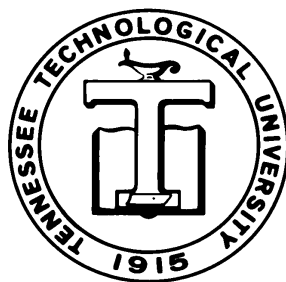

DEPARTMENT OF MATHEMATICS
TECHNICAL REPORT

ERRORS AND MISCONCEPTIONS IN
COLLEGE LEVEL THEOREM PROVING

DR. ANNIE SELDEN
AND
DR. JOHN SELDEN

AUGUST 2003

No. 2003-3



TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

This paper first appeared as
"Errors and misconceptions in college level theorem proving"

by

Annie Selden and John Selden

in the

*Proceedings of the Second International Seminar
on Misconceptions and Educational Strategies
in Science and Mathematics*

Joseph D. Novak, Editor

Vol. III, Cornell University, July 1987, 457-470.

[N.B. It has been made available in this form
because it is difficult to obtain from many libraries.]

ERRORS AND MISCONCEPTIONS IN COLLEGE LEVEL THEOREM PROVING

Annie Selden and John Selden
Mathematics Department
Tennessee Technological University

INTRODUCTION

In this paper we describe a number of types of errors and underlying misconceptions that arise in mathematical reasoning. Other types of mathematical reasoning errors, not associated with specific misconceptions, are also discussed. We hope the characterization and cataloging of common reasoning errors will be useful in studying the teaching of reasoning in mathematics.

Reasoning in mathematics is no different from reasoning in any other subject. Mathematics does, however, contain many exceedingly long and intricate arguments. The understanding of such arguments is essential to the correct application of mathematics as well as to its continued development. Indeed, in mathematics it is the proofs that provide the strong consensus on the validity of basic information that is characteristic of a science. From this point of view, the idea of proof in mathematics corresponds to those special techniques of observation and experimentation which are essential to the other sciences.

The examples presented here arose as student errors in a junior level course in abstract algebra, taught several times at universities in the United States, Turkey, and Nigeria. A principal goal of the course is the improvement of reasoning ability. Before giving a brief description of the course and the methods used in teaching it, we will comment on a widely held conception of the nature of mathematics and its teaching. The inadequacy of this concept interferes with a student's ability to benefit from mathematics instruction, and in particular, delays improvement in his reasoning.

THE STATIC VIEW OF MATHEMATICS

One way to view lower division college mathematics is that it consists of (1) concepts, (2) algorithms for solving problems, and (3) implementations of these algorithms, including the selection of the proper algorithm. As part of this view, the correctness of an answer depends entirely on selecting the right algorithm and on implementing its steps correctly. We will call this the static view because it suggests that most important work on this level depends only on the implementation of unchanging algorithms. Increased mathematical competence is seen as equivalent to knowing more algorithms.

Here is an example of an algorithm as we mean it in this context. Consider a continuous function f defined on an interval $[a,b]$. To find the maximum value of f on $[a,b]$: (1) find the derivative f' of f , (2) set $f' = 0$ and solve the equation, (3) find where f' does not exist, (4) evaluate f at all of the above points and at a and b , (5) the largest of these values is the maximum.

Of course this simple view of the nature of lower division college mathematics leads to a simple view of how it should be taught. The concepts should be explained, the algorithms should be provided, and the implementation of the algorithms should be practiced.

We think this is the conception of mathematics and its teaching held by many college students. We do not claim that they consciously ascribe to a description such as this or that they think about this topic at all. Rather, we note that they act as if they did and that they do so persistently.

Any teacher of lower division college mathematics can observe this by asking students to solve problems for which they have not been provided algorithms. Even if the teacher explains that this will require combining familiar techniques used on previous problems, many students will be disturbed. Often students who are reluctant to characterize mathematics or its proper teaching are definite about what it is not; it is not solving problems without instructions.

Turning now to the textbooks used in teaching lower level college mathematics, we see that their structure is in agreement with the static view. Emphasis is placed on explaining concepts and algorithms clearly and on sample solutions and practice problems. It is true that a scattering of theorems and proofs is usually included. Most teachers, however, think this material is more or less ignored by students and it is tempting to think authors agree. Often theorems are written in a style which makes them shorter and more memorable but harder to link to their proofs. It is second nature for mathematicians to expand such writing, adding missing quantifiers etc., so that they can understand the proofs. Such amplification is beyond most students who have no training or practice with it. Indeed we have occasionally found a theorem in a calculus book which could not be unambiguously understood by anyone not already familiar with the theorem.

