CBMS Issues in Mathematics Education
Volume 6, 1996

Dihedral Groups: A Tale of Two Interpretations

RINA ZAZKIS AND ED DUBINSKY

This note is concerned with mathematical objects and their naming schemes, that is, means of assigning to an object some algebraic quantity. We start with an empirically-based examination of a phenomenon observed in undergraduate Abstract Algebra courses and then we continue with the analysis of related phenomena that are found in the literature.

Our specific interest is with mathematical and psychological aspects of constructing a group on the set of symmetries of a regular polygon of n sides and on the set of permutations of n objects. The points we wish to make are amply illustrated in a specific example and so we will mainly consider the case $n = 4$ with the corresponding set D_4 of 8 symmetries, and the symmetric group S_4 of 24 permutations. As is well known, the elements of D_4 correspond to elements of S_4 and so it would seem that there are two ways of constructing (representing, if you like) a group structure on D_4: one as motions of the square with composition (of transformations) as the operation and the other as a set of permutations with multiplication (of permutations) as the operation. They should be isomorphic.

One can see the first construction as a geometric or visual way of thinking about a certain group and the second as a symbolic or analytic way of thinking of the same group. We were (and are) interested in how students used these two modes of thinking when they were trying to understand the construction of these groups. We found that students tended to make a certain error in moving back and forth between the two representations and, in analyzing this error we came to the view that the situation was a little more complicated (both mathematically and psychologically) than one might think. Indeed, we even came to question our belief that the group D_4 can be constructed in a completely visual manner.

We think that these questions are important because they are related to the pedagogical issue of students' visual and analytic approaches to making sense out of mathematical situations. The observations we are making in this note form a small piece of a larger study of general visual/analytic issues [11].

In the following pages, we explain how we were motivated to think through the issues discussed in this paper, state the main problem, propose an explanation of the students' difficulty, suggest an analysis that could form the basis for eliminating the difficulty, and describe how we think students may be thinking about the relationship between dihedral groups and groups of symmetries.

Having done that, we attempt to clarify certain complex relationships between mathematical objects and their names. We analyze several related problems (of double interpretation) described in literature, such as Birkhoff and Mac Lane's alibi/alias dichotomy [3, 9]. Our main purpose in such an analysis is to reveal complexities that are often not acknowledged by mathematics experts. We explain in what way our problem is related to other problems from the literature and in what way it differs. We also suggest possible avenues for future research that will further explore the mathematical and psychological phenomena we have encountered.

Our Motivation

Our considerations were motivated by ten individual interviews with undergraduate mathematics majors in the middle of a first course in Abstract Algebra. In one of the tasks presented to the students in the interview, they were asked to list the elements of D_4, and then to calculate a product of two specific elements of D_4. The interviewer made no attempt to suggest a particular representation of D_4 and each student made one of two choices of how to do it: geometrically, using a physical model of a square, or analytically, multiplying permutations. The interviewer then asked the student for another way to do it and to see that both methods gave the same answer. As it turned out, eight of the ten students did not get the same answer and felt they had made an error (they each did the same thing) and this note resulted from our attempts to understand what they did and why it seems wrong.

Let us introduce the notation used in the Abstract Algebra course which these students were taking. The elements of D_4 were familiar to them as four clockwise rotations, denoted R_0, R_{90}, R_{180}, R_{270}, and four reflections across horizontal, vertical and diagonal axes, denoted H, V, D_L (left diagonal) and D_R (right diagonal). (See Figure 1.) The permutations of S_4 were denoted as sequences of four digits, that represent the "first floor" of a standard "double decker" notation. For example, [4123] represented the permutation "$1 \to 4$, $2 \to 1$, $3 \to 2$, $4 \to 3$," usually listed as $\binom{1234}{4123}$.

The following excerpt from an interview with Peter[1] was typical of the student performance of the task of computing $R_{90} * V$ and their reactions upon its completion. Note that the student applies the convention of starting with the element on the right in computing a product, that is, $R_{90} * V$ is interpreted as "V followed by R_{90}." (See Figure 2.)

[1] The names of the interviewees have been changed to protect their identities.

