Main Goal: To interpret the product of matrices as a composition of linear transformations

By the end of the lesson students should be able to:
– Coordinate input and output vectors across a sequence of transformations
– Determine the matrices that yield a desired sequence of transformations
– Conceive of the product of two matrices as a composition of two linear transformations
– Understand how the order of matrix multiplication coordinates with the order in which linear transformations are composed
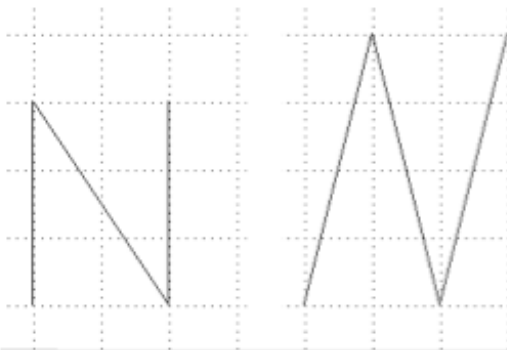
The instructor should:
– Connect students' work to composition of functions as an interpretation of matrix multiplication
– Introduce the terms commutative and non-commutative after the students complete the task

Handout 3: Pat and Jamie

An image of Handout 3 is given below. See for a printable student version of the handout.

PAT AND JAMIE

Name _____ Group Members _____

Suppose the "N" on the left is written in regular 12-point font. Find a matrix A that will transform N into the letter on the right, which is written in italics in 16-point font.

Last semester, two linear algebra students—Pat and Jamie—described their approach to the Italicizing N Task in the following way:

"In order to find the matrix A, we are going to find a matrix that makes the "N" taller (from 12-point to 16-point), find a matrix that italicizes the taller "N," and the combination of those will be the desired matrix A."

1. Do you think their approach allowed them to find a matrix A? Does it seem sensible? If so, do you think it was the same matrix A we found this semester?

2. Try Pat and Jamie's approach. You should either: (a) come up with a matrix A by using their approach, or (b) be able to explicitly explain why this approach does not work.

Use your group's whiteboard as a space to write out our work as you work together on this problem.

Handout 3: Pat and Jamie

Learning Goals and Rationale

Goals

The main goal of the problem on Handout 3 – Pat and Jamie – is to help students come to interpret matrix multiplication as a composition of linear transformations.

Rationale

This task sets up a scenario in which students must first decide if the approach of two "fellow students," Pat and Jamie, is valid and, if so, determine the matrices that represent the transformation via Pat and Jamie's approach. Note that in the problem statement the order in which Pat and Jamie transform the N is not vague (they make it taller first and then italicize it), but the way in which to computationally accomplish this is purposefully left vague. It is anticipated that students will struggle with how to combine and symbolize one transformation followed by another. This requires students to coordinate input and output vectors across a sequence of transformation, and to relate that coordination to the order in which the sequence of corresponding matrices is multiplied.

The Pat and Jamie task sets the stage for a shift in students' mathematical activity and goals; students are asked to *combine matrices* that define transformations in addition to determining what those matrices are. This is a shift from the goal of constructing a single transformation matrix by coordinating sets of input and output vectors, such as in the Italicizing N task.

Handout 3: Pat and Jamie

Student Thinking

Group 1:
Group 1 correctly found the matrix that stretches the 12-pt font to 16-pt font, but incorrectly found Pat and Jamie's "italicize" matrix. Note that the matrix $\begin{bmatrix} 1 & 1/3 \\ 0 & 1 \end{bmatrix}$ italicizes the original 12-pt font N, not the taller 16-pt font N. Having found these two matrices, the group was unable combine them in a way that would yield matrix A from Task 1.

Group 2:
This group found the same matrices as Group 1, but also realized that matrix multiplication would yield A. The students did realize that the order in which you multiply the matrices does make a difference, but did not realize that their matrices do not transform N in the same way that Pat and Jamie did. Instead, Group 2 found matrices that lean (or tilt) N and then make N bigger. It is unclear whether these students are aware of this.

Group 3:
This group correctly found the two matrices to "make N bigger" and "italicize N" and also multiplied them in the correct order to yield the matrix A. Notice, though, that the same matrix produced by Groups 1 and 2 is crossed out in the upper left-hand corner of their whiteboard. We also see from their work that the students found first attempted to add the matrices, then multiplied them in the incorrect order before finally multiplying them in the correct order. From this, it is unclear whether the students understood why this using this operation in this order produced a correct solution.

