The “Fundamental Theorem" of Statistics: Classifying Student Understanding of Basic Statistical Concepts

Julie M. Clark
Hollins University
Roanoke, Virginia

Gertrud Kraut
University of Texas at Tyler
Tyler, Texas

David Mathews
Southwestern Michigan College
Dowagiac, Michigan

Joseph Wimbish
Huntingdon College
Montgomery, Alabama

1 Introduction

Numerous researchers and mathematics educators have noted that there is often a wide disparity between what was taught in a mathematics course and what remains in the students’ minds after the final exam (see [2], [12] and [26]). A previous study ([12]) analyzed interviews of a group of college freshmen immediately upon their completion with a grade of ‘A’ of an elementary statistics course. The goal of that study was to determine as precisely as possible the conceptions of the mean, the standard deviation, and the Central Limit Theorem held by the most successful students shortly after completing a course in elementary statistics . To understand the interview responses, the researchers in [12], Mathews and Clark, relied heavily upon the Action-Process-Object-Schema (or APOS) epistemological framework ([1] and [12]).

According to APOS theory, an *action* is any transformation of objects to obtain other objects. It is perceived by an individual as being at least somewhat external, as it has the characteristic that at each step, the next step is triggered by what has come before. An individual is said to have an *action conception* of a given concept if her or his depth of understanding is limited to performing actions relative to that concept. As noted in [12] an individual with a deep understanding of a concept would perform these same actions when appropriate, but he or she would not be *limited* to performing actions.

When an action is repeated and an individual reflects upon it may be *interiorized* into a *process*. That is, an internal construction is made that performs the same action. In contrast to an action, a process is perceived by the individual as being internal and under one’s control, rather than as something one does in response to external cues. An individual who has a *process conception* of a transformation can reflect on, describe, or even reverse the steps of the transformation without actually performing those steps. All eight students studied in [12] demonstrated that they had moved beyond an action conception to a process conception of mean. They had internalized the notion of numbers coming in, a certain computation taking place, and a number being output.

When an individual reflects on operations applied to a particular process, becomes aware of the process as a totality, realizes that this entity has properties, that transformations (whether they be actions or processes) can act on it, and he or she becomes able to actually construct such transformations, then he or she is thinking of this process as an *object*. We say that the process has been *encapsulated* into an object, and that the individual has an *object conception* of the concept. For many of the students interviewed in [12] the mean was neither a measure of central tendency, an object with properties, nor a cognitive entity. Of the eight students, three demonstrated the lowest level of understanding of standard deviation - an action conception. At this cognitive level, standard deviation was only a rule or formula to be given and followed, and students were unable to describe the algorithm. In sharp contrast to the situation for mean, none of the remaining five students demonstrated an *appropriate* object conception of standard deviation. Two of the five had developed inappropriate object conceptions of the standard deviation as representing distance between members of an ordered set. The other three students had inappropriate object conceptions of standard deviation as tied to single data points.

The most sophisticated cognitive structure described in APOS theory is that of a *schema*. An individual's schema for a certain piece of mathematics is that person's own cognitive framework which connects in some way all of the ideas that the individual either consciously or subconsciously views as related to the piece of mathematics. Researchers applying the APOS theoretical perspective attempt to describe a generic road map of understanding which includes individual mental constructs, their origins, and their relationships to one another. Such theoretical models are used to describe both elementary constructs and schemas and are referred to in APOS Theory as *genetic decompositions*. In [12] it was proposed that in order to develop an understanding of the Central Limit Theorem a student would need to have at her or his disposal object conceptions of mean, standard deviation, set, distribution, and sample, as well as an understanding of the differences between populations and samples. However, none of the eight students reported in [12] demonstrated a working knowledge of the Central Limit Theorem. One of the major goals of this study is to verify, or modify, the genetic decomposition of the Central Limit Theorem proposed in [12]. Naturally this required identifying students who had developed a viable understanding of this crucial theorem.

Another goal of the present paper is to add to the body of knowledge concerning the development of statistical knowledge in college students by building on the work done by Mathews and Clark.

2 Description of Study

The present paper extends the methods, ana lyses and findings of Mathews and Clark in the following four ways. First, the student population was expanded. Instead of selecting students from just one campus, three other campuses were involved. In addition, the population was no longer restricted to freshmen students. In this study the population was expanded to include upperclassmen and graduate students in both science and non-science areas.