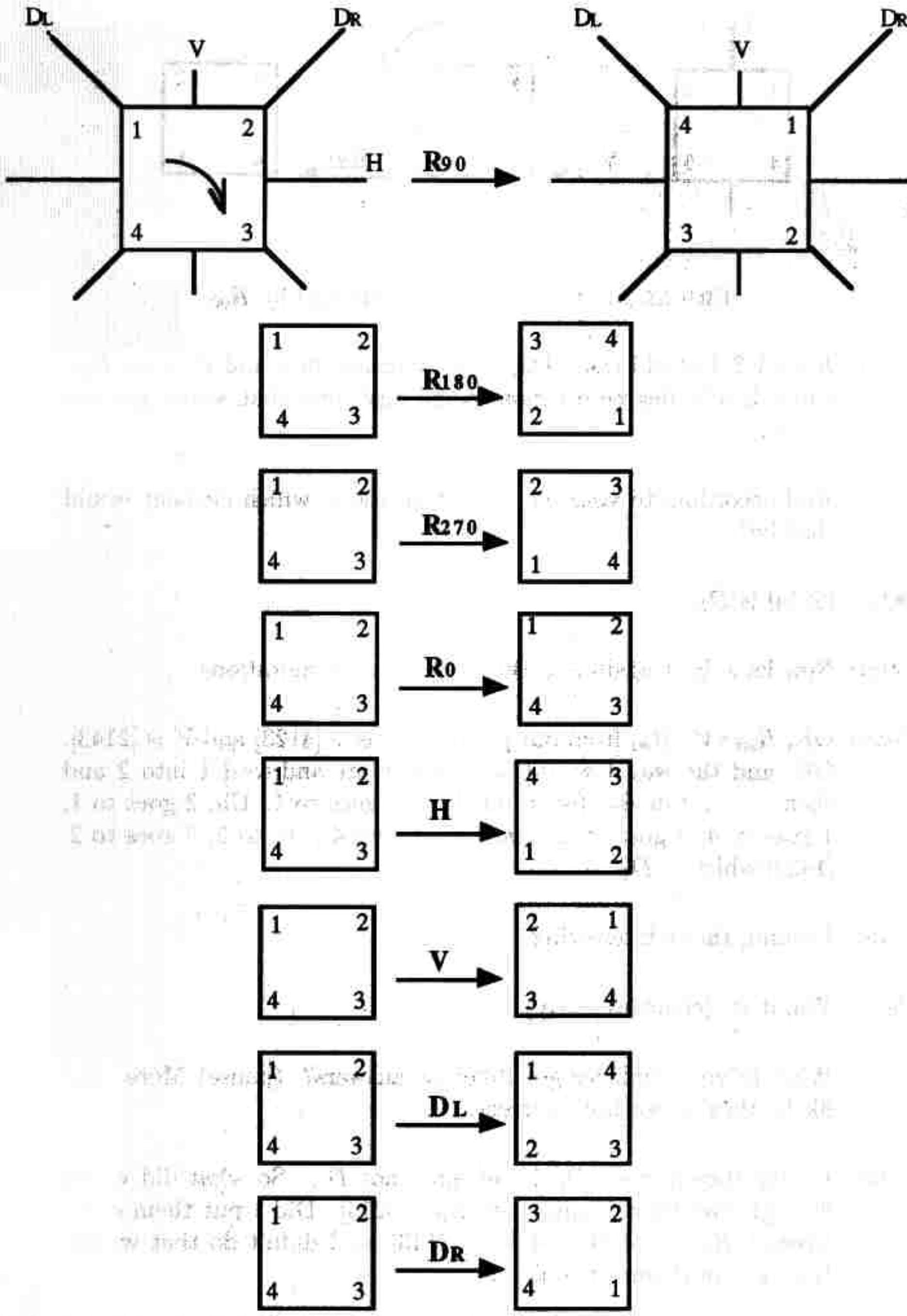

FIGURE 1. Elements of D_4—Global interpretation.

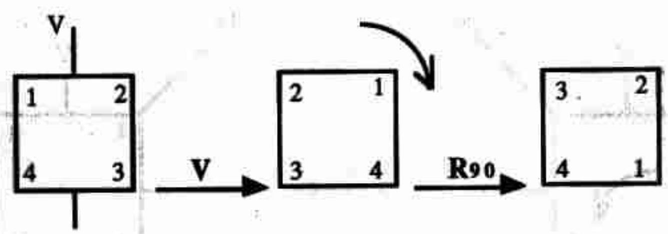

FIGURE 2. Carrying out V followed by R_{90}.

Peter: $R_{90} * V$? I would take this, do the vertical flip, and then do R_{90}, which is a 90 degree rotation to the East, and that would give me [3214].

Int: And according to your element table there, which element would that be?

Peter: [3214] is D_R.

Int: Now let's do it again and this time with permutations.

Peter: OK, $R_{90} * V$. R_{90} from our permutations is [4123] and V is [2143]. OK, and the way I would do it was start and feed 1 into 2 and then . . . , 2 in R_{90} feeds into 1, so 1 goes to 1. Uh, 2 goes to 1, 1 goes to 4; 3 goes to 4, 4 goes to 3; and 4 goes to 3, 3 goes to 2. [1432] which is D_L.

Int: Ummm, that's interesting.

Peter: Yes, it is. (chuckles softly).

Int: Why do you think we got different answers? (pause) More than likely, they're not both correct.

Peter: I'd say they're not. D_L is certainly not D_R. So what did we do wrong? We did a V and then R_{90}, [3214]. Did I put them down wrong? R_{90} is [4123] and V is [2143], so I didn't do that wrong. (pause) I'm stumped (chuckle).

Peter appears to perform the manipulations and calculations correctly, but using two different methods, he gets two different answers. He sees this and, in listening to the tape we took his reaction to be a kind of nervous laughter indicating some discomfort with the situation or what is called in learning theory, disequilibration [4]. Since the calculations are performed correctly, the difficulty appears to lie in the two representations and their connections.

The Problem

The conventional view is that the dihedral groups D_n are established from geometrical considerations of rigid motions of a regular polygon of n sides and the symmetric groups S_n are established from considerations of permutations of n objects. Then, there is an isomorphism of D_n into S_n. Textbooks on group theory, however, usually don't spell out these isomorphisms to the learner explicitly. The books prefer to leave this as an exercise to the reader (e.g., Armstrong [2, p. 36] or to list the elements of D_3 or D_4 as permutations, without making explicit correspondence to the transformations they represent (e.g., Fraleigh [6, p. 70]. As demonstrated in the previous section, learners who try to use an explicit correspondence, can be surprised by unexpected results. This raises pedagogical and attitudinal issues. Will students try to resolve these contradictions or will they give up and eliminate their disequilibration by making unnecessary conclusions regarding the nature of mathematical anomalies? What pedagogical strategies are promising for getting more of the former and less of the latter?

In trying to analyze the mathematical situation from a psychological point of view, we tried to understand how one might think about the connection between D_4 and S_4. Following the traditional point of view we began with introspection into our own way of constructing the group D_4. The idea was to do this first and then think about the connection with S_4.

The eight symmetries of a square are clear. They are dynamic processes which with a little mental activity that is non-trivial, but reasonable for the students we have in mind, can be thought of as eight individual objects and it is reasonable to name them as in Figure 1. Do these eight objects form a group? Previous research has reported [5] that beginners in Group Theory have a tendency to ignore the operation and relate only to the set of elements of the group. However, even when the need for operation is recognized, applying the binary operation of composition to the symmetries of a square is not as obvious as it may seem. To avoid the undue influence of permutations we tried to do this without labeling the vertices. It is not so easy, for example, for the composition $R_{270} * H$. That got us to wonder about the "easy" cases, say $R_{270} * R_{180}$. Obviously the answer is R_{90}. But is it?

Take a square which is completely blank and looks exactly the same on both sides. Rotate it by 180 degrees and then by 270 degrees. In what sense can one say that you have rotated the square 90 degrees? You can't tell by the result, because the square looks exactly like it would look if you had not rotated it at all. You can't tell by watching the action because, in reality the composition of the two acts of rotation is indeed an act of rotation—of 450 degrees, not of 90 degrees! A combination of more thought and choice of convention has to go into a decision to call it rotation by 90 degrees.

In this sense, one can suggest that specifying the binary operation on a set of symmetries is not sufficient to define a group, since from a purely visual point of view, unless some analytic structure and representational conventions are introduced (labeling of vertices, addition mod 360, etc.) the eight rigid motions of a square are not closed under composition. There is a real difficulty in specifying the net result of a sequence of transformations without some form of labeling. It seems that labeling vertices implies an embedding of the set D_4 in the set S_4 and carrying the group structure to D_4 from (a subset of) S_4. Our conclusion from this is that the connection between D_n and S_n, if considered carefully, must be established on the level of sets, not as groups. One constructs a one-to-one map of D_n into S_n (using a considerable amount of visualization) and then the operation on D_n is constructed without ambiguity by pulling back the operation on S_n. It may seem that paying explicit attention to the construction of this injection could help students avoid pitfalls related to this connection.