Groups 4 and 5:

Group 4's work (below) highlights how he students found the two transformations and the correct order of multiplication. The students also wrote "$CB = A$ which corresponds to bigger first and leaning second." Group 5's work (left) highlights a bit of the compositional nature of the task, where the "$(B(N))$" seems similar to the notion of "$f(x)$," with the matrix C then acting on the output of $B(N)$. Note that both of these responses only begin to hint at the compositional nature of the two transformations, with the output of the transformation defined by B becoming the input for that defined by C. It is the role of the teacher to facilitate explicit discussion of this issue. (*Note*: These students had agreed that the "lean" matrix should be called C.)

Handout 3: Pat and Jamie

Implementation

Getting Started with the Task

The Pat and Jamie task sets the stage for a shift in students' mathematical activity and goals; students are asked to *combine* matrices that define transformations in addition to determining what those matrices are. This is a shift from the goal of constructing a single transformation matrix based on inputs and outputs, such as in the Italicizing N task.

Encourage students talking in their small groups about whether Pat and Jamie's approach seems sensible and if that approach should give the same final matrix A that was found in the initial Italicizing N task. (See Questions 1-2 from the Handout.)

Working in small groups towards a solution

The class begins the Pat and Jamie task by talking in their small groups about whether Pat and Jamie's approach seems sensible and whether that approach should give the same final matrix A that was found in the initial Italicizing N task. The teacher plays an important role in facilitating the subsequent whole class discussion, eliciting students' initial ideas. Students tend to take the view that Pat and Jamie's approach seems sensible, and that it should yield the same matrix A as the previous task. We have found it useful to make these explicit in whole class conversation prior to students' work on Question 2 from the Handout.

During small group work, the instructor should notice the various approaches that the students are taking - noting which examples of student thinking might be leveraged in whole class discussion. Here we provide typical student responses/difficulties and recommend how they might be leveraged in whole class discussion:

Students will likely begin with the one of the approaches they used in Tasks 1 and 2. For instance the students might first determine the image of "N" under a desired transformation and then use an unknown matrix to find a system of equations (a, right). Another approach would use the generalizations discussed in Task 2 to determine what matrix would distort $\square^2$ in a desired way (b).

a) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \Rightarrow \begin{matrix} 0a + 3b = 0 \\ 0c + 3d = 4 \end{matrix}$

and $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} 2a + 0b = 2 \\ 2c + 0d = 0 \end{matrix}$

b) $\begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix}$ stretches $\square^2$ vertically by a factor of $\frac{4}{3}$

Between these two approaches, each group should be able to determine correctly a matrix that would make the "N" taller (from 12-pt font to 16-pt font, as described in the task set-up). This matrix is $\begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix}$. The more difficult part of this task (from the students' perspective) is usually the coordination of this "taller" transformation with a new transformation that would italicize the new, 16-pt font "N." Specifically, students using approach (a) from above often have difficulty determining appropriate input-output pairs to use with the unknown matrix to generate a system of equations. For instance, students may try to find a matrix that italicizes the original 12-pt N (Groups 1 and 2 in Student Thinking section). Because of this, students often do not view the output of the first transformation as the input for the second transformation, which Pat and Jamie's method requires. The difficulty with using the second approach (anticipating the entries of the matrix by observing the geometric distortion of the plane) is that the italicizing matrix skews the plane, and students may not yet know the form of a matrix for that geometric transformation.

Short break into Whole Class Discussion

If many groups' progress is stalling here (trying to determine the "leaning" matrix in Pat and Jamie's approach), the instructor should reconvene everyone into a whole class discussion.

This "mini whole class discussion" would have two main purposes:
- To allow students to describe their progress, articulate what they are finding difficult about the problem, and bounce ideas off of each other of how to proceed
- To allow the teacher to suggest that the class agree upon a shared notational system for the problem

In service of the first goal, as students start to share their reasoning and progress, the instructor may wish to have everyone agree on names and notation for the various matrices (which serves the second goal), which the class should readily agree would help in their progress.

> For instance, the instructor could ask students what label they gave the matrix that made the N "taller" and use that. Alternatively, the instructor could suggest they all agree to use "B" (for "make bigger") to notate this matrix.
>> **NOTE:** In the past, we have avoided using "T" for the matrix that makes the N "taller" because of the use of "T" to denote a linear transformation throughout this section.

