

## Examining consistency of student errors in vector operations using module analysis

Nekeisha Johnson and John B. Buncher

*Department of Physics, North Dakota State University, Fargo, ND, 58108, USA*

It has been well documented that introductory physics students struggle with vector addition and subtraction. We use the results of a multiple-choice assessment on one-dimensional vector addition and subtraction, administered to students in a large-enrollment algebra-based physics sequence, to explore the consistency in the types of errors students make. The assessment was analyzed using Module Analysis for Multiple Choice Responses, a type of network analysis that highlights groups of responses commonly given together. We find evidence for five distinct modules that are consistent across multiple semesters. Of these groups, two support the method of “closing-the-loop”, two are consistent with students performing the wrong operation, and one suggests that students provide an answer independently of whether they were prompted to add or subtract two vectors, though only for vectors that are anti-aligned.

## I. INTRODUCTION

Successful manipulation of vectors and vector quantities is critical to success in introductory physics courses, and extensive literature exists examining student difficulties with vector operations, specifically addition and subtraction [1–11]. Prior work has shown that students’ vector abilities are largely unchanged even after a year of physics instruction [2], students’ solution methods are influenced by the relative position and orientation of vectors [6, 7], and that students tend to stick with the same solution method across a variety of problems, even though a different method may be more appropriate [5].

Prior research has investigated the types of errors students make when adding vectors as arrows, and the frequency of those mistakes. Hawkins et al. hypothesized that the solution method students would use when solving vector addition questions would depend on the specifics of the problem; how the vectors were initially drawn (head-to-head, tip-to-tail, tail-to-tail, separated), the presence or absence of a grid on which the vectors were drawn, and the relative orientation of the vectors to a grid [7]. In interviews involving 8 students solving 10 separate questions, they found 7 out of 8 students used the same solution method through all problems presented. In a follow-up study, they found that the initial arrangement of vectors (tip-to-tail or tail-to-tail) had a significant effect on students’ solution method, with more students using a “head-to-tail” method when the vectors were presented in a tip-to-tail format and more students choosing a “bisector” method when the vectors were presented tail-to-tail. Barniol and Zavala later investigated the effect that problem context and vector positioning have on student solution methods for vector addition [6]. Problems 1-3 were presented in three different contexts (displacement, force, and no explicit context), and problems 4-6 were presented with different initial positioning (tail-to-tail, head-to-tail, and separated, all presented on the same grid). They found that students were more likely to use a “head-to-tail” solution method in the displacement context, but were more likely to draw vectors tail-to-tail in the force context. When vectors were presented to the students “head-to-tail”, students were more likely to make a “closing the loop” error than in the other orientations. In both the tail-to-tail and “separated” orientations, students were more likely to make a tip-to-tip error, draw a general bisector, or make an error by subtracting only one of the components of the second vector [6]. We extend this work by investigating if certain errors are likely to be chosen together, across questions, using Module Analysis for Multiple Choice Responses (MAMCR).

Brewe et al. were the first to propose MAMCR to analyze the relationships among students’ incorrect answers on the Force Concept Inventory (FCI)[15]. By examining the types of incorrect answers that are closely tied together in a network, it can provide insight into the consistency of student reasoning across different questions. An example is the “impetus” module found by Brewe et al., which is dominated by answers students could arrive at following a line of reasoning

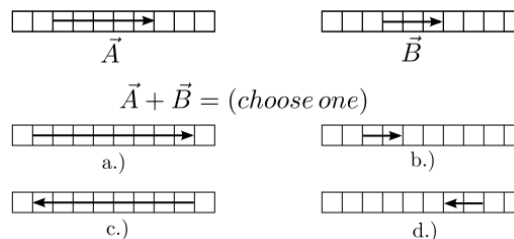


FIG. 1. Students were presented first with this addition problem. Each of the four possible responses comes from the various combinations of  $\pm\vec{A}$  and  $\pm\vec{B}$ . In this case, a) is  $+\vec{A} + \vec{B}$ , b) is  $+\vec{A} - \vec{B}$ , c) is  $-\vec{A} - \vec{B}$ , and d) is  $-\vec{A} + \vec{B}$ .

where a contact force (either constant or diminishing) continues to act on an object after contact has stopped. Recent work has expanded MAMCR to work with larger FCI data sets [19], applied this modified MAMCR to the FMCE [17], and extended it to include students’ correct answers on the FCI [18].

In this work, we apply Brewe et al.’s original MAMCR to the “arrows-on-a-grid” questionnaire originally designed to compare students’ ability to add and subtract vectors in an “arrows-on-a-grid” representation and an  $ijk$  representation [10]. By using MAMCR, we hope to determine if students make a particular type of mistake across multiple iterations of the same type of vector addition and subtraction problems. While we can use this methodology to examine if student errors are consistent across problems, it does not tell us explicitly what mistakes students are making. The coupling of understanding what kinds of incorrect answers are commonly given together, along with explanations of how students arrive at these incorrect answers, should guide us to targeted instruction aimed at each potential mistake.

## II. METHODS

### A. Data collection

The data for this study was collected in introductory algebra-based courses (both first and second semester). The courses met for 150 minutes per week with an optional 2-hour lab component. The courses used a mixture of Peer Instruction and traditional lecture, and included weekly online assignments [14]. Graphical vector addition was covered explicitly in the first-semester course, and was used in the second-semester course. One of the authors of this study (JB) was the instructor for all sections where data were collected.