INADEQUACY OF THE STATIC VIEW

Problems arising outside a mathematics course often do not match exactly the algorithms covered in the course. This is inherent in the general nature and broad applicability of mathematics and cannot be avoided merely by expanding the list of algorithms covered. Applications of mathematics, even on the lower division college level, will often require the creation of a new algorithm, at least in the sense of altering and combining familiar procedures.

Since, even without instruction, everyone has some ability to adapt algorithms to new problems, we will describe an example illustrating the impoverished level of this ability in a typical first calculus class. It is this that renders the static view of mathematics inadequate.

Consider two related problems: (1) Given a curve and a point on it, find where the tangent line at that point crosses the x -axis. (2) Given a curve, find a point on it so that the tangent line at that point passes through the origin. The solutions to these problems involve the same techniques: finding and evaluating derivatives and formulating and solving linear equations. Suppose an algorithm for the first problem is provided in class and the second problem is not mentioned. Our experience suggests one can expect at least 75% of the students to be able to work the first problem and less than 20% the second.

Clearly, if mathematics is to be widely useful, the static view of it is inadequate. Even on the lower college level, mathematics should include the creation of algorithms, at least in the sense of combining and altering known techniques. Since correctness of a newly created algorithm cannot be ascertained by appeals to authority, reasoning should also be regarded as an integral part of mathematics and its teaching. We do not mean to suggest that reasoning on this level should be in the form of proofs, but some examination of the correctness of algorithms should be included.

Perhaps this point can best be illustrated by looking outside mathematics to programming courses. Students are required to produce their own algorithms and it often happens that these algorithms do not perform as expected. As a result, the validation of programs through testing is regarded as an important and necessary component of the discipline.

The inadequacy of the static view of mathematics not only limits the usefulness of early college mathematics courses, it also inhibits the development of informal reasoning skills. These provide the foundation for the more formal reasoning required in proofs.

THE ROLE OF PROOFS

Proofs are not only essential to the development of new mathematics, their proper reading is an integral part of the understanding of advanced topics. Reading a proof is much more active than generally supposed. Subtle questions must be asked and answered. It is difficult to know when a student is doing this correctly, but it is easy to see if he has written a proof correctly. Since skills in reading and writing proofs are interdependent, often students are asked to write proofs as the only clear indication of their understanding of a topic.

ABSTRACT ALGEBRA

Abstract algebra, the course from which our examples are selected, is a three hour per week year sequence usually taken in the third year of college. At this stage in a student's education the rapid development of reasoning skills is important and we have taken it as a major objective of the course.

Abstract algebra is a particularly good vehicle for teaching reasoning because the proofs are less complicated than, say, advanced calculus, and the notation is not particularly complex. It is also useful that students have not seen this topic before and it is difficult for them to lift proofs from textbooks, so they must rely on their own reasoning and ideas.

THE SOCRATIC-MOORE METHOD

We provide the students with a short set of notes containing definitions, theorems, problems, and occasional examples, but no proofs or solutions. We do not give formal lectures although we answer questions. Students present their work in class and we provide evaluations, detailed criticisms and suggestions¹.

In the mathematical community this technique is sometimes called the Moore method after the late R. L. Moore who practiced it with remarkable success (Forbes, 1971). Of course in a broader setting this sort of teaching has a long history and calls to mind the methods of Socrates. There are many versions of this method, and seemingly insignificant variations in it may greatly alter its effectiveness.

REASONING ERRORS

Students make a great variety of reasoning errors in attempting proofs. We feel some of these errors are based on underlying misconceptions, while others, although repeatedly observed, are of a technical or other nature. Students persist in making both types of errors². For the former, we offer our views as to the possible underlying misconceptions, that is, we give a general rule or idea which, if believed by a student, would result in that type of error. These errors are taken to have a rational basis, and we comment on how they might come about³. For the latter, we sometimes speculate on the underlying causes of the errors, but do not see them as conceptual in nature. Each type of error is illustrated with one or more actual student “proofs”⁴.

REASONING ERRORS BASED ON MISCONCEPTIONS

M1. Beginning with the conclusion, arriving at an obvious truth and thinking the proof is complete. Of course, this provides a valid argument if and only if the steps are reversible.