Unfortunately, this is not quite enough. There is more than one reasonable way to embed the set D_n in the set S_n.

Two Embeddings

There are (at least) two possible ways to correspond a symmetry of a (labeled) square to a permutation. One way is to look at the vertices of the square as objects being moved. In performing R_{90}, vertex 1 is moved to vertex 2, vertex 2 is moved to vertex 3, 3 is moved to 4 and 4 is moved to 1. Therefore we may represent R_{90} as [2341]. We will refer to this correspondence as "object interpretation."

Another way of interpreting the situation is to look at the environment of the square and think not about vertices moving but about positions and which vertices they contain. In this interpretation, after R_{90}, position 1 contains vertex 4, position 2 contains vertex 1, position 3 contains 2, and 4 contains 1. This is represented by a permutation [4123].

We will refer to this correspondence as "position interpretation."

Clearly, these two correspondences are inverses in the sense that, given an element of D_4, the permutation to which it corresponds under the object interpretation is the inverse of the permutation to which it corresponds under the position interpretation. That is, there is an underlying anti-automorphism of the group S_4 and hence there are two, essentially different, ways in which groups can be constructed on the set D_4 by using permutations. The difference, of course, is small in the case of D_4 since all but two of its elements are idempotents. For example, the fact that R_{180} has order two assures that when vertex 1 moves to vertex 3, position 1 contains vertex 3; when vertex 2 moves to vertex 4, position2 contains vertex 4; and so on. The same situation exists with the reflections. We can predict from this analysis that the error which our students make will never appear unless we ask them to form a product involving R_{90} or R_{270} as either a factor or a result.

We can use our analysis to give a plausible explanation of the errors made by our students. Peter made a square, labeled the sides and performed the manipulations as indicated in Figure 2. There is no question about what to do with the first motion—the square is flipped across its vertical axis. The problem arises with the second motion because the square is no longer in its original position. In order to decide whether applying R_{90} means that the square should be rotated clockwise or counter-clockwise, it is necessary to describe the correspondence more carefully and we will do this below. We will see that a clockwise rotation is consistent with the object interpretation. The final position of the square can only be interpreted as the motion D_R and so, Peter's first response is correct.

Why did Peter get a different answer when he did it by permutation product? Notice that in choosing [4123] for R_{90} he says that "1 goes to 4," that is, position 1 contains vertex 4 which is the position interpretation. Indeed, as can be seen by comparing the excerpt with Figure 3, Peter used the position interpretation in going from symmetries to permutations, and therefore, he was in effect calculating this time with group elements that are the inverses of the V and R_{90} used in his first calculation.

Element of D4	Object interpretation	Position interpretation
R_0	[1234]	[1234]
R_{90}	**[2341]**	**[4123]**
R_{180}	[3412]	[3412]
R_{270}	**[4123]**	**[2341]**
H	[4321]	[4321]
V	[2143]	[2143]
D_L	[1432]	[1432]
D_R	[3214]	[3214]

FIGURE 3. Corresponding transformations to permutations.

We can see explicit descriptions of how students used the position interpretation in corresponding symmetries to permutations. For example, Stacey explains how she corresponded a permutation to R_{90}.

Stacey: [1234] was my beginning position. And then I rotated it once , 90 degrees (clockwise), yes. OK, and then again by reading the top left all the way down through bottom left, my first rotation of 90 degrees was [4123].

John mentions "position" in his answer.

Int: How would you write out R_{90} as a permutation?

John: OK, um, R_{90} would map the corners [1234] to the new positions [4123].

And Jeff explicitly acknowledges the choice of "positional" interpretation in his explanation.

Int: Alright, you made a rotation 90 degrees in a clockwise fashion. Alright, now how did you get the permutation to correspond to the set?

Jeff: OK, the first element in this permutation would be where the 1 used to be. And the second one would be where the 2 used to be, the third one where the 3 used to be and the fourth one where the 4 used to be. So we got [4123] from there.

These excerpts suggest that for many students the position interpretation is natural for setting up the correspondence. However, we submit that it is not so easy to use this interpretation in actually manipulating a square. For example, consider the middle square in Figure 2. We will argue below that to be consistent with the position interpretation the correct implementation here of R_{90} is what appears to be a counterclockwise rotation. At the very least, the reader might agree at this point that for the middle square in Figure 2, the position interpretation does not tell us how to interpret R_{90}. The reason is that this interpretation involves "vertices repositioned" from an initial state. But after performing one motion, the initial positions have changed. Thus we suggest that the interpretations we have given are reasonable for deciding which permutations to assign to a single motion, but can be confusing to use in a context in which more than one motion is being performed. We need a more powerful model.

Local And Global Transformation—A Possible Solution

We would like to suggest such a model. The position interpretation corresponds to a different perspective on symmetries that we will call "local." According to this perspective, a symmetry is a "local" transformation of an object in the sense that the object is moved according to axes and directions that are transformed along with the object. The term "local," as well as the idea of local representation, is borrowed from Turtle Geometry [1, 8] where transformations are functions of the position of the object transformed.

More precisely, in the case of D_4, the idea is to construct the four axes and directions on the square and move them along with the square as shown in Figure 4.