Second, the nature of the sampling was altered. In [12] a census of all students receiving a grade of 'A' in any section of a beginning statistics course the previous semester had been taken. In this study different methods were employed to compile the sample. On campus T, the interviews were with all three willing 'A' students, on campus E the sample consisted of the three 'A' students for that semester, and on campus H all the 'A' students were contacted by campus mail and given an appointment. Those who responded were interviewed. Those who failed to respond to the first appointment were given the opportunity to reschedule. This resulted in a total of nine interviews across the three campuses. As in [12], the population from which the sample was collected was restricted to students who had received an ‘A’ in their statistics courses in the hopes of generating data that could be used both to verify the hypothesized genetic decompositions and to provide some information about the effectiveness of traditional pedagogy on these students in terms of cognitive development.

The interviews were conducted as audio-taped clinical interviews. The current investigators did attempt a uniformity in the interviewing process. The primary questions were:

1. What is meant by the word “mean” in statistics?
2. What is meant by “standard deviation? "
3. What is the “Central Limit Theorem ? "

The students were given the opportunity to answer these questions without prompting. If they had difficulty, they were given as much prompting as was necessary to elicit meaningful responses. The students were given the opportunity to write down anything they wanted. The written work then became a part of the student’s transcript. In many interviews, the discussions also touched on much broader topics, such as sampling, distributions, *Z*-scores, normality, and the nature of statistics.

As in [12] the primary framework used for interpreting responses was the APOS model and the following characterization of mathematical knowledge [1]:

> An individual’s mathematical knowledge is her or his tendency to respond to perceived mathematical problem situations by reflecting on the problems and their solutions in a social context and by constructing or reconstructing mathematical actions, processes and objects and organizing these in schemas to use in dealing with the situations.

To extend the analysis, other frameworks were incorporated to give multiple views of the same interview. In addition to APOS theory, Skemp’s ideas of instrumental and relational understanding [20] and Gray and Tall’s notion of a *procept* [4] were used as “lenses" through which the students’ interviews were viewed. Incorporating these perspectives provides a third way in which the methods and analyses of [12] are extended in the present paper.

Finally, by interviewing a more diverse sample of students, more comprehensive descriptions of cognitive developmental structures in statistics were possible, which led to new and expanded recommendations for future studies and the way statistics might be taught and learned.

3 Results and Revised Analysis

3.1 Revised APOS Analysis: Mean

As in the previous study ([12]), we found no students who were limited to an action conception of the mean. This was not surprising, as students limited to this understanding would be unable to calculate a mean in the absence of a formula. The action of averaging is relatively simple, and most students have performed this action many times before they enroll in a college-level statistics course.

However, moving beyond a process conception of mean is much more difficult. Three of the students in this study had not progressed beyond a process conception of the mean. Although they could perform the necessary actions, describe the process of computing the mean of a set of numbers, and in some cases reverse this process, these students appeared unable to conceive of the mean of a data set as an entity itself. They were unable to perform any actions on the output of their processes or to associate any meaningful properties with the means they computed.

For example, in the following excerpts from their interviews, in spite of prompting from the interviewer, Cathy and Gina[1] were unable to articulate any understanding of how a mean might describe a data set:

> **Interviewer:** But, when we talk about the mean as being a measure of sorts, what kind of measure is it?
>
> **Cathy:** Um, ..., besides an average I'm not sure what you're looking for.
>
> **Interviewer:** What does the mean tell us about a data set? What does it measure?
>
> **Gina:** Just the average of the data. That's it.

Both Cathy and Gina were able to provide the synonym of average for mean but neither appeared to be able to consider the mean of a data set as an object which could have properties.

The remaining six students all exhibited object conceptions of the mean. For example when asked, Helen was able to act on the mean by comparing it with another measure of center (the median), and she demonstrated that she sees the mean as an object with some properties – namely the mean tells us something about the corresponding set of numbers as a whole.

> **Interviewer:** Why would you use the median over the mode or median over the average or the mean, or can you think of when you might use one over the other?
>
> **Helen:** I think that the median is more when you want to know the absolute middle or absolute center of something, whereas the average is more of then you want to be able to compare everything, to compile everything into one number.

Most of the students' object conceptions of the mean, though readily apparent, were also weak. Many of the students repeatedly described the mean in a circular manner; they appeared able to use only the word "average" to describe what the mean is or tells us about a set.

> **Interviewer:** Ok. So you've written 1,2,3,4,5, and found a mean of 3. Great. Um, what is, what is that number 3 telling you about the data set?
>
> **Lou:** That, that is, well basically considered to be the average number for that. It is a, it is a projected value, I don't know how to put this. Uh, I don't know. I've over used the term average. But the average of the numbers, it's a number that you could use to reasonably identify with all the numbers of that group.