The misconception consists in thinking that one valid technique of proof begins with the conclusion and ends with a known fact. However, this is not acceptable as it is often difficult to arrange a proof into a sequence of discrete steps, each of which can easily be checked for reversibility. This misconception may have arisen from methods learned in secondary school for verifying trigonometric identities and solving equations. Also, since a good heuristic for discovering a proof is to analyze the meaning of the conclusion, college students may have seen this presented in class, along with a statement that all steps are reversible. They, thus, could easily be confusing discovery with proof.

Example

Theorem: Let G be a group such that for all g in G , $g^2=e$, where e is the identity of the group. Then,

- (i) for all g in G , $g = g^{-1}$ and,
- (ii) G is commutative.

“Proof” of (ii) having proved (i): To show G is commutative means, for all a and b in G , it must be that $ab=ba$. Multiplying $ab=ba$, by the appropriate inverses, and using part (i), one gets $a=bab$, and $b=aba$. Now,

$$(1) b * = aba * = (bab)ba = ba(bb)a = baea = baa = be = b,$$

and

$$(2) a * = bab * = (aba)ab = ab(aa)b = abeb = abb = ae = a.$$

Comment

There are no errors present in lines (1) and (2). In this case, the steps indicated with a "*" are not reversible; they are equivalent to $ab=ba$.

M2. Names confer existence. This error occurs from failing to distinguish between symbols for things whose existence is established and symbols for things whose existence is not established. Often this error occurs when a student attempts to solve an equation, without questioning whether a solution exists. We asked a class of precalculus college students to solve the equation $\cos x = 3$ in order to test whether they could apply the fact that the range of $\cos x$ is $[-1,1]$. Most tried unsuccessfully to manipulate the equation to find x and were unable to reach the proper conclusion.

A different example of this type of error occurred when an abstract algebra student attempted to prove that a semigroup in which the equations $ax=b$ and $ya=b$ always have solutions is a group. One must first establish the existence of an identity element, e , and then show that each element, g , in the group has an inverse. The student attempted the second part by contradiction, supposing that g had no inverse. Then, in the very next line, he used the symbol g^{-1} , which he just assumed didn't exist, and made calculations with it.

The underlying misconception is that names always represent existing things. Writing $\cos x$ or g^{-1} seems to confer existence on and the right to manipulate the symbol. Perhaps this comes from secondary school algebra where x is referred to as "the unknown", that is, as something which exists and should be found.

If one asks an abstract algebra student to prove that the equation $ax=b$ has a solution in a group, he will often proceed as follows: $ax=b$, so $a^{-1}ax=a^{-1}b$, so $x=a^{-1}b$. He does not realize that by manipulating the equation he is tacitly assuming x exists. Of course, he should produce a group element, in this case $a^{-1}b$, which when substituted in the given equation yields a true statement.

One final simple-minded example illustrates the difficulties that can occur. Given $(x+3)(x+2) = (x+1)(x+4)$, one can conclude $6 = 4$, by supposing there is such an x .

M3. Apparent differences are real. This error occurs when things which have different names are taken to be different. The underlying misconception is there is a one to one correspondence between names and mathematical objects.

Although students realize that the same real number can be written in many different ways, for example, $1/2 = 2/4$ or $3 = 1+2$, often it does not occur to them that two different abstract expressions may represent the same thing. This happens even though they know two apparently different trigonometric expressions can be equal from having verified identities.

Example

Theorem: If a commutative group has an element of order 2 and an element of order 3 then it must have an element of order 6.

“Proof”: Let g be the element of order 3 and let h be the element of order 2. The $g^3 = e$ and $h^2 = e$ where e is the identity of the group.

Consider the subgroup generated by hg . Since $h^6g^6 = (h^2)^3(g^3)^2 = e^3e^2 = e$, this subgroup is $\{hg, h^2g^2, h^3g^3, h^4g^4, h^5g^5, h^6g^6\}$ which simplifies to $\{hg, g^2, h, g, hg^2, e\}$ using $g^3=e$ and $h^2=e$. So hg has order 6.

Comment

Something is missing here, namely an argument showing that the 6 symbols hg, g^2, h, g, hg^2, e represent 6 distinct elements. This can be shown, but one shouldn't assume the student can show it or is even aware that he must.

Example

Lagrange's Theorem: Let G be a group of order n . Let H be a subgroup of order m . Let r be the number of distinct right cosets of H in G . Then $n=rm$.

“Proof”: Let the distinct right cosets be H, Hg_1, \dots, Hg_{r-1} . These form a partition of G into equivalence classes. Let $H = \{h_1, \dots, h_m\}$. Then $Hg_s = \{h_1g_s, h_2g_s, \dots, h_mg_s\}$ has m elements in it. Thus G is partitioned into r classes, each with m elements, so G has $n=rm$ elements.