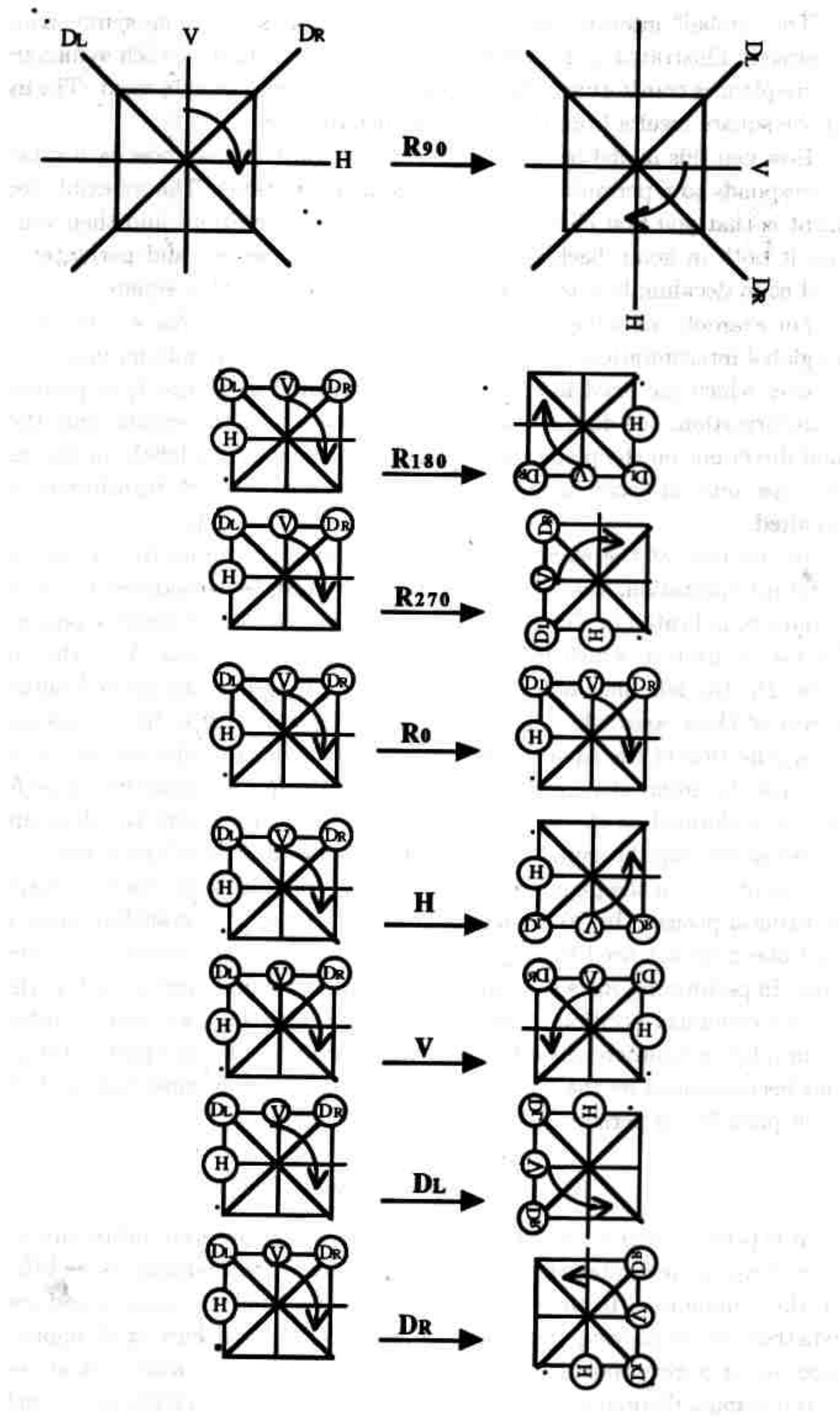

FIGURE 4. Elements of D_4—Local interpretation.

The "global" interpretation corresponds to a perspective on symmetries that is already illustrated in Figure 1. Here we consider that in each symmetry the entire plane is transformed, but the axes and directions remain fixed. The motion of the square results from this global transformation.

How can this model be used in practice? Figure 3 shows how each symmetry corresponds to a permutation in the two interpretations. The inflexible requirement is that you first choose one of the two interpretations and then you must use it both in going back and forth between symmetries and permutations as well as in deciding how to implement a composition with a square.

For example, we have seen in Figure 2 how to perform $R_{90} * V$ in the object or global interpretation. One performs the motions relative to an immutable set of axes which are fixed in the plane. As indicated in Figure 1, to perform the transformation, one does not need any information on the square, only the axes and directions on the plane that contains the square. The labels on the vertices are used only in deciding relative to these fixed axes which transformation has resulted.

In the case of the same operation performed according to the position or local interpretation, the axes and directions must be considered to be on the square as indicated in Figure 4. We choose the "canonical starting position" to be the position in which the names of the axes make sense (V is the vertical axis, D_L the left diagonal, etc.). The transformations are defined entirely in terms of these axes which are fixed on the square. Thus, the transformation V is reflection of the square about the V-axis and a rotation moves the V-axis towards the adjacent D_R-axis. In performing $V * R_{90}$ as shown in Figure 5, first R_{90} is performed on the canonical position. Then, if we want to follow up with V, we should flip the square according to the new position of the V-axes, that is, "horizontally" in the plane, since the original "vertical" axis was transformed to horizontal position by 90 degrees rotation. The resulting transformation is D_R and one does not need to keep track of the labels on the vertices to determine this. In performing $R_{90} * V$ as shown in Figure 6, we first start with V performed on the canonical position. Then, to follow up with R_{90}, we have to rotate the square *left* or counterclockwise, since the direction, which is a part of the square, has been changed by the flip. The result is D_L, which is consistent with Peter's multiplication of permutations.

Connecting The Two Models

It is perhaps reasonable to think about the object-position dichotomy as more a property of permutations and the global-local characterizations as belonging to the symmetries. In the former case, one must label the vertices and consider whether one is moving these four objects to new positions or changing what appears at a given position. For the latter case, it is the axes that are labeled and one must distinguish between moving the plane which happens to contain a square versus moving a square which happens to sit in a plane. In either case,

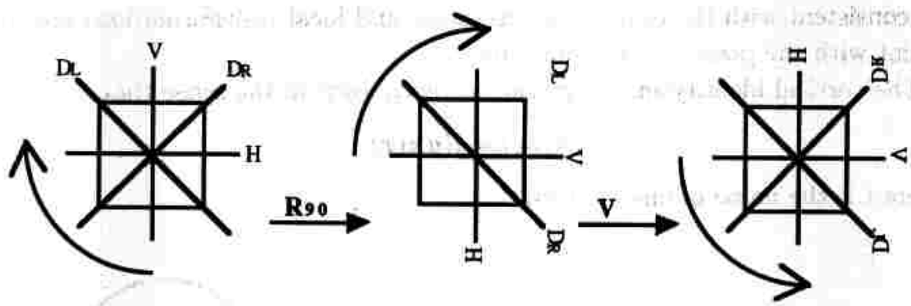

FIGURE 5. Carrying out R_{90} followed by V—Local interpretation.

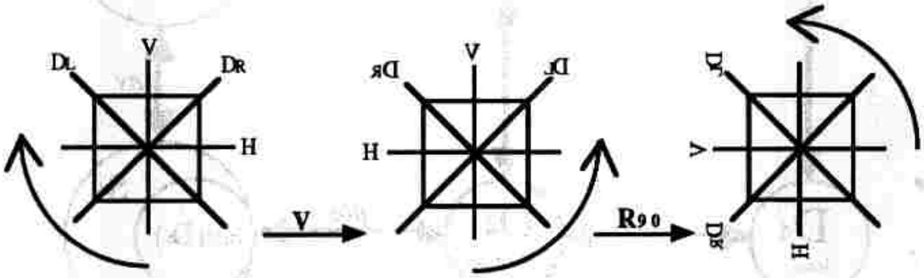

FIGURE 6. Carrying out V followed by R_{90}—Local interpretation.

some analytic structure must be added to the square. This gives further support to our earlier observation that the analytic structure of labeling vertices or axes is necessary to formally define the group of symmetries of a regular polygon with the operation of composition.