Students like Lou and Helen realize that the mean summarizes the entire set, but they fail to understand *how* it summarizes the set, or what it indicates about the set. Their understanding of this object does not necessarily carry with it the notion of the mean as a measure of central tendency.

In addition, the language that these students used indicates that although they can think of the mean as an object, their initial responses about the mean are dominated by focusing on the process of computing a mean. Although he later ascribes several significant properties to the mean and declares that the mean is "a number that you could use to reasonably identify with all the numbers of that group," when Lou is asked what is meant by the word 'mean', he immediately responds with the following process description:

> **Lou:** It's another word for average. Meaning, uh, what does it mean, ...all of them, uh, all pieces of information added together and divide by the number of pieces of information.

Jack also appears to think of the mean as something one 'does'.

---

[1] The names of the students have been changed to protect their privacy.

Jack: So. *To do* the mean you just add the points... and divide by the number[2] .

As does Helen:

Interviewer: Now suppose that your roommate was studying for a statistics test ...and your roommate has to know everything there is to know about the average. What would you try to tell her?

Helen: Um, how to *do it*, you know, how to work with the numbers.

Clearly, the statistical concept of mean is what Tall and others [4] (CHECK CITE) have described as a *procept* - an entity which has both process and object (or concept) characteristics. To understand the concept of mean as a procept, the student must become aware of this duality of characteristics and be able to distinguish between the process of computing the mean and the result of that process.

The APOS framework provides language to model the individual cognitive constructs available to a given individual. However, since an individual's mathematical knowledge is her or his *tendency* to respond to a perceived problem situation, he or she may or may not respond in a particular situation by accessing the appropriate mental construction even if it is at her or his disposal. For example, a student who possesses an object conception of the mean may tend to reason almost exclusively with the underlying process and therefore be unlikely in certain situations to invoke the mean as an object. For such cases we find it useful to apply Vinner's model of an *evoked concept image* [22]. This gives us language to describe what we see in these six students – they often bring to bear a piece of their understanding which is less powerful than required for the situation. In APOS terminology students are considered to have an object conception of the mean because they have demonstrated this understanding at some point. But since the predominance of their thinking tends to indicate the corresponding process, we may often observe their *evoked concept image* containing merely a process conception of the mean.

3.2 Revised APOS Analysis: Standard Deviation

As in [12], one-third of the students in this study appeared limited to an action conception of the standard deviation. In some cases, they did not manage to explicitly indicate this conception in the interview, but as they were 'A' students, we make the assumption that they have successfully performed the action of computing standard deviation at some point in the past. All three of the students in this situation demonstrated an inability to compute or discuss the standard deviation in the absence of a formula for computing it. Consider the comments of Matt:

Interviewer: Are you trying to think of a formula?

Matt: Yeah, is it...the mean ...wait

Interviewer: You can write it down if you want.

Matt: Sum of...Is it the sum over...no...yeah...is it the sum of...It's the sum of something divided by the mean.

None of the six students who had progressed beyond an action conception of standard deviation appeared to be limited to a process conception. They all had object conceptions of standard deviation, but the richness of their conceptions varied greatly. This appeared to be a reflection of two distinct tracks in the development of their process conceptions. In the language of Skemp [20] some students appear to have progressed to a *relational understanding* of the process of computing the standard deviation, while others have only an *instrumental understanding* of this process. The process of standard deviation, if understood relationally,

---

[2]Italics added for emphasis.

would involve the measurement of successive distances from the mean of the data. Instrumental understanding of the process might involve merely subtraction and squaring of numbers with no cognitive link to the concept of distance. In either case we would say that the algorithm for computing standard deviation would have been interiorized to a process, and with such understanding a student would be able to conceive of calculating a standard deviation in the abstract in the sense of: input data, perform calculation, output number.

If the process of computing standard deviation is understood only instrumentally then we might expect that the cognitive object resulting from encapsulation of this process would not carry with it a conscious connection to mean. If, on the other hand, the process is understood relationally, the encapsulated process should carry with it this connection. (Note that in either case the resulting object is a number and students are likely to associate rudimentary object characteristics with any number. Indeed many of the characteristics of a cognitive object come from cognitive connections with associated concepts - not merely from the process which has been encapsulated. This can mislead students, teachers, and researchers about the true level of a student's cognitive development.)

Gina provided us with a striking example of this instrumental understanding. She was able to describe the process of computing a standard deviation, and she demonstrated that the resulting numerical output had some properties. Thus she had an object conception of the standard deviation. But the process she had interiorized was flawed, and she demonstrated no awareness of the connection between the resulting object standard deviation, and the mean which she used explicitly in the algorithm.