Comment

Again, the student has concluded there are m distinct elements by counting up m distinct symbols. He is tacitly assuming the symbols represent different elements.

M4. Using the converse of a theorem. This is a classic and extremely persistent reasoning error. The misconception consists in equating an implication and its converse. Even students who have been taught that there is a difference make this error, especially when a theorem is complicated.

The basis for this misconception seems to be the imprecision of everyday language. People often use the “if, then” construction when they mean “if and only if”. When someone says “if it rains, I won't go”, he often also means, but doesn't explicitly say, “and if it doesn't, I will”, which is logically equivalent to the converse. Mathematicians and textbook authors may reinforce this confusion when they state definitions using “if”, but actually mean “if and only if”.

Example

Theorem: Let G_1 and G_2 be two groups contained in a semigroup S such that $G_1 \cap G_2$ is nonempty. Then $e_1 = e_1e_2e_1$, where e_1 is the identity of G_1 and e_2 is the identity of G_2 .

“Proof”: By an argument with another type of error, the student concludes $e_1 = e_2$. Then $e_1e_2 = e_1e_1e_2 = e_1e_2e_1$, so $e_1e_2 = e_1e_2e_2 = e_1e_2e_1e_2$, so e_1e_2 is an idempotent. The identity of a semigroup is an idempotent. Therefore, $e_1e_2x = x$ for all x in S . Let $x = e_1$, $e_1e_2e_1 = e_1$.

Comment

The student has used the converse of the theorem that the identity is an idempotent in his penultimate line, not to mention the fact that the semigroup under consideration was not given as having an identity. As is often the case, this student has made multiple errors.

M5. Real numbers laws are universal.

Students who take abstract algebra at the junior level have very little idea that mathematics deals with objects other than geometric configurations and real and complex numbers. Thus, the examples of abstract concepts like group must be rather simple, and we tend to stick to real and complex numbers, matrices, and functions with which they have some familiarity. Even this limited collection of examples is rich enough for students to see that not everything behaves the way the more naive students expect. What they expect is really a misconception, namely, that the rules they know for dealing with real numbers are universal.

Example

Theorem: A semigroup can have at most one identity.

“Proof”: Suppose the semigroup has two identities, e and e' . Then $es=s$ and $e's=s$, so $es=e's$. Hence $e=e'$.

Comment

When it was pointed out to the student that the last step of his argument amounted to cancelling the s , and he was asked why that was permissible, he said, "That's not cancelling; that's logic."

Example

Theorem: Given a group G of finite order n . For each g in G , $g^n=e$ where e is the identity of the group.

Further, if H is a normal subgroup of order r , then for each g in G , g^m is in H , where $m=n/r$.

“Proof” of the second part: From Lagrange's Theorem, we know the number of distinct cosets of H in G is m . Now $g^m=g^{n/r}=(g^n)^{1/r}$. By the first part, $g^n=e$, so $g^m=e^{1/r}$. This is followed by an argument attempting to show $e^{1/r}$ is in H .

Comment

It apparently never occurred to the student to question whether the $1/r$ power of an arbitrary group element made sense; after all, there is nothing strange about taking the r -th root of a positive number.

Example

Theorem: A group G in which every element is of order 2 is commutative.

"Proof": Let g, h be elements of G . By hypothesis, $(gh)(gh)=1$ and $(hg)(hg)=1$, where 1 is the identity of G . Then $(gh)(\cancel{gh})=(hg)(hg)=h(\cancel{gh})g$ by the cancellation law for groups. So $gh=hg$.

Comment

Here the cancellation law for groups was not used correctly. The error could easily have resulted from thinking of the cancellation law in the real numbers, where cancelling this way is allowed, and even, routine.

M6. Conservation of relationships. This type of error occurs when students act as if doing the same thing to both sides of any relationship preserves the relationship. For example, given $h \neq k$, they will conclude $gh \neq gk$. They will do this in an abstract algebra proof, even though they are aware that in the real numbers, one must know $g \neq 0$. They may also be aware that in matrices, one must know g is nonsingular to conclude $gh \neq gk$.

This seems to be an instance of improper generalization from past experience. In secondary school algebra, one can say if $a=b$, then $ac=bc$, and if $a < b$ and $c > 0$, $ac < bc$. The misconception is that relationships can be preserved by operating on both sides in the same way.