As we have indicated above, the important thing is to use a single interpretation throughout any particular discussion. This interpretation must remain invariant as one moves back and forth between permutations and symmetries or switches from listing elements to performing the group operation. The overall situation is described in Figure 7. The maps *obj* and *pos* are bijections of D_4 onto subgroups of S_4 , defined respectively by columns 1,2 of Figure 3 for *obj* and columns 1,3 of Figure 3 for *pos*. The map *inv* sends an element of S_4 into its inverse. All of the other maps in the diagram are the identity on D_4 . The entire diagram commutes. For the identity maps this is clear and for the right hand portion it follows from the relation,

$$pos(x) = inv(obj(x)) \qquad \text{for all } x \text{ in } D_4$$

The horizontal identity maps are automorphisms of D_4 with the indicated operations because the operation obtained from the object (respectively, position) interpretation is the same as the operation obtained from using global (respectively, local) axes. Another way of expressing this is that global transformations are consistent with the object interpretation and local transformations are consistent with the position interpretation.

The vertical identity maps are anti-isomorphisms in the sense that

$$i(xy) = i(y)i(x)$$

where i is the name of one of them.

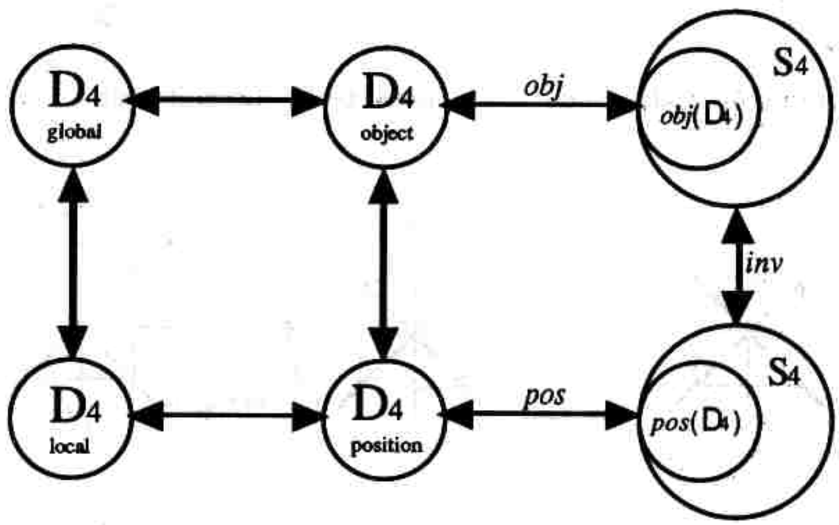

FIGURE 7. Interpretations of binary operations on D_4.

Following through the interview segment with Peter, we suggest that he used global interpretations for performing transformations with a physical square in his first calculation, obtaining the answer D_R. In his second calculation he used the position interpretation to correspond permutations to transformations, obtaining the answer D_L. Put another way, in his first calculation he worked at the upper level of the diagram and in his second calculation he worked on the lower level. The two structures are connected by maps which do not always preserve the operations and this explains why he obtained two different answers.

Of course, in the case of commuting elements the inconsistent choice of global symmetries and position interpretation doesn't lead to an error and disequilibration but leaves "an illusion of isomorphism."

Dealing With Disequilibration

As we indicated above, the error and the inconsistent interpretations which caused it were seen explicitly in eight of the ten students that were interviewed. Actually, neither of the other interviewees did the problem correctly. One student was confused and failed to carry out the composition at all. The tenth student reversed the order of elements when carrying out the transformations, and therefore, didn't "have a problem."

Following are some indications of the eight students' attempts to reequilibrate. The first reaction usually was to look for a computational error. When such was not found, the students concluded that the order of applying transformations should have been different from the order of multiplying permutations. Jeff explains:

> Jeff: I think, using the actual square we didn't do them in the right correct order. For some of them it didn't really matter . . .

John explains:

> John: When you do it in permutations, you have to work right to left, but if you're actually going to manipulate the square, you perform the operations left to right.

The idea of "starting on the right" when composing functions may appear strange for novices, and makes sense only with an appropriate interpretation of function composition as $fg(x) = f(g(x))$. A decision to alternate the order, once starting on the left and once on the right, indeed helps with temporary reequilibration. On the other hand it reveals a poor understanding of functions. Thinking of transformations and of permutations as functions would have eliminated the possibility of such a solution.

Resolving disturbing inconsistencies by accepting a "mid-stream order switch" is an approach to which even experienced mathematicians can be reduced. For example, Halmos [7, p. 66] in discussing matrices that correspond to linear transformations acknowledged the "unpleasant phenomenon of indices turning around." Halmos also wrote that "it is a perversity not of the author, but of nature" that makes us use an equation that "works" instead of the "more usual equation." We will return to this example in the following section.

Let us consider briefly the robustness of the students' interpretation. On several occasions, the interviewer pointed out to the student that the rotation R_{90} should have corresponded to [2341], rather than [4123]. In trying to make sense of this remark, the typical response was "Oh, that should have been left/counterclockwise rotation. I thought it was right/clockwise rotation. That's why it didn't work." An excerpt from a student's protocol is given below:

> Int: Let's talk about R_{90}. Why don't you tell me how you came up with [4123] as a permutation.

> Mark: Well R_{90}, we got [4123] because we had a square labeled [1234] starting at the upper left-hand corner. And then we rotated ours clockwise, which gave us, 4 would be in the upper left-hand corner and you'd go [4123] around that way [clockwise]. We rotated ours clockwise. I think you said it should have been, what? [2314]?

[2341], yes. Which would have been a counterclockwise rotation of once around.

The "coincidence" resulting from the fact that R_{90} has order 4 is that the position interpretation of the right turn leads to the same result as the object interpretation of the left turn. When a different correspondence was suggested, the students had a tendency to change the transformation (from right rotation to left rotation) rather than to change the interpretation (from position interpretation to object interpretation). Such behavior suggests that the position interpretation is not just an occasional preference—it may be a very salient interpretation by novices in group theory. If so, then why? Is it an artifact of instruction or is it based in deeper perceptual and conceptual factors? We regard this as a matter deserving close empirically-based attention.