Interviewer: Can you tell me anything about the standard deviation? How do you compute it, what does it mean?

Gina: The standard deviation is how much the data, is a, deviates from each other. And how you compute it is you, you take the mean

Interviewer: Yeah, do it. Do it on the paper, and then describe what you're doing.

Gina: Okay, you take the mean, and then you subtract each data from the mean.

Interviewer: All right. Just, yeah, you can just...

Gina: Ok. And then we write...I can't remember if it was $\bar{x}$ minus that, but you take the difference between the mean and the data.

Interviewer: Go ahead.

Gina: So it'd be 1,1,2, and then you would square them?

Interviewer: Mmm hmm.

Gina: (pause) And you would add up those squares.

Interviewer: Mm hmm.

Gina: And that is your, um, variance. That's right. I forgot how you denote variance, s.

Interviewer: Right.

Gina: And then you would take the square root of that, which is whatever.

In fact, although Gina does realize that standard deviation is a measure of how much the data deviates, she never recognizes it as an *average* deviation from the mean, in spite of extensive prompting from the interviewer. For her, standard deviation remains a measure of deviation for each individual number within the set.

Interviewer: So the standard deviation is the square root of the variance. And you've named both of them. Now when you computed the mean, you added them all up, and divided by the number of items.

Gina: Mm hmm.

Interviewer: Was there anything like that when you were compute the, variance? Did you just add up the differences, or, was there also division involved, somewhere?

Gina: I think we just added up variances. Maybe we did, I don't remember.

Gina does not seem to have a relational understanding of standard deviation. It was possible for her to compute $(x-\bar{x})$ without consciously linking the cognitive objects of mean and standard deviation. It is even possible that the majority of the "object status of standard deviation" for her results from her understanding of numbers as objects and/or the object status of "output". Her interview clearly indicates a lack of relational understanding of the process of computing a standard deviation.

A crucial distinguishing factor between understanding the algorithm for standard deviation instrumentally versus relationally appears to be the richness of understanding of the expression $(x-\bar{x})$. This quantity can be calculated mechanically without awareness that a distance from the mean is being measured. In the worst case, $\bar{x}$ is simply a number that was computed first, and x is the number in the first column of a table. In fact students in this situation may express confusion about whether to compute $(x-\bar{x})$ or $(\bar{x}-x)$ as Gina did. If a student views $(x-\bar{x})$ as merely a difference (not as a distance), then he or she may also have difficulty viewing the remaining process as an average.

Several of the students who exhibited a weak object conception of the standard deviation had clearly linked this concept with that of a Z-score. This connection is certainly understandable since the first step in both processes is to subtract the mean from a single piece of data in order to measure its distance from the mean. The picture is further muddled since the Z-score computation also requires dividing by the standard deviation in order to tell us *how many standard deviations away* from the mean this piece of data lies. Consider the initial response of Helen:

Interviewer: Okay another term you may have used is the standard deviation, do you remember what the standard deviation was?

Helen: Um, how many it is away from the ... Like I can picture it in my head..

Interviewer: Well, write it down.

Helen: It's like more than 2 standard deviations away from the mean.

As seen in the excerpt below, Gina suggests that standard deviation and Z-scores are the same type of measure, although later in the interview she is unclear about whether she believes they measure spread or position.

Interviewer: Besides the standard deviation, do we have another measure that would kind of tell us how the data is spread out?

Gina: Mm, (pause)...Z-score maybe. Is that what you're trying to get at?

And much like some students in [12], Jack initially describes the standard deviation as if it were in fact a Z-score.

Jack: Okay standard deviation is... variation from the mean. How far...like take a particular point and ... how far from the mean it is.

Interviewer: Okay. Uh, so can you sorta give me an example?

Jack: Uh huh. Let me think. Let me think. I, did this on my calculator, so...(laughs).

Interviewer: Yeah. That's true. You're probably gonna have a hard time actually computing it, but if you could ball park it for me.

Jack: Okay, so, so we've got 2,3,4,5, and 10, and the mean is around 4. So say we look at 5 which is about, almost one standard deviation away.

Later in the interview, Jack correctly discusses the standard deviation as a measure of 'consistency' which tells one something about the entire set. His confusion may be because he has a vague notion that the standard deviation and *Z*-score both are supposed to measure distance from the mean, but he is unclear about how they do this, or what the difference between them is. All three of these students appear to have a strong *concept image* of standard deviation as a *Z*-score.