> Similarly, the instructor could ask students what label they gave the matrix that made the N "lean" and use that. Alternatively, the instructor could suggest they all agree to use "L" (for "make lean").
>> **NOTE:** In the past, we have avoided using "I" for the matrix that "italicizes" the N because "I" is reserved for the identity matrix and identity transformation in Task 4.

> It is often the case that all of the students develop the same B matrix, but not all students develop the same L matrix. The instructor may recommend naming these matrices L_1 & L_2, or L & L', etc.

What's Happening Here Mathematically and How Do I Use It?

> The mathematics below surfaces eventually through students' work – most often after students are sent back into small groups after the above mini-whole class discussion. We include it here so that instructors can be prepared for the variety of student responses.

> We stress that stress that students may not even think to use matrix multiplication at first, much less to know why it is sensible in this task. These mathematical relationships **do** need to come out in whole class discussion by the end of Task 3.

The students who use the 16-pt "N" as input for L will probably find the matrix L (or L_1) $= \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & 1 \end{bmatrix}$.

The students who use the 12-pt "N" as input for L will probably find the matrix L' (or L_2) $= \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 1 \end{bmatrix}$.

All students will probably find the matrix $B = \begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix}$. Notice that:

$BL \neq A$ but $BL' = A$:

$$BL = \begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{4} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & \frac{4}{3} \end{bmatrix} \neq A = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 1 \end{bmatrix} = BL'$$

and $LB = A$ (which matches what Pat and Jamie did) but $L'B \neq A$:

$$LB = \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & \frac{4}{3} \end{bmatrix} = A \neq \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{4}{3} \end{bmatrix} = L'B.$$

Once issues of notation are balanced, the instructor should prompt students for general strategies for moving forward with the Pat and Jamie task. Some groups may have (correctly) used the vectors from the 16-pt "N" as the inputs for the italicizing transformation to correctly determine the "lean" matrix L. The teacher may choose to summarize and revoice this explanation using mathematical symbolism that highlights how the output of the first transformation becomes the input for the second and illustrate this with particular input-output pairs (e.g., see Figure 3.1). This discussion is important because it lays a foundation for students' understanding matrix multiplication as composition of linear transformations.
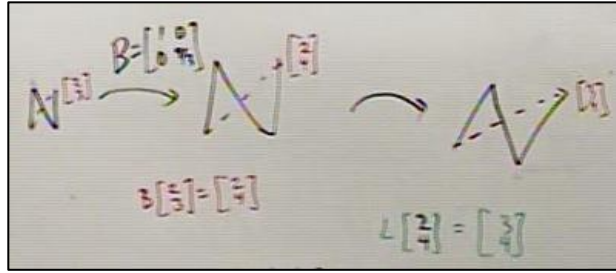
Figure 3.1. The instructor built from students' verbal explanation by sketching the intermediate N, writing $B\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}2\\4\end{bmatrix}$ and $L\begin{bmatrix}2\\4\end{bmatrix} = \begin{bmatrix}3\\4\end{bmatrix}$

The students should be able to continue in small group work toward explicitly finding the matrices B and L (if they have not already done so) and determining how the two matrices are combined (and why) to produce matrix A from the original Italicizing N task.

Example Prompt for Small Group Work
"Thanks for sharing your initial ideas. We still have some work to do to confidently decide what "bigger" matrix B and "lean" matrix L represent Pat and Jamie's approach. Also, we need to decide how to combine these matrices — we know (from #1 on the handout) that the matrices should combine to equal A. But what method of combination makes sense, and why? If we didn't know they should equal A, how could we know what to do?"

Back in Small Groups
If they hadn't done so already, all students should now be ready to find at least one lean matrix. The task is now to find a way to sensibly combine these two matrices to produce A, the matrix that transforms the normal size 12 N to the italicized size 16 N. It is likely that a few groups will try to add the matrices component-wise. Other students may recall the matrix multiplication algorithm from previous classes. In the latter case, these students may multiply both BL and LB (or $L'B$ and BL'). The students should see that only one of these products (in each pair) results in A, which the instructor could leverage into a discussion of the noncommutativity of matrix multiplication.

Specifically, the students who italicized the 12-pt "N" will need to realize that, although they can multiply their matrices to create A ($BL' = A$), these matrices do not represent the transformations described by Pat and Jamie's approach.