As a variant of this, we note that occasionally students will act as if expressions are preserved, even when there aren't two sides. Students will simplify the polynomial $4x^2+2$ to $2x^2+1$.

M7. Element set interchanges. Students understand statements involving elements more easily than equivalent statements about sets. Our beginning abstract algebra students had difficulty making proofs when the notion of a subsemigroup T of a semigroup S was defined to be a nonempty subset T of S such that $TT \subseteq T$, the product of two sets having been defined previously. However, when a subsemigroup was defined to be a nonempty subset T of S such that for all a and b in T , ab is in T , students made better proofs.

This strategy of delaying potential confusion is helpful, however, set concepts must eventually be introduced. The usual definition of a normal subgroup H of a group G is given in terms of the equality of left and right cosets; that is, a subgroup H is normal if and only if $gH=Hg$, for all g in G . Students often convert this to $gh=hg$; they incorrectly think they can merely substitute h for H . Even when told explicitly that $gH=Hg$ means that given gh in Hg , there is an h' in H so that $hg=h'g$, they revert to writing $gh=hg$ in making proofs.

The misconception is that information about a set is interchangeable with information about a typical element of that set. So each time the set H appears, it is permissible to replace it by h .

OTHER ERRORS

E1. Overextended symbols. This error occurs when one symbol is used for two distinct things, often because the distinction was unobserved. Such errors can indicate an incomplete grasp of a mathematical structure, such as group, and first appear when the structure is used several times in the same setting.

Example

Theorem: Let G_1 and G_2 be two groups contained in a semigroup S such that $G_1 \cap G_2$ is nonempty. Then $e_1 = e_1 e_2 e_1$, where e_1 is the identity of G_1 and e_2 is the identity of G_2 .

"Proof: There is an element g in $G_1 \cap G_2$; $g = e_1 g$ and $g = e_2 g$, so $e_1 g = e_2 g$. Since g is a group element, g^{-1} exists, and $e_1 g g^{-1} = e_2 g g^{-1}$, $e_1 e_1 = e_2 e_2$, $e_1 = e_2$. Multiplying on the left and right by e_1 yields $e_1 = e_1 e_1 e_1 = e_1 e_2 e_1$.

Comment

The error consists in not distinguishing the two different kinds of inverses that exist. There is an inverse of g in G_1 , one might call it g_1^{-1} , and an inverse of g in G_2 , one might call it g_2^{-1} .

In the same way, students often fail to distinguish between equivalence classes coming from different equivalence relations.

Example

Theorem: Let G be a group, H a subgroup of G , and K a normal subgroup of G . Then $f: HK/K \rightarrow H/H \cap K$ defined by $f(hkK) = f(hK) = h(H \cap K)$ is an isomorphism.

"Proof" that f is one-to-one: Suppose $f(h_1K) = f(h_2K)$. Then $h_1(H \cap K) = h_2(H \cap K)$ so $[h_1] = [h_2]$ so $h_1K = h_2K$.

Comment

What is needed here is two different symbols for the two different equivalence classes, such as $[]_{H \cap K}$ and $[]_K$. One then sees that an additional argument is required,

E2. Weakening the theorem. This error occurs when what is used is stronger than the hypothesis or when what is proved is weaker than the conclusion. Often a student thinks it's clear he has proved the theorem. Adding to the hypothesis is a well-known technique of practicing mathematicians. If one cannot prove a conjecture as it stands, one can add to its hypothesis and attempt to prove a weaker result. However, students rarely realize they are proving a weaker result.

Errors of this kind occur when a student tacitly assumes a group is finite although nothing is stated about the order of the group, when a student assumes a semigroup has an identity although that is not given, or when a student assumes a group is cyclic.

Example

Theorem: Given a semigroup S with identity 1 , and left cancellation. The S has only one idempotent, 1 .

“Proof 1”: Let's suppose we have a group with 1 . Let e be an idempotent. Then $ee=e$. Multiplying by e^{-1} gives $e^{-1}ee=e^{-1}e=1$, but $e^{-1}ee=1e=e$, so $e=1$.

Comment

What the student has actually shown is that a group can have only one idempotent, the identity. The hypothesis was strengthened. Another student weakened the conclusion as follows.

Example

“Proof 2”: Suppose there are two identities s and t , then $as=a$ and $at=a$, so by left cancellation $s=t$. This unique identity is an idempotent as $a \cdot 1=a$ for all a in S , so in particular $1 \cdot 1=1$.

Comment

This student has shown the weaker result that a semigroup identity is an idempotent, rather than that a semigroup identity is the only idempotent, given left cancellation.