A fourth student, Lou, also demonstrated an object conception of standard deviation. His standard deviation was tied to the entire set, did not appear to be confused with his understanding of *Z*-scores, and he articulated that the smallest possible standard deviation was zero. He even described the standard deviation as average distance from the mean. However, when he attempted to reconstruct the actual process, he averaged only the two extreme distances from the mean. Apparently he developed a "hollow' or what Sfard has termed a "pseudo structural" object. As Sfard notes [19] it is, in general, difficult for students with a pseudo structural object to develop a rich conception of the object.

Two of the students in this study did appear to have rich, appropriate object conceptions of the standard deviation. It is worth noting that both students had strong mathematical backgrounds (Rick had already obtained an electrical engineering degree, and Will had completed two semesters of calculus at the time of his statistics course.) Both of these students provided almost textbook perfect descriptions of the standard deviation:

Will: Well, the standard deviation, um, is seeing how spread out the group is from the mean.

Interviewer: Mm hmm.

Will: Um, the further they are from the mean, the larger the standard deviation is. It's a measure of, I guess it's the average distance of the squared difference from, the mean.

Interviewer: Can you ball park what the standard deviation of that set, which was again, 7,8, 9, and 10. What it would be?

Will: About one.

Interviewer: About one. Um, and can you sort of tell me how you're ball parking that?

Will: Uh, because the end values are 2 away, and the, uh, the middle values are 1 away, and the, it's about an average distance, uh, since you're squaring it and then you're going to take the square root of it.

Interviewer: Okay.

Will: You're coming out with about an average distance.

Rick could even relate the standard deviation to the mean via the coefficient of variation:

Interviewer: All right. Do you have any other measures that could tell you a little bit about a data set?

Rick: Uh, we can look at the, the standard deviation. Which would tell, which would give me an indication of the uh, what would be a proper word, the dispersion of the data set.

Interviewer: Mm hmm.

Rick: In other words, it would tell me if most of my data is, is within an error range or if it's spread out very far.

Interviewer: Okay.

Rick: I can look at, uh that'd give me a numerical number. Uh, a standard deviation of a hundred might not be very much if most of my datas are within values in the millions.

Interviewer: Mm hmm.

Rick: I can look at the, there's a coefficient that I recall. Coefficient variance, I forget the exact term, but I can get the coefficient of the relationship of that dispersion. So I might end up with a very small number. Now no matter what my numerical values are, if my coefficient is small or large, I can get the, the, the proportional dispersion.

Rick and Will would be said to have a proceptual understanding of standard deviation. They have apparently come to terms with the dual nature of this measure of spread by synthesizing the process and concept natures as two different sides of the same object.

3.3 Revised APOS Analysis: The Central Limit Theorem

None of the students interviewed in [12] made useful statements about of the Central Limit Theorem. In this follow-up study of nine 'A' students (including freshmen, sophomore, juniors, and post-graduate students) only one demonstrated a reasonable understanding of the Central Limit Theorem. Three of the students had no recollection of the theorem or insisted that this topic had not been covered in their course. These students had no evoked concept image of the Central Limit Theorem at all. Another three students remembered hearing the name and said their instructors had stressed the importance of this theorem, but they could not give any description of the theorem or place it in any meaningful context. For two students, the name 'Central Limit Theorem' actually invoked the Empirical Rule.

All of the eight students who did not have a viable understanding of the Central Limit Theorem were missing one or more of the cognitive constructs proposed by the researchers in [12]. For example, Rick had strong object conceptions of both mean and standard deviation, but did not demonstrate an understanding of distributions. In fact, when asked about requirements for a distribution to be normal, he indicated that whether or not a distribution was normal is determined by whether the book states it is. Rick was one of the students who had "no recollection of the Central Limit Theorem."

Jack had object conceptions of set, mean, and standard deviation (though he had a concept image of the standard deviation as a *Z*-score), and as evidenced by the excerpt below was clearly able to distinguish between sample and population quantities:

Jack: The population mean takes the entire population. Your entire, what your, every data point in the population. And a sample is just one part of the population.

Interviewer: And would they be the same?

Jack: No. They would be different because...Well it can be representative. The sample mean can be representative of the population, but it is not the entire population. It's not entirely accurate.

With these pieces, Jack has begun to build a Central Limit Theorem schema:

Jack: Okay the Central Limit Theorem deals with the um, ...I'm confused. Either it's mean of means, to do with the mean of means...

Interviewer: Uh huh.

