Mathews & Clark, 2007), etc., we must first start with an object familiar to the student. However, unlike the Action conception of function where a student needs only numbers and operations, sampling requires the student to have in mind more than numbers and operations. In the case of sampling, the familiar object starts with the student's conception of the population but also requires the student to have some conception of a stochastic process, as well as a conception of a statistic of interest. Additionally, the student must also have at least level one of the MCS (a proportional relationship between the sample and the population). Consider the following excerpt from Saldanha and Thompson (2002):

> D: Ok. It's asking…the question is…like "do you like Garth Brooks?". You're gonna go out and ask 30 people, it's[4] gonna ask 30 people 4500 times if they like Garth Brooks. The uh…(talks to himself) what's this? Let's see…the actual…like the amount of people who actually like Garth Brooks are…or 3 out of 10 people actually prefer like Garth Brook's music. (p. 262)

This excerpt demonstrates the Action conception of sampling. The population of familiar objects is people (who like or don't like Garth Brook's music). D wants a count of people who like Brook's music, thus the mental process he (or rather he has the computer perform) is a count of a number of people who like Brook's music out of a sample of 30 people. The stochastic process that D has in mind here references what Saldanha and Thompson (2002) refer to as the first two phases of the sampling process. D takes a sample of thirty people from the population, counts those who like Garth Brooks (phase one) and then continues this process 4499 more times (phase two). Thus the Action conception of sampling involves interiorizing (in Sfard's usage) phase one and two of the sampling process in relation to a collection of familiar objects and getting a collection of statistics[5].

[3] "Proportional" here should be thought of as 'quasi-proportional' which stems from envisioning a bounded variation of outcomes if the sampling process is repeated.
[4] The "it" that D is referring to is a computer program that will draw multiple samples where the population is set to have 30% of people prefer Garth Brooks.
[5] It should be noted that the generation of a statistic of interest pre-supposes that the statistic is an embedded object of the selected sample as well as another set of mental processes.

Just as with APOS-Function, the distinction between the Action conception of Sampling and the Process conception of Sampling involves a difference of vantage points. For Sampling, when a student is inside the process, the student appears to focus on the process of generating the multiple sample statistics. These sample statistics are simply results of a process. However, when the student's vantage point is outside of the process, his/her attention moves away from the process that yields the sample statistics to the statistics. Recall that condensation refers to turning a mental process into a self-contained whole. As the student continues to interiorize phases one and two of the sampling process, the phases, the statistic, and population begin to condense into a sophisticated meaning for the sample statistics. It is at this point that the student has moved to the Process conception of Sampling. The condensation shifts the student's focus from inside to outside of the process and allows for the student to engage in the third phase of sampling process as described by Saldanha and Thompson (2002). The third phase is a rather sophisticated scheme of ideas. Here the student creates categories that each sample statistic can fit into. This shift of focus to the sample statistic is consistent with the student progressing to the second level of MCS. The creation of categories is driven by the student's conception of a desired judgment he/she wants to make about a statistic. An example would beneficial here.

> D: When you go out and take one sample of 30 people, the cut off fraction means that if you're gonna count, you're gonna count that sample, if like 37% of the 30 people preferred Garth Brooks. And then it's going to tally up how many of the samples had 37% people that preferred Garth Brooks. (Saldanha & Thompson, 2002, p. 262)

Here D is exemplifying the process conception of sampling. In particular, D is focused not on the process that generates the sample statistic, but rather manipulating those statistics by comparing each proportion to a cut-score (the 37%) through the computer. Thus, D has created two categories, those sample statistics above or below the cut-score. Just as with other instances of APOS Theory, just because a student is able to speak/work in a way that is consistent with the Action conception to an observer, does not mean that the student is locked at that level. D is clearly able to talk about sampling at the Action conception, however, he is also able to move beyond this level of understanding and speak at the Process conception level of Sampling.