E3. Notational inflexibility. This error arises from an inability to adapt notation from one context to another. As is customary in abstract algebra, we use multiplicative notation in all definitions about noncommutative groups and semigroups. However, additive examples are considered in class, and commutative groups, especially in the case of rings, are written in additive notation.

For example, the cyclic subgroup generated by the element h in a group G is defined to be $H=\{h^n:n \text{ is an integer}\}$. On being asked to find the cyclic subgroup of the additive reals generated by 1 , students occasionally answer $H=\{1^n\}=\{1\}$, which is incorrect. In additive notation, the cyclic subgroup generated by h is $\{nh:n \text{ is an integer}\}$.

A similar error occurred when a student was asked to find the kernel of a homomorphism. The group was the nonzero reals under multiplication and the function was defined by $f(x)=|x|$. One student wrote: $K(f)=\{x : f(x)=0, x \in \mathbf{R}-\{0\}\}=\{x : |x|=0, x \in \mathbf{R}-\{0\}\}=\emptyset$. He used the additive identity, 0 , in place of the multiplicative identity, 1 . The student should have suspected his answer was incorrect, as a previous theorem stated that the kernel is always a subgroup, and hence, nonempty. Students often fail to notice if their answers are reasonable, that is, in agreement with previously obtained results.

E4. Misuse of theorems, other than the converse. In applying a theorem, an error of this type arises from misunderstanding or partly neglecting the hypothesis or misinterpreting the conclusion. A student's inability to read precisely⁵ combined with a desire to finish the proof

quickly may result in such mistakes. This error has similarities with E2. In this case, the theorem being used is misunderstood, and in E2, the theorem being proved is misunderstood.

One student stated that if a group has subgroups of orders r and s , then it must have a subgroup of orders, which is false. The student was probably thinking of the following: If a group has cyclic subgroups of orders r and s , where r and s are relatively prime, then it has a subgroup of order rs .

Another student stated that a subsemigroup H of a group must necessarily be a subgroup, forgetting the additional requirement that H must contain the identity and the inverse of each element. A counter example was well within the student's mathematical experience; the open unit interval is a subsemigroup, but not a subgroup, of the positive reals under multiplication. Once again, the student has not checked to see if what he claims is reasonable.

E5. Circularity. This error consists in reasoning from a statement to itself. Often the reasoning is from one version of the conclusion to another, already known to be equivalent to the first.

Example

Theorem: Let G and K be groups. If $f:G \rightarrow K$ is an onto homomorphism and H is a normal subgroup of G , then $f(H)$ is a normal subgroup of K .

"Proof": It has already been proved that $f(H)$ is a subgroup of K , so it remains to show that $kf(H)k^{-1} \subseteq f(H)$ for all k in K . Now $kf(H)k^{-1} \subseteq f(H)$ implies $kf(H) \subseteq f(H)k$. We also need to show $f(H)k \subseteq kf(H)$. This implies $k^{-1}f(H)k \subseteq k^{-1}kf(H) = f(H)$, which in turn implies $kk^{-1}f(H)k = f(H)k \subseteq kf(H)$. Thus, $f(H)k = kf(H)$, so $f(H)$ is normal in K .

Comment

The student began with the conclusion. In the first part of the argument, he went from one equivalent version of normal to half of another. He then took the remaining half and got back where he started. Each instance is an example of circularity. The student doesn't understand what constitutes a proof but realizes arguments progress from one piece of information to another.

E6. The locally unintelligible proof. In this error neither the proof as a whole nor most individual sentences can be understood. The format is acceptable, the words "theorem" and "proof" occur together with many symbols, and most sentences are syntactically correct, however, the assertions are incomprehensible or incorrect. Such "proofs" may be first approximations at imitating textbook and classroom proofs⁶.

Example

Theorem: A commutative group which has an element of order 2 and an element of order 3 must have an element of order 6.

“Proof”: Let G be a commutative group, $H \subseteq G$. Let $g_1 \in G$, $g_2 \in H$, $g_3 \in G$. According to Theorem 55 of the notes (on using a subgroup H to define the right coset equivalence relation on G), $H g_1 \subseteq G$. Then $H g_1 = g' \in G$. Let a be the order of g' , which is the number of distinct right cosets of H in G . Let b be the order of g_2 , c be the order of g_3 . Since $g_3 \in G$, c is the order of G . We want to show $c=ab$. According to Lagrange's Theorem, the order of a group is equal to the order of a subgroup times the number of right cosets of that subgroup in the group. Therefore, $c=ab$. In this theorem $a=2$, $b=3$, $c=2 \cdot 3=6$. This gives an element of order 6.