Jack: and there's a formula to compute, find the mean of means. And the Central Limit Theorem, does it deal with...um, I'm trying to think...et me think, let me think, let me think...I'm blanking. (laughs).

Interviewer: Not surprising. Do you have any recollection of uh when you might use the Central Limit Theorem?

Jack: Maybe when...when you have a uh, when you want to estimate a mean, or you want to estimate a sample, a size you need...

Interviewer: Uh huh.

Jack: Use the theorem to determine what size you need to be accurate, or to be close to the correct answer.

But Jack's notion of distributions was weak. He seemed to assume that 'everything is normal,' and therefore would most likely find it difficult to understand the significance of the theorem.

Will was the only student interviewed who exhibited a reasonable understanding of the Central Limit Theorem. However, he was an unusually talented student by any traditional measure (GPA, GRE scores, teacher evaluations, etc.) At the time of this course he was a sophomore, double majoring in chemistry and biology, and his science and mathematics background was quite strong. In short, he was an intrinsically motivated, gifted student.

Throughout his interview Will demonstrated strong object conceptions of set, mean, standard deviation, sampling, and distribution. When asked about the theorem he responded :

Will: Okay, the Central Limit Theorem says if you take, if you take groups, if you take a, a group of groups from a population, if you have a population and then you take out groups, and you take their, um, the mean of each of those groups, then you will come out with, and you average those group means, then, then that should be the mean of the population, of the original, the whole population.

Interviewer: Uh, okay, so the Central Limit Theorem has a very obvious connection to this concept of mean which we talked about earlier. Is it, does it have any connection to the standard deviation?

Will: Um, the standard deviation is related in terms of, how many groups you have.

Interviewer: Mm hmm.

Will: An the number of um, of each group. I mean the, the number of each sample size.

Interviewer: Okay.

Will : So the larger the sample size, um, uh, the smaller your deviation from the or, the population standard, standard deviation, I mean the population, um, standard deviation to the smaller group's standard deviation.

Interviewer: In general how do you know when you're supposed to use the Central Limit Theorem?

Will: Um, only when you have groups of, of, of, a well only when you've multiple samples from the same population. And generally, it's some sort of test of, the uh, population's actual mean. You know. You know there's usually some kind of claimed mean, or you're trying to estimate the mean, and you can only do that from the samples.

Will's success in recalling the Central Limit Theorem, and the other students' lack of such success, lend support to the contention that a schema interpretation of the Central Limit Theorem is useful to teachers and researchers. And with such an interpretation the apparent necessity of the individual cognitive constructs described by the authors of [12] has pedagogical implications.

On the other hand, understanding the individual constructs — even understanding them as procepts — without the connections provided by a schema does not provide a rich enough model of understanding the Central Limit Theorem to guide an improvement of our pedagogical treatment.

4 Conclusions and Pedagogical Suggestions

The revised APOS analyses given in the previous section, when combined with the results reported in [12] and when viewed through the expanded set of theoretical perspectives, provide an extremely rich model of the cognitive development of 'A' college students' understanding of mean, standard deviation, and Central Limit Theorem as a result of traditional pedagogical approaches. None of the combined seventeen students from four distinct campuses were limited to action conceptions of the mean. The action of computing a mean appears to be fairly easy to interiorize into the process of "adding the data then dividing by the number of items in the list," and 'A' students consistently appear to have interiorized this action into a process. However there were three distinct trends observed beyond this level. Some of the 'A' students were limited to this process conception of the mean, some had developed a weak object conception that seemed to coexist with a predominately process concept image, and some had developed a robust, rich object conception that is compatible with the understanding that mathematicians and statisticians would expect.

It is worth noting that only two of the seventeen students in the combined study appeared to have progressed to this rich object conception. Such disappointing results indicate that the primary goal of new pedagogical treatments of the mean must be to overcome the strongly invoked process image of mean which most students bring with them to their college courses. We need pedagogical strategies that will create conflicts between the students' expectations and their results, thereby dis-equilibrating the students, and promoting the necessary reflection and maturation that will enable them to encapsulate the process (of mean) into an object. What specifically helps dis-equilibrate students who already have a process conception is the necessity for working with mean as an entity. We must create situations involving the mean which require its use as an object. Improved results might be obtained if students confront situations in which the mean is more appropriate (as a measure of center) than the median, or mode, and vice versa. Note that for many of the data sets students see in traditional course, all the measures of center are essentially equal, so having several is redundant and only serves to reinforce their misperception that these measures are equivalent. Students should be repeatedly exposed to data sets which are *not* normal. In particular students must be forced to deal with sets of data for which there is no mean (categorical data), or in which the mean is meaningless as a descriptor of the set.