Similar Phenomena

We would like to situate our considerations of two ways to interpret the symmetries of a square with a collection of other phenomena which have some similarities to, but also differences with our situation. We will consider

(1) the alibi/alias dichotomy introduced by Birkhoff and Mac Lane [3, 9] considered in the case of translation in $\mathbf{R}^2$ [3, p. 238] and rotation in $\mathbf{R}^2$ [10, p. 75].
(2) the alibi/alias dichotomy for quadratic forms [3, pp. 250–251; 9, pp. 387–388],
(3) an observation of Halmos on the matrix of a linear transformation [7, p. 65] and
(4) the effect of an automorphism of a vector space on the matrix of a linear transformation.

First, we will describe these phenomena and then compare and contrast them with our object/position dichotomy.

Example 1: Alias/alibi for translation and rotation in $\mathbf{R}^2$.
Birkhoff and Mac Lane [3, p. 238] suggest that an affine transformation of $\mathbf{R}^2$ into itself, can be interpreted as an *alibi*: a transformation, in which each point x is carried into a point y on the same coordinate system, or an *alias*: a change of coordinates, in which the original coordinate system is replaced by a new one. For example, the equations

$$y_1 = x_1 + 2\ , \quad y_2 = x_2 - 1$$

can be seen as a translation of every point in the plane two units east and one unit south (alibi) or as a change of the coordinate system to a parallel system with the origin two units west and one unit north of the given origin (alias). We note that both interpretations change the representation, or name, of the point to the same new name: e.g., the point $(0, 0)$ is either carried to the point whose coordinates are $(2, -1)$ in the original coordinate system or the point remains in the same place but, in the new coordinate system, it has coordinates $(2, -1)$. We also note that the way this is done in the alibi case is by applying the transformation to points and it is done in the alias case by applying the inverse of the transformation to the coordinate axes (or basis).

A similar example was discussed by Synge [10] who considered the following equations in the Euclidean plane

$$x' = x\cos\theta + y\sin\theta$$

$$y' = y\cos\theta - x\sin\theta$$

as a transformation which can be interpreted in two different ways. "First, we may think of fixed axes, and of a point which at first has coordinates (x, y) and later has coordinates (x', y'), all coordinates being measured with respect to the fixed axes. In this way of looking at the transformation, the axes are fixed and the plane rotates through an angle θ, in the sense from Oy towards Ox. Secondly, we may think of a fixed point, referred first to axes Oxy and secondly to axes $Ox'y'$, the axes $Ox'y'$ being obtained from Oxy by rotating them through an angle θ in the sense from Ox towards Oy. Now the plane is fixed and the axes "rotate"[10, p. 75]. We note here that even though Synge didn't name the interpretations alibi/alias, his two interpretations are essentially the same as those of Birkhoff and Mac Lane. In both cases discussed in Example 1, both the alias and the alibi interpretations lead to the same result in that both change the coordinates to the same new coordinates, and in order to do this, it is necessary to make use of the inverse of the original transformation.

Example 2: Quadratic forms under automorphism.

In a later work, Mac Lane and Birkhoff [9] applied the alibi/alias analysis to several situations such as similar matrices and the signature of a quadratic form. Our next example is adapted from Birkhoff and Mac Lane, [3, pp. 250–251] and Mac Lane and Birkhoff, [9, pp. 387–388]. It discusses what happens to the (symmetric) matrix of a quadratic form with respect to a basis when an automorphism is introduced. Recall that a quadratic form q on a vector space V over $\mathbf{R}$ is a function $q : V \to \mathbf{R}$ such that the following expression defines a bilinear function on $\mathbf{R} \times \mathbf{R}$.

$$q(v + u) - q(v) - q(u), \quad v, u \text{ in } V.$$

Consider a finite dimensional vector space V, a quadratic form q with the symmetric matrix A relative to some basis (x_i). This means that if X is the coordinate function which assigns to a vector v in V the coefficients of its expansion in terms of the basis (x_i), then we have in vector and matrix notation,

$$q(v) = X(v)A(X(v))^T$$

for all v in V. (Here, $X(v)$ is considered to be a "row" vector and $(X(v))^T$ a "column" vector.) Now, suppose we have an automorphism $s : V \to V$, and that its matrix relative to the basis (x_i) is P. This means that $P = (p_{ij})$ and

$$s(x_i) = \sum_j p_{ij}\, x_j.$$

What is the effect on the (symmetric) matrix of the quadratic form v if the automorphism s is applied to V?

We can again make either an alibi interpretation in which we consider that s changes the vectors in v or an alias interpretation in which the vector remains the same, but the coordinates are changed.

For the alibi interpretation, we consider that s changes a vector v to $s(v)$. The coordinates of v are thereby changed from $X(v)$ to the coordinates of $s(v)$, which are $X(s(v))$. Using the transition matrix, we have

$$X(s(v)) = X(v)P$$

Then we can write,

$$\begin{aligned} q \circ s(v) &= q(s(v)) \\ &= X(s(v))A(X(s(v)))^T \\ &= (X(v)P)A(X(v)P)^T \\ &= X(v)(PAP^T)(X(v))^T \end{aligned}$$

That is, relative to the basis (x_i) the matrix of q is changed to the matrix of $q \circ s$, which is the symmetric matrix PAP^T.

Now for the alias interpretation we may consider that the coordinates of a given vector v are changed from $X(v)$, its coordinates with respect to the basis (x_i), to its coordinates with respect to a new basis (y_i) given by

$$y_i = s(x_i)$$

The new coordinates of a vector v, with respect to the basis (y_i) satisfy

$$X(v) = Y(v)P \quad \text{or} \quad Y(v) = X(v)P^{-1}.$$

So we may write,

$$\begin{aligned} q(v) &= X(v)A(X(v))^T \\ &= (Y(v)P)A(Y(v)P)^T \\ &= (Y(v)P)A(P^T(Y(v))^T) \\ &= Y(v)(PAP^T)(Y(v))^T. \end{aligned}$$

That is, the matrix of q relative to the basis $(s(x_i))$ is the symmetric matrix PAP^T.

We note that both interpretations result in the same new matrix of the quadratic form. We also note that in going from the alibi to the alias interpretation it is necessary to replace the transformation of coordinates $X(v)P$ to the inverse transformation, that is, $X(v)P^{-1}$.

Example 3: Matrix of a linear transformation.

The following example is discussed in Halmos [7, p. 66]. Consider a finite dimensional vector space V and a basis (x_i) for V. If T is a linear transformation on V, then it can be represented by a matrix. One way to do this is to write out the expansion of each $T(x_i)$ in terms of the basis and lift the coefficients.