This is an example of two distinct matrices (L and L') that, when multiplied by B, yield A but not when multiplied in the same order (i.e., $BL \neq BL'$). Even more, $BL \neq LB$ and $BL' \neq L'B$.

Reconvening in Whole Class Discussion
Again, students' main difficulties in this task are:
- Using the outputs from the "make taller" transformation as inputs in the "lean" transformation to find the correct matrix for the "lean" transformation
- Realizing the order of matrix multiplication that matches Pat and Jamie's approach is multiplying the matrix B on the left by "lean" matrix L.

These common challenges serve as valuable points for whole class discussion that provide students with opportunities to consider ways to coordinate input and output vectors across multiple transformations, as well as how to interpret matrix multiplication as composition of functions.

The instructor should, if possible, find representatives of student work to discuss in whole class: one that tried $BL' = A$, and one that tried $LB = A$.

Example Discussion Question:
"Ok here we have two groups that represent the two main approaches I saw out there: it

looks like this group decided on $BL' = A$ as Pat and Jamie's approach, and the other decided on $LB = A$. Can both approaches be correct? If not, how to do we decide?

Note: Some students may not have experience multiplying matrices. If this is the case, it is often useful to discuss matrix multiplication as an extension of the product of a matrix and a vector. Extending this, the product of the two matrices is the concatenation of the individual matrix-vector products. While it is not the only approach the instructor may choose to discuss matrix multiplication, this discussion does rely on the students' current ways of operating with matrices and vectors (i.e., making use of their knowledge of $Ax = b$ to understand $AX = B$ as $A[x_1 \quad x_2] = [b_1 \quad b_2]$. In a more proof-oriented course, the instructor may view this discussion as insufficient or unnecessary.

Emphasizing the Connection to Composition of Functions
The instructor should wrap up Task 3 by connecting students' work to the concept of function composition. Through this mini-lecture, the instructor can use students' work to highlight how, in Task 1, $Ax = y$ was interpreted as the linear transformation defined by A, acting on the input vector **x** to produce the output vector **y**. Now, in Pat and Jamie's approach, that vector **x** would first be acted on by the matrix transformation B, then that result (let's call it **z**, so $Bx = z$) would be acted on by L to produce **y** (so $Lz = y$). In other words:

x could be turned into **y** in one step via A: $Ax = y$
x could be turned into **y** in two steps via B and L: $L(Bx) = L(z) = y$
Therefore, $Ax = y$ and $LBx = y$
This works for any input vector **x**, implying $A = LB$.

The instructor should state that, when considering matrices as linear transformations, matrix multiplication is interpreted as a composition of functions.

- Instructors may wish to bring up a few examples from high school algebra to remind students that they are familiar with how composition of functions works:
  - $f(x) = x + 4$ and $g(x) = x^2$: $f(g(2)) = f(4) = 8$ but $g(f(2)) = g(6) = 36$. Thus $f(g(x)) \neq g(f(x))$ for every **x** in the domain.
  - Some classes may choose to further emphasize that the codomain of the first function serves as the domain for the second function. This gives a nice rationale (in terms of linear transformations) for why an m x n matrix can only be multiplied by an n x p matrix.

Defining Commutative and Noncommutative Matrices
Finally, the instructor should define the term **commutative**: We say two matrices D and F are commutative if $DF = FD$. If not, we say F and D are **not commutative**. Students should connect this to noncommutativity of function composition as well.

EXTRA RESOURCE: **Here is PDF of a ppt file** we have used in class to summarize the conclusions of Task 3.

Task 3 of the Italicizing N Sequence (Back to Top)

Handout 3: Pat and Jamie

Suggestions for Homework

Below are problems we have used in the past for homework. Instructors should feel free to use any subset of these that they feel would be beneficial to their students.

1. Suppose that $T: \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

a. Find the images of the following vectors under T:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

b. Sketch the original vectors on one graph and the images of the vectors on a second graph. Do something like the following, where the graph on the left sketches $\mathbf{x}$, $\mathbf{y}$, $\mathbf{w}$, $\mathbf{u}$, and $\mathbf{v}$, and the graph on the graph sketches their images under the transformation.
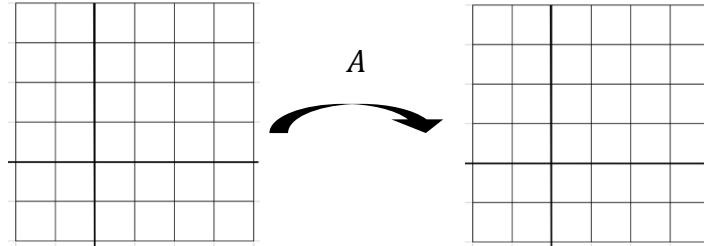
c. Describe in words what the transformation defined by A does to the vectors. What does the transformation do to the entire space? How did you reach this conclusion?