Students also need to be put into situations which encourage reflection about the mean as an object. Ideally students would work through activities which demand that they find and use properties and limitations of this measure of center.

With respect to standard deviation, the results of this study suggest that the discouraging cognitive understandings of 'A' level students seen in [12] may, in fact, be representative of the outcomes of most traditional courses nationally. More than a third of the students in both studies had not progressed beyond action conceptions of standard deviation, and in some cases these action conceptions were not demonstrated in the interviews, but only inferred from the students success in their courses. As in [12] the same inappropriate object conceptions of the standard deviation as either a measure of distance between members of an ordered set or as a Z-score, which measures something for individual data values, were seen in many of the nine students interviewed for this study. It was quite striking that only two of the entire group of seventeen students had developed relational understandings of the process of computing a standard deviation and moved onto rich object conceptions of this concept.

As we inferred from the process for computing standard deviation, the key distinguishing factor in the richness of the students' object conceptions was whether or not an intermediate step (averaging distances from the mean) was relationally understood. So we need activities which will foster a relational understanding of the process of computing a standard deviation by stimulating reflection on the meaning of (x-$\bar{x}$) as a distance. Students should be asked to experiment with measuring spread before they are introduced to the idea of a standard deviation or shown a formula for computing it.

As an extension of this principle, algorithms which do not carry with them the process (e.g. - the alternative formula for standard deviation) probably promote instrumental understanding and actually inhibit relational understanding. This is an argument for using calculators or computers for calculations with large data sets: otherwise students are forced to use algorithms that do not engender relational understanding of the object.

We also need to devise a strategy for untangling *Z*-scores from standard deviation. That this phenomena was seen in these interviews as well as in those of [12] suggests a strong need for pedagogical strategies that will promote *distinct* object conceptions of standard deviation and *Z*-score. Traditional pedagogical approaches appear to create a pedagogical obstacle for the students. **(REFERENCE NORMAN AND PRICHARD.)** After a brief section which requires computations of standard deviations, most texts simply supply students with the numerical value of any standard deviations they need in subsequent chapters and exercises. The remainder of the course often makes extensive use of *Z*-scores, but never again asks students to understand the meaning of the standard deviation.

The pseudo-structural object conceptions of standard deviation demonstrated by some students may have been the results of classroom pedagogies which depended almost exclusively on a calculator (or computer) to compute standard deviations. This may indicate that first students need extensive practice computing average distances from mean for themselves, then reflection on the significance of the calculation of (x-$\bar{x}$), before moving to use of calculator or computer. (Difficulty of the algorithm results in memorization - REFERENCE THOMPSON (COPING, 1994).

New pedagogical strategies should also include activities that would promote the formation of an object conception of distributions. Students should be repeatedly exposed to more than just the standard normal and *t*-distributions. Activities should ask students to compare various distributions, use *large* data sets, and continually require the use of tests for normality. In fact a significant portion of standard exercises might be dead-end problems - situations in which the assumption of normality is violated, and therefore techniques based on this assumption would be invalid or inappropriate.

Students also need to understand samples as cognitive objects. Activities which ask them to compare samples with other samples would help. A pedagogical tool that is vastly underused is multiple sampling from the same population. This tool can also help reinforce the distinction between the process of sampling and the object of sample. Implicit in all of this is the constructivist perspective that says it is important for the student to be actively engaged in doing and reflecting on the mental constructions called for by the course. This refection is not necessarily a by-product of 'seeing' examples -as evidenced by Glover (p. 19 complete reference)

With all of these concepts the most important consideration may be to create a need for the cognitive object. In [6] Harel defines the Necessity Principle:

> For students to learn, they must see a need for what we intend to teach them, where by "need" is meant intellectual need, as opposed to social or economic need. [6]

Nowhere is this need more crucial than in developing Central Limit Theorem schema. Most of these students in this study appeared to view the subject of statistics as merely the collection, organization and manipulation of data. When asked why we study statistics, they inevitably made reference to descriptive statistics, never to the inferential nature of statistics. Although they had been successful in creating confidence intervals and performing hypothesis tests in their courses, they did not connect these ideas with their understanding of descriptive statistics. They appeared to view these two branches of statistics (descriptive and inferential) as unrelated and independent. They had little need for the Central Limit Theorem because, in their experience, virtually everything was normal anyway.