Thus, for example if D is the differentiation transformation on the vector space of polynomials of degree $< n$ and we take the monomials for the basis, then we can write

$$\begin{aligned}
Dx_1 &= 0x_1 + 0x_2 + \cdots + \quad 0x_{n-1} + 0x_n \\
Dx_2 &= 1x_1 + 0x_2 + \cdots + \quad 0x_{n-1} + 0x_n \\
Dx_3 &= 0x_1 + 2x_2 + \cdots + \quad 0x_{n-1} + 0x_n \\
&\cdot \qquad \cdots \qquad \cdot \\
Dx_n &= 0x_1 + 0x_2 + \cdots + \quad (n-1)x_{n-1} + 0x_n
\end{aligned}$$

so that deleting everything but the explicit numbers leads to the matrix,

$$[D] = \begin{vmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 2 & 0 & \dots & 0 & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & n-1 & 0 \end{vmatrix}$$

On the other hand, if we wish to use the standard convention of applying a linear transformation to a vector by placing the coefficients of the vector on the right of the matrix as a column vector and multiplying, then we have to first transpose the matrix. Thus, we define the matrix of a transformation T with respect to a basis (x_i) to be (a_{ij}) where i indicates the row, j indicates the column and a_{ij} is the j^{th} component of the expansion of $T(x_i)$ with respect to the basis (x_i) . This choice of the transpose of $[D]$ and not $[D]$ to represent the linear transformation was referred to by Halmos [7, p. 66] as "the unpleasant phenomenon of indices turning around."

Example 4: Effect of an automorphism on a matrix of a linear transformation.

Let $T : V \to V$ be a linear transformation and (x_i) a basis for V. Here we will couch our discussion of bases and the representations they provide in the language of duality, rather than transposes and row versus column vectors. We will denote by (a_{ij}) the matrix of T relative to the basis (x_i). It is given by

$$a_{ij} =< Tx_i, x^*j >$$

where (x^*j) is the dual basis of (x_i) for the dual V^* of V. Let s be an automorphism of V and we may consider two interpretations of how s might change the matrix. In the first interpretation, s changes the basis (x_i) to the basis (y_i) where $y_i = s(x_i)$. This gives rise to a new matrix (bij) given by

$$b_{ij} =< Ty_i, y_j^* >$$

where (y_j^*) is the dual basis of (y_i) for the dual V^* of V.

In the second interpretation, we consider that s changes T by conjugation: that is, instead of applying T to V, one first applies s, then T and then s^{-1}. In this case, the matrix of T is now the matrix of the conjugated transformation $s^{-1}Ts$ relative to the original basis. Thus we have the new matrix (c_{ij}) given by

$$c_{ij} =< s^{-1}Tsx_i, x_j^* > .$$

Using the fact that the transpose of the inverse of an automorphism is the inverse of the transpose and also the fact that $sy_j^* = x_j^*$, we get:

$$\begin{aligned} c_{ij} &=< s^{-1}Tsx_i, x_j^* > \\ &=< Tsx_i, s^{-1^*} x_j^* > \\ &=< Ty_i, y_j^* > \\ &= b_{ij}. \end{aligned}$$

Thus we have an example that is very similar to the alias/alibi situations except that here the two interpretations still give the same new names but it is not necessary to switch to an inverse. One could, of course, argue that there is an inverse present in that transforming the basis (x_i) to the basis $s(x_i)$ amounts to applying the inverse of s to the elements of V. It is not necessary, however, to consider this inverse whereas in the alias/alibi situations, the inverse plays an essential role.

Comparing with Examples 1 and 2, we observe that the use of the inverse transformation in alias situations, to obtain the same new name as in alibi situations, arises only when the objects in the discussion are vectors in a finite-dimensional vector space (points in $\mathbf{R}^2$ in Example 1). This is due to the fact that applying a transformation to a basis results in applying its inverse to the coordinates of elements of a vector space. In Example 2 however, the objects in the discussion are quadratic forms and not vectors. The same automorphism s is applied either to the quadratic form (alibi) and to the basis (alias). The inverse of s appears in the discussion because we chose to perform the calculation at the level of vectors and their coordinates. In the following paragraph, we repeat the analysis of Example 2 avoiding the introduction of the inverse of a transformation by performing the calculations at the level of quadratic forms and their matrix representations.

Revisiting Example 2: Quadratic forms under automorphism.

Let V be a finite dimensional vector space and q a quadratic form on V. Given a basis (x_i) for V, there is a unique symmetric matrix (a_{ij}) which satisfies,

$$q(v) = \sum_{i,j} < v, x_i^* > a_{ij} < v, x_j^* >, \qquad \text{for all } v \text{ in } V.$$

Now suppose that s is an automorphism of V. Then we can interpret the effect of s in two ways. The alias interpretation is that s changes the basis (x_i) to the basis (y_i) where $y_i = s(x_i)$. This gives rise to a new symmetric matrix (b_{ij}) uniquely determined by the relation

$$q(v) = \sum_{i,j} < v, y_i^* > b_{ij} < v, y_j^* >, \qquad \text{for all } v \text{ in } V.$$

In the alibi interpretation, it is considered that s changes q to $q \circ s$. In this case, the matrix of the new quadratic form, relative to the original basis is the unique symmetric matrix (c_{ij}) which satisfies,

$$q \circ s(v) = \sum_{i,j} < v, x_i^* > c_{ij} < v, x_j^* >, \qquad \text{for all } v \text{ in } V.$$

Now, write $w = s(v)$. Using the fact that $s^* y_j^* = x_j^*$ and that the dual of the inverse of an automorphism is the inverse of the dual, we get

$$\begin{aligned} q(w) &= \sum_{i,j} < s^{-1}w, x^* i > c_{ij} < s^{-1}w, x_j^* > \\ &= \sum_{i,j} < w, s^{-1^*} x_i^* > c_{ij} < w, s^{-1^*} x_j^* > \\ &= \sum_{i,j} < w, y_i^* > c_{ij} < w, y_j^* >, \qquad \text{for all } w \text{ in } V. \end{aligned}$$

By the uniqueness, (b_{ij}) is identical to (c_{ij}). Thus we see that the two interpretations lead to the same new matrix for the quadratic form and in both the alibi and alias interpretations, the automorphism s and not its inverse is used to make the change.

It may be that applying a transformation to points, but the inverse of the transformation to the basis (as in Examples 1 and 2) will appear arbitrary to students and cause some confusion. If that is the case, the fact that in an analysis such as this there is no such switch to the inverse may cause this situation to be less confusing. These are research questions which should be investigated.

Similarities in these situations.