2. Suppose that $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is defined by $T(\mathbf{x}) = A\mathbf{x}$ for each of the matrices listed below. For each given matrix, answer the following questions:

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \quad E = \begin{bmatrix} 4 & 0 \\ 0 & 0 \\ 0 & 2 \end{bmatrix} \quad F = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}$$

a. Rewrite $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ with correct numbers for m and n filled in for each matrix. What is the domain of T? What is the codomain of T?

b. Find some way to explain in words and/or graphically what this transformation does in taking vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$. You might find it helpful to try out a few input vectors and see what their image is under the transformation. This might be difficult, but an honest effort will give you credit.

c. For the transformation, can you find two different input vectors that will give the same output vector?
    i. If so, give an example.
    ii. If not, why do you think it isn't possible?

d. For the transformation, can you get any output vector? (Any vector in $\mathbb{R}^m$)
    i. If so, explain why you can get any vector in $\mathbb{R}^m$
    ii. If not, give an example of an output vector you *can't* get with the transformation and explain why.

3. In class, we discussed matrix multiplication, when wanting to consider what happens to vectors or the whole space under the transformations, as composition of functions. We thought about how, if you have $AB\mathbf{x}$, where A and B are matrices and $\mathbf{x}$ is a vector, $AB\mathbf{x}$ could be thought of as B transforming $\mathbf{x}$ first, and then A transforming the result of $B\mathbf{x}.$
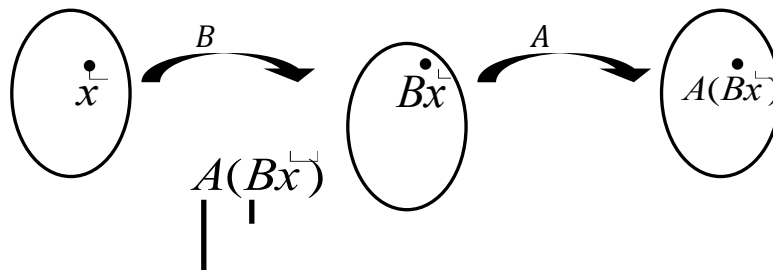
Now consider the composition of real-valued → Transformed by *B* first / The result transformed by *A* functions, familiar from high school. Use the given functions *f*, *g*, and *h* below and carry out the requested computations.

$$f(x) = 2x + 4 \quad g(x) = x^2 - 3x \quad h(x) = \frac{\sqrt{x}}{x - 2}$$

a. $f(g(x))$ and $g(f(x))$
b. $f(g(2))$ and $g(f(2))$
c. $h(f(x))$
d. $h(f(0))$ and $h(f(-1))$
e. $f(g(h(x)))$
f. Reflect on how composing these functions and evaluating at a value (such as at 2 in part b) is similar to what we saw with matrix multiplication through the Pat & Jamie Task on Wednesday. Write at least two thoughtful sentences.

4. Let the matrices *F* and *G* be defined as below. Answer the following questions accordingly.

$$F = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 0 \\ 0 & 3 & 4 \end{bmatrix}, G = \begin{bmatrix} 2 & 4 & 1 \\ 0 & 3 & -2 \\ 5 & 0 & 1 \end{bmatrix}$$

a. Let $x = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, and let **Gx = y**. Compute **Gx** and compute **Fy**.

b. Let $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, and let **Fx = u**. Compute **Fx** and compute **Gu**.

c. Explain how parts a) and b) illustrate matrix multiplication as composition of functions.

d. Explain how parts a) and b) illustrate that matrices *F* and *G* are not commutative.

5. Let *A* be a 5 x 4 matrix and *B* be a 4 x 3 matrix.
    a. What is the domain of the transformation defined by *B*? What is the codomain of the transformation defined by *B*?
    b. What is the domain of the transformation defined by *A*? What is the codomain of the transformation defined by *A*?
    c. What is the domain of the transformation defined by *AB*? What is the codomain of the transformation defined by *AB*?
    d. Explain why the product *BA* is not defined. Ground your explanation in the concepts of composition of functions, domain, and codomain.