We have already seen that the words 'Central Limit Theorem' do not evoke a useful concept image in most students. However, as educators, we view its understanding as *the* crucial step from descriptive to inferential statistics. Without that understanding students will be limited to an instrumental focus of inferential techniques such as hypothesis testing and estimating with confidence intervals. Since we have plenty of readily accessible machines that can do the instrumental calculations for us, understanding is what the humans must bring to this situation. So, rather than skipping over or reducing emphasis on this theorem, instructors should incorporate activities that will promote both a rich understanding of the theorem and an intellectual need for the theorem. Perhaps it should also be re-named The "Fundamental Theorem" of Statistics, so that its very name suggests its importance, as the Fundamental Theorem of Calculus does.

References

[1] Asiala, Mark, Brown, Anne, DeVries, David, Dubinsky, Ed, Mathews, David, & Thomas, Karen (1996). A framework for research and curriculum development in undergraduate mathematics education. *Research in Collegiate Mathematics Education II*, *6*, Conference Board of the Mathematical Sciences: Issues in Mathematics Education, 1–32.

[2] Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44–63.

[3] Glasersfeld, E. von. (1995). *Radical constructivism: A way of knowing and learning.* London: Falmer Press.

[4] Gray, and Tall, D.

[5] Hardiman, P.T. (1984). Usefulness of a balance model in understanding the mean. *Journal of Educational Psychology, 76*, 792–801.

[6] Harel, G. (1997). The Necessity Principle and Its Implications to the Teaching of Mathematics. Presented at MAA, 1997. [7] Hong, E., & O'Niel, H.F. (1992). Instructional strategies to help learners build relevant mental models in inferential statistics. *Journal of Educational Psychology, 84*, 150-159.

[8] Landwehr, J., Swift, J., & Watkins, A. (1987). *Exploring surveys and information from samples*. Palo Alto, California: Dale Seymour Publications.

[9] Li, K., & Shen, S. (1992). Students' weaknesses in statistical projects. *Teaching Statistics, 14*, 2–8.

[10] Mason, J. ICME 1996 JOE -

[11] Mason, J. (1985). *Thinking Mathematically*. Workinham, England: Addison-Wesley Publishing Company.

[12] Mathews, D., & Clark, J. Successful Students' Conceptions of Mean, Standard Deviation and the Central Limit Theorem. Manuscript submitted for publication.

[13] Mokros, J., & Russell, S. (1995). Children's Concept of Average and Representativeness. *Journal for Research in Mathematics Education, 26*, 20–39.

[14] Piaget, J., & Garcia, R. (1989). *Psychogenesis and the History of Science*. (H. Feider, Trans.) New York: Columbia University Press. (Original work published 1983).

[15] Romberg, T. (1992). Perspectives on scholarship and research methods. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning*. (pp. 49–64). New York: Macmillan.

[16] Schoenfeld, A. (1994). Some notes on the enterprise (Research in collegiate mathematics education, that is). In *Research in Collegiate Mathematics Education I, 4*, 1–19.

[17] Sfard, A., (1987). Two conceptions of mathematical notions, operational and structural. In A. Borbàs (Ed.), *Proceedings of the 11th Annual Conference of the International Group for the Psychology of Mathematics Education* (162–169). Montreal: University of Montreal.

[18] Sfard, A., & Linchevski, L. (1996). The gains and the pitfalls of reification: The case of algebra. *Educational Studies in Mathematics*, (in press).

[19] Sfard, A., (1991). On the dual nature of mathematical conceptions. *Educational Studies in Mathematics, 22*, 1–36.

[20] Skemp, R. (1976). Relational Understanding and Instrumental Understanding. *Mathematics Teaching*. Association of Teachers of Mathematics.

[21] Steffe, L., & Glassersfeld, E. (1988). On the construction of the counting scheme. In *Construction of Arithmetical Meanings and Strategies*. New York, Springer.

[22] Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*, 151–169.

[23] Thompson, P. (1994). Students, functions, and the undergraduate curriculum. In *Research in Collegiate Mathematics Education I, 4*, 21–44.

[24] Thurston, W. (1994). On proof and progress in mathematics, *Bulletin (New Series) of the American Mathematical Society*, *30*, 161–177.

[25] Velleman, P., & Hoaglin, D. (1992). Data analysis. In D.C. Hoaglin and D. Moore (Eds.). *Perspectives on contemporary statistics, MAA notes 21* (pp. 19–39). Washington, DC: Mathematical Association of America.

[26] Vinner, S. (1992). The function concept as a prototype for problems in mathematics learning. In G. Harel and E. Dubinsky (Eds.). *The concept of function: Aspects of epistemology and pedagogy, MAA notes 25* (pp. 195–213). Washington, DC: Mathematical Association of America.