In each of the situations we have described, there are certain objects and in each case there are two naming schemes or means of assigning to an object some algebraic quantity. Thus in Example 1, the objects are the points in $\mathbf{R}^2$ and the first naming scheme is the Cartesian coordinate system. The second naming scheme comes from the introduction of the translation or the rotation. In Example 2 the objects are the quadratic forms and the first naming scheme is the matrix of a quadratic form with respect to a basis. The second naming scheme comes from introducing an automorphism of a vector space. In Example 3 the objects are the linear transformations on a vector space V and the naming schemes are the two ways of lifting the matrix elements from a set of equations—with and without a transpose. In Example 4 the objects are the linear transformations on a vector space V and the first naming scheme is the matrix of a linear transformation with respect to a basis (one of the choices in Example 3). The second naming scheme results from the introduction of an automorphism of a vector space. Finally, in our situation, rigid motions of a square, the objects are the symmetries of the square and the naming schemes are the two ways of assigning a permutation to a symmetry that come from the object/position interpretations.

Before moving over to focus on the differences, we should acknowledge that because of the existence of two naming procedures, there is a potential for mixing them up with a resulting confusion and even error.

Differences among the situations.

One very concrete difference is that in some of these situations such as position/object, and matrix of a linear transformation, the two procedures for naming give different names, whereas in the others, the two new names are the same. In fact, in the first stages of our investigation, after considering Example 1 only as a case of alibi/alias dichotomy, we assumed that the same name was due to the choice to use one transformation in alibi interpretation and its inverse in alias interpretation, which may seem a bit unnatural. If the inverse were not introduced, the two interpretations would give different results. However, our further investigation revealed that the use of the inverse is not essential when moving to a higher level objects, such as quadratic forms or linear transformations.

There is another significant difference, that appears to us more fundamental when comparing the above examples. Consider the object/position situation as opposed to the alibi/alias dichotomy. In the former case, attention is focused on a single object (a symmetry) which is not in any way changed in the discussion. There are, however, two different interpretations that lead to two procedures for naming this object as a permutation. In the latter case, however, attention is again focused on a single object (point on a plane) and a single naming scheme (assignment of Cartesian coordinates), so that no ambiguity exists at first. But then, a transformation is introduced and the question arises of whether to apply it to the objects or to the naming schemes. The two possible answers to this question then give rise to two new procedures for naming the object. Thus, in this second case, one has objects with original names (before the transformation) and then (after the transformation) two possibilities for new names for a total of three names (two of which may be the same) connected with each object. In the former case, there is, from the beginning two names for each object, but that is all that we have. It seems to us that the matrix of a linear transformation is similar in this last respect to the object/position dichotomy while the situation in Example 4 could be another example of the alias/alibi dichotomy.

This last kind of difference seems to us to be essential and we can see no way of embedding all of these types of situations in a single analysis. That identifies one more question to which future research might be directed. However this research comes out, the similarities among the situations appear to be enough to warrant considering them together when searching for effective pedagogical strategies.

We have shown in this paper how difficulties can arise in our situation of object or position interpretations used to assign permutations to symmetries of the square. Synge expresses the opinion that these difficulties do not have to arise in the case of rotations of axes which he considered. He claims that when operations are easily followed intuitively, the change between interpretations is not expected to cause confusion. He makes a point of the importance of making one's interpretation clear in space-time transformations, where "our intuition is not so active" [10, p. 75].

It seems that research is called for regarding the other situations to see if they can be the cause of any student difficulties. If there are such indications, then the question arises as to what pedagogical strategies might help students make sense out of these situations and their multiple interpretations.

Conclusion

Our investigation revealed confusion when Abstract Algebra students attempted to connect symmetries of a square with permutations of S_4. Based on these few students we can say that one possibility is suggested. The position/local interpretation may be more natural for students to use when making a formal correspondence between symmetries and permutations whereas the object/global interpretation may be more natural to use when deciding how to actually move a square (after one or more transformations have been made). Such a situation would, of course, represent an error waiting to be made.

Clearly, transformation groups as well as permutation groups are important examples in the introductory Abstract Algebra course and establishing relationships between the two can be very beneficial for the learner. "The great essential is to try to be quite clear which view we are taking in any particular argument, because otherwise great confusion may result." [10, p. 76]. Therefore, we suggest that the issues discussed above should not be "swept under the rug" or left as exercises, but treated with sufficient attention paid to, and acknowledgment of, the difficulties involved. Of course that is easy to say, but we cannot forget that simply pointing things out to students has very little effect on helping them understand something. Pedagogical strategies must be devised that will help students become aware of these subtleties and use them to make sense out of these important early examples of groups.

However, a more general goal of an Abstract Algebra course is to use the D_n and S_n relationship to introduce mathematics students to broadly applicable skills and invoke or increase their awareness of complexities and non-uniqueness of mathematical interpretations. Our presentation is only an overture to a variety of areas in which such skills or awareness could prove useful. This article is a step in the direction of clarifying a kind of complexity that undergraduate mathematics instructors should be attuned to in order to facilitate the success of their students.

References

1. Abelson, H. and diSessa, A., *Turtle Geometry: the Computer as a Medium for Exploring Mathematics*, MIT Press, Cambridge, MA, 1981.
2. Armstrong, M. A., *Groups and Symmetry*, Springer/Verlag, New York, 1988.
3. Birkhoff, G. and Mac Lane, S., *A Survey of Modern Algebra*, 3rd Edition, MacMillan, New York, 1965.
4. Dubinsky, E., *Reflective abstraction in advanced mathematical thinking*, Advanced Mathematical Thinking (D. Tall, ed.), Kluwer Academic Publishers, Boston, MA, 1991.
5. Dubinsky, E., Dautermann, J., Leron, U., and Zazkis, R., *On learning fundamental concepts of group theory*, Educational Studies in Mathematics **27** (1994), 267–305.
6. Fraleigh, J. B., *A First Course in Abstract Algebra*, 4th edition, Addison Wesley, Reading, MA, 1989.
7. Halmos, P. R., *Finite Dimensional Vector Spaces*, D. Van Nostrand Company, Princeton, NJ, 1958.
8. Leron, U. and Zazkis, R., *Of geometry, turtles and groups*, Learning Mathematics and Logo (C. Hoyles and R. Noss, eds.), MIT Press, Cambridge, MA, 1992, pp. 319–352.
9. Mac Lane, S. and Birkhoff, G., *Algebra*, MacMillan, New York, 1967.
10. Synge, J. L., *Relativity: the Special Theory*, 2nd Edition, North-Holland Publishing Company, Amsterdam, 1964.
11. Zazkis, R., Dubinsky, E., and Dautermann J., *Coordinating visual and analytic strategies: A study of students' understanding of the group D_4*, Journal for Research in Mathematics Education (in press).

SIMON FRASER UNIVERSITY

PURDUE UNIVERSITY