Comment

On the surface this appears to be a proof. It starts and stops in an expectable way and quotes Lagrange's Theorem correctly. It is also syntactically correct, except for one small place. However, it is impossible to find any basic underlying idea which the student might have started with or to follow the individual sentences⁷.

The student may have selected Lagrange's Theorem almost arbitrarily⁸ and tried to develop it into a proof.

E7. Substituting with abandon. This error consists in obtaining one statement from another using an unjustifiable substitution. One fixed element is replaced by another unequal fixed element. Of course, it is permissible to substitute for a universally quantified variable; and perhaps, this error results from confusing the two situations.

This error often occurs when a student attempts to prove a theorem which begins “For all s in S , ...”. The standard way to start the proof is to write “Let $s \in S$.” With these words, s becomes a fixed element, and one is no longer free to substitute for s . Occasionally, to emphasize this point, one writes, “Let s be an arbitrary, but fixed element of S .”

Examples

Theorem (Cancellation Law): Let G be a group. Let g, h, k be elements of G . If $gh=gk$, then $h=k$.

“Proof”: Let e be the identity of G . Substitute e for g in $gh=gk$. The $eh=ek$, so $h=k$.

Theorem: Let G be a group with identity e . If g, h, k are elements of G so that $gh=e=hg$ and $gk=e=kg$, then $h=k$.

“Proof”: We have $gh=e=hg$. Since $k \in G$, substitute h for k to get $gk=e=kg$.

Comment

In the first theorem, we have this error in its purest form; g was arbitrary, but fixed from the beginning, but the student substituted e for g , believing g to be a universally quantified variable. In the second theorem, both h and k are arbitrary, but fixed elements from the start: the only way to substitute one for the other is to know they are equal.

Example

Theorem: If G is a group of order n , then $g^n=e$ for all g in G , where e is the identity of G .

“Proof”: Since G is finite, and one wants to show $g^n=e$ for all g in G , one can choose n to be zero or a value which gives the identity element.

Comment

The student has failed to realize n is given as an arbitrary, but fixed positive integer in the statement of the theorem. He is not free to choose it, rather he must prove something about it.

E8. Ignoring and extending quantifiers. This error results from failing to notice restrictions on variables. Often a variable is thought to be universally quantified when it isn't.

Example

Theorem: Given a group G and an element a in G . Then $H=\{g \in G:ag=ga\}$ is a subgroup of G .

Comment

One student concluded incorrectly that H was commutative because $ag=ga$, failing to notice that a is not a universally quantified variable.

Another student proved H was closed under multiplication as follows. Let $g, h \in H$. Then $ag=ga$ and $ah=ha$. Then $a(gh)=(ag)h^*=h(ag)=(ha)g^*=g(ha)=(gh)a$. In the equalities marked with a “*”, the student assumed that since h and g commute with a , they also commute with ag and ha , respectively. He failed to notice that a is fixed, and not universally quantified, in the definition of H .

E9. Holes. This type of error consists in claiming that a statement follows immediately from previously established results when in reality a considerable argument is required.

Example

Theorem: Given a group G of order n and a normal subgroup H of order r , then for all g in G $g^m \in H$ where $m=n/r$.

“Proof”: First it was shown that $g^n=e$ for all g in G , where e is the identity of G . Then $g^n=g^{mr}=(g^m)^r=e$. Letting $k=g^m$, one gets $k^r=e \in H$. *** So $k \in H$.

Comment

The hardest part of the argument goes where we have inserted stars. This student grasps the concept of proof and has a reasonable overall approach, but has difficulty distinguishing between statements which follow immediately and those requiring justification.

E10. Using information out of context. In this type of error information from one argument is improperly used in another, often because identical symbols appear in both. This error is most likely to occur when proofs are organized into independent sections, for example, in theorems involving case analysis, set equality, or equivalence of statements. In such situations, a student may unjustifiably transfer information from one section to another. The next example is rather unusual in that the error involves two theorems, rather than two independent sections of one proof.

Example

Theorem (Cancellation Law): Let G be a group. Let $g, h, k \in G$. If $gh=gk$, then $h=k$.

“Proof”: Since G is a group, it has an identity e and $gh=e=hg$. (A previous theorem with the letters g and h is invoked here.) Since $k \in G$, we can substitute k for h and get $gk=e=kg$. Then, by the previous theorem on uniqueness of group inverses, $h=k$.