While there is some sense of variation within the Action conception of Sampling, a student at the Process conception must realize that variation of the sample statistic is not just possible but inevitable. At the Action level, a student is much more focused on the process that yields the sample statistic and is thus not likely to focus on differences in the values of the statistic. If he/she does, he/she may attribute the variation to causal features in sense of Konold's (1989) outcome approach. However, the variation of the sample statistic is an integral part of the Process conception of Sampling. Without the expectation of variation amongst the sample statistics, the student cannot create categories of the statistics.

As a student continues to partition the collection of sample statistics, the student begins to develop a distribution of the sample statistic. The student can start seeing the sampling distribution as an object in its own right. When the sampling distribution is reified by the student, he/she could now be said to be at the Object conception of Sampling. In APOS-Function, for example, when a student is at the Object conception, he/she is able to take a function (and that student's conception of the encapsulated process) and manipulate that function as if it were a familiar object, like a number. The student can add two functions, multiply, or compose two functions together. The student gets another function out of this manipulation.

However, in APOS-Sampling, the manipulations take a different form. Instead of having a pair of similar objects (two functions), the student has a sampling distribution (the reification of multiple sampling processes) and a sample statistic. The manipulation that the student now uses is a comparison between the sampling distribution and the sample statistic. This manipulation of the sampling distribution allows for the student to get a sense of what to expect when taking a sample of size *n* from the population prior to sampling. In addition, the comparison allows the student to decide how representative a sample of size *n* is of the population after sampling. The result of this manipulation/comparison is not necessarily another object but rather a characterization/judgment of samples and the population. The Object conception of Sampling is consistent with the third level of MCS. Consider the following excerpt:

> D: If like…if you represent—if you give it like the split of the population and then you run it through the how—number of samples or whatever it'll give you the same results as if—because in real life the population like of America actually has a split on whatever, on Pepsi, so it'll give you the same results as if you actually went out, did a survey with people of that split.
> I: Ok, now. What do you mean by "same results"? On an particular survey at all—you'll get exactly what it—?
> D: No, no. Each sample won't be the same but it's a…it'd be…could be close, closer…
> I: What's the "it" that would be close?
> D: If you get…if you take a sample…then the uh…the number of the like whatever, the number of "yes's" would be close to the actual population split of what it should be.
> I: Are you guaranteed that?
> D: You're not guaranteed, but if you do it enough times you can say it's with like 1 or 2% error depending upon uh how many times—I think—how many times you did it. (Saldanha & Thompson, 2002, pp. 265–266)

The above excerpt shows that D is progressing into the Object conception of Sampling. At times it is a little murky about what exactly D refers to when he says "it", however it is clear that D is talking about the purpose of resampling. While D knew the actual population proportion, there is evidence that D thinks about the sampling distribution ("Each sample won't be the same."). In the last line, D talks about the representativeness of a sample based off of the sampling distribution he is imagining.

Comparisons to Other Perspectives of Sampling

There has been some work done focusing on the (mis-) conceptions that adults have on sampling and other statistical topics (e.g., Kahneman's, Solvic's, and Tversky's 1982 book Judgment Under Uncertainty: Heuristics and Biases) as well as some work on children's conceptions (Lajoie, Jacobs, & Lavigne, 1995) of statistical topics. With reference to children's understanding of sampling, there have been several articles that focus on different aspects of sampling than this proposed perspective. Watson and Moritz (2000a, 2000b) looked at a collection of Australian students at various grade levels. Using Watson's statistical literacy framework, they developed six categories of individuals understanding of sampling. However, Watson and Mortiz's six categories are based on the size and type of sample that the student constructed. When sampling was examined within the frame of expectation, variation, and decision making, the student's conception of sample size was used as the point of reference for the student's understanding of sample (Watson & Kelly, 2006). Jacobs (1997) also looked children's understanding of sampling; however Jacobs focused on how students (fourth and fifth graders) evaluated different sampling methods and how those students drew conclusions based upon presented samples (with sampling methods). The APOS-Sampling theoretical perspective sets aside the size of the sample that the student constructs and instead focuses on what understandings the student has for sampling.