Comment

Not only has the student used a piece of a previous theorem out of context in the first line; he has also made a substitution error (E7). This serves to illustrate that several of the reasoning errors described in this paper can occur within a single proof.

ANOTHER PERSPECTIVE

The reasoning errors analyzed above have been classified according to whether or not they are misconception based. It is also possible to classify reasoning errors according to their logical characteristics, that is, according to whether they arise from difficulties in generalization, use of theorems, notation and symbols, nature of proof, or quantification. We summarize these two classifications in the following table and note their independence.

ERRORS	Misconceptions Based	Other
Generalization	M5, M6	
Use of Theorems	M4	E4
Notation and Symbols	M2, M3, M7	E1, E3, E10
Nature of Proofs	M1	E2, E5, E6, E9
Quantification		E7, E8

GENERAL COMMENTS

The analysis and classification of reasoning errors presented in this paper suggests a number of questions:

1. How complete is the error list? It would be useful to have a list sufficiently complete that each incorrect student proof contains at least one of the errors.

2. How does each of these reasoning errors arise and how could it be prevented?
3. Is the making of one type of reasoning error correlated with making others? Perhaps students who make a particular type of error always make another type.
4. Which types of reasoning error occur most frequently?
5. Do certain types of reasoning errors occur more in one course than another, for example, in algebra as compared with topology?
6. Are any of these reasoning errors correlated with particular sections of students' earlier coursework?

If lower division mathematics courses were to ignore the static view and include significant instruction on creating and validating algorithms, it is possible that reasoning would be improved, as well as applications extended. Evidence concerning this point would be useful.

Finally, we note that there is remarkably little correlation between the reasoning errors we have observed and classified and the topics emphasized in an introductory logic course or even in one of the newer courses as on transitions to advanced mathematics⁹.

END NOTES

- ¹ A detailed description of this method is given in Selden and Selden (1978, p. 69-71).
- ² This agrees with the general research on misconceptions (Novak, 1983, p.2).
- ³ An assumption shared by Confrey. (1983, p.30, p.25).
- ⁴ Occasionally extraneous material was edited out.
- ⁵ Lack of precision and accuracy have been observed in first year university students' attempts to solve physics problems (Mehl and Volmink, 1983, p.228). Lin (1983, p.202) has suggested that beginning physics students are unaccustomed to the necessary precision of expression.
- ⁶ Students indicate they solve mathematics problems by imitating textbook examples (Confrey, 1983, p.23).
- ⁷ The authors have made a painstaking attempt to find some basic underlying idea, plus a line-by-line analysis of this "proof" (Selden and Selden, 1978, p.78-9).
- ⁸ In making this "proof", the student appears to have acted in an unsystematic and somewhat impulsive way (Mehl and Volmink, 1983, p.228). He does not take what computer scientists would call a top-down approach.
- ⁹ We recently taught such a course several times and found the treatment in a typical text (Smith, Eggen, and St. Andre, 1986) only superficially related to these reasoning errors.

REFERENCES

- Confrey, Jere. "Implications for Teaching from the Research on Misconceptions." Proceedings of the International Seminar Misconceptions in Science and Mathematics, 1983, 21-31.
- Forbes, Douglas Robert. The Texas System: R. L. Moore's Original Edition, Ph.D. dissertation, University of Wisconsin. Ann Arbor: University Microfilms International, 1980.
- Lin, Herb. "A 'Cultural' Look at Physics Students and Physics Classrooms. An Example of Anthropological Work in Science Education." Proceedings of the International Seminar Misconceptions in Science and Mathematics, 1983, 194-213.
- Mehl, Merlyn C. and John D. Volmink, "Influence of Cognitive Instruction on Misconceptions in Physics and Mathematics." Proceedings of the International Seminar Misconceptions in Science and Mathematics, 1983, 228-232.
- Novak, Joseph D. "Overview of the Seminar." Proceedings of the International Seminar Misconceptions in Science and Mathematics, 1983, 1-4.
- Selden, Annie A. "Teaching the Doing of Mathematics." Abacus, September 1981, 41-53.
- Selden, Annie A. and John Selden. "Errors Students Make in 'Mathematical Reasoning.'" Bosphorous University Journal, Vol. 6, 1978, 67-87.
- Smith, Douglas, Maurice Eggen and Richard St. Andre. A Transition to Advanced Mathematics. Monterey, California: Brooks/Cole, 1986.