This theoretical perspective is related to that of Saldanha and Thompson (2002, 2007) in that this perspective takes a distributional approach to sampling. This is to say that the student's understanding of sampling is placed within his/her ability to keep in mind a sampling distribution. In fact, the work of Saldanha and Thompson greatly influenced this theoretical perspective. While Saldanha and Thompson laid out a three phase sampling process and a multiplicative conception of sampling, this perspective serves as an attempt to extend and reframe their work. The original MCS defines a hierarchy of levels but does not address how a student might move from one level to another. APOS-Sampling serves as one perspective that allows for a development framework to extend the MCS and suggest one way to get students to develop a richer understanding of sampling.

References

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the Process Conception of Function. *Educational Studies in Mathematics*, *23*(3), 247–285.

Clark, J., Kraut, G., Mathews, D., & Wimbish, J. (2003). The Fundamental Theorem of Statistics: Classifying Student understanding of basic statistical concepts. *Retrieved October*, *20*, 2007.

Dubinsky, E., & McDonald, M. A. (2002). APOS: A constructivist theory of learning in undergraduate mathematics education research. *The teaching and learning of mathematics at university level*, 275–282.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005a). Some historical issues and paradoxes regarding the concept of infinity: An Apos-Based analysis: Part 1. *Educational Studies in Mathematics*, *58*(3), 335–359.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005b). Some historical issues and paradoxes regarding the concept of infinity: An APOS analysis: Part 2. *Educational Studies in Mathematics*, *60*(2), 253–266. doi:10.1007/s10649-005-0473-0

Jacobs, V. R. (1997). Children's Understanding of Sampling in Surveys. (p. 37). Presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. Retrieved from http://www.eric.ed.gov/PDFS/ED411247.pdf

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction*, *6*, 59–98. doi:10.1207/s1532690xci0601_3

Lajoie, S. P., Jacobs, V. R., & Lavigne, N. C. (1995). Empowering children in the use of statistics. *The Journal of Mathematical Behavior*, *14*(4), 401–425. doi:10.1016/0732-3123(95)90039-X

Lipson, K. (2003). The role of the sampling distribution in understanding statistical inference. *Mathematics Education Research Journal*, *15*(3), 270–287.

Mathews, D., & Clark, J. M. (2007, July 25). *Successful Students' Concpetions of mean, Standard Deviation, and the Central Limit Theorem*. Unpublished. Retrieved from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf

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

Saldanha, L., & Thompson, P. W. (2002). Conceptions of Sample and Their Relationship to Statistical Inference. *Educational Studies in Mathematics*, *51*(3), 257.

Saldanha, L., & Thompson, P. W. (2007). Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, *2*(3), 270–297.

Sfard, A. (1992). Operational origins of mathematical notions and the quandary of reification - the case of function. In E. Dubinsky & G. Harel (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy*. Washington, D.C.: Mathematical Association of America.

Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. *Educational Studies in Mathematics*, *12*(2), 151–169.

Thompson, P. W. (1994). Students, functions, and the undergraduate curriculum. *Research in collegiate mathematics education*, *1*, 21–44.

Watson, J. M., & Kelly, B. A. (2006). Expectation versus variation: Students' decision making in a sampling environment. *Canadian Journal of Math, Science & Technology Education*, *6*(2), 145–166.

Watson, J. M., & Moritz, J. B. (2000a). Development of understanding of sampling for statistical literacy. *The Journal of Mathematical Behavior*, *19*(1), 109–136. doi:10.1016/S0732-3123(00)00039-0

Watson, J. M., & Moritz, J. B. (2000b). Developing Concepts of Sampling. *Journal for Research in Mathematics Education*, *31*(1), 44–70.