Students were administered an online, multiple choice quiz at the end of the semester. The quiz was offered on an opt-in (2015) and an opt-out (2017 and 2018) basis, and consisted of 12 one-dimensional (1-D) and 16 two-dimensional (2-D) vector addition and subtraction problems. In this paper we examine only the 1-D problems: additional analysis of the 2-

D responses is in progress, including the use of handwritten work on some of the questions. For the 1-D questions there were six combinations of vectors  $\vec{A}$  and  $\vec{B}$ , and students were prompted to first add them (questions “+1” through “+6”) and then subtract them (questions “-1” through “-6”). Each problem had four possible choices, corresponding to  $+\vec{A} + \vec{B}$ ,  $+\vec{A} - \vec{B}$ ,  $-\vec{A} + \vec{B}$ , and  $-\vec{A} - \vec{B}$ . In all problems, vectors were presented on separate grids. The image accompanying question “+1” is presented in Fig. 1.

### B. Analysis

We employed Brewe et al.’s MAMCR process to analyze the student responses. All analysis was implemented using the R programming language and the igraph networking package. A network of the students’ responses was created, with the possible choices represented as nodes weighted by the number of students who selected that response. Edges connecting two nodes were weighted by the number of students who selected both responses. The “backbone” (or underlying structure) of the network was extracted using locally adaptive network sparsification (LANS) [20]. LANS calculates how critical each edge is to both connected nodes. If an edge is deemed insignificant at both nodes, the edge is cut from the network, leaving behind the backbone. For our purposes, we use a significance level of  $\alpha = 0.05$ .

As an analogy, consider using Google Maps to travel between city A and city B. The cities are the nodes of our network, and the roads connecting them are the edges. If there are multiple ways to travel between these two cities - an interstate, a state highway, and a county road - then the backbone from city A to city B would be the interstate, and LANS would remove the state highway and the county road. If city C is connected to city A by only a small county road, LANS leaves that as part of the backbone.

The backbone network was run through igraph’s built-in implementation of the InfoMap community detection algorithm [21]. From examining the nodes in each group (community), we aim to classify and describe the kinds of mistakes that students make on one-dimensional vector addition and subtraction problems.

## III. RESULTS

The data collected from four courses is summarized in Tab. I and Fig. 2. Since MAMCR uses only the incorrect answers in the analysis, we list this for each section in Tab. I. Figure 2 shows the percent of students answering correctly for each question. Figure 2 also shows that vector subtraction was consistently more difficult for students than vector addition in all courses, even in the 1-D case. We note that the results for question “-2” are excluded from Fig. 2 as there was an error in the version of the question that was administered during the Spring 2015 semester.

TABLE I. Vector addition and subtraction data collected from algebra-based introductory physics I (Sem1) and II (Sem2). Fall is considered on-sequence for Sem1, and Spring is considered on-sequence for Sem2. The students performed comparably on any one-dimensional question when compared across courses, regardless of if graphical vector addition was covered explicitly in the course.

Semester - Course	On/Off Sequence	N (Students)	n (incorrect)
Sp15 - Sem1	Off	63	109
Sp15 - Sem2	On	52	109
Fa17 - Sem1	On	175	491
Fa18 - Sem1	On	201	568

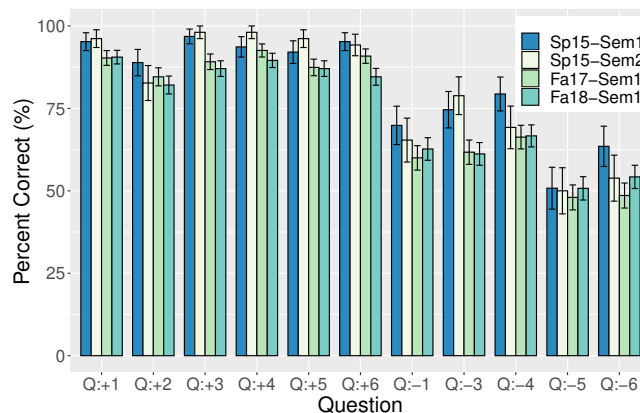


FIG. 2. Percent of correct responses on each one-dimensional problem. Students performed much better on vector addition than subtraction. Error bars represent  $\pm 1$  SE.

The remaining eleven questions are comprised of six addition and five subtraction problems, with each addition question having a corresponding subtraction version. Questions 1, 3, and 4 had vectors in an “aligned” orientation, with both vectors pointing to the right, while questions 2, 5, and 6 had vectors “anti-aligned”, with  $\vec{A}$  pointing to the right and  $\vec{B}$  pointing left. Students in these courses perform better on vector addition than subtraction, and perform better on vector subtraction when the two vectors are parallel (“aligned”) than when they are anti-parallel (“anti-aligned”), consistent with previous literature [9, 10].

In order to categorize groups, we applied a naming scheme that would allow us to find consistent patterns in answer types. This naming scheme is detailed in Tab. II, with each response referred to by one method students use to combine  $\pm\vec{A}$  and  $\pm\vec{B}$ . As an example, back in Fig. 1, option a) is +1:+A+B, option b) is +1:+A-B, option c) is +1:-A-B, and option d) is +1:-A+B. We emphasize that, without written work, we cannot be sure how students are arriving at a given answer. However, if a student arrives at an answer of -A+B on one problem and then uses the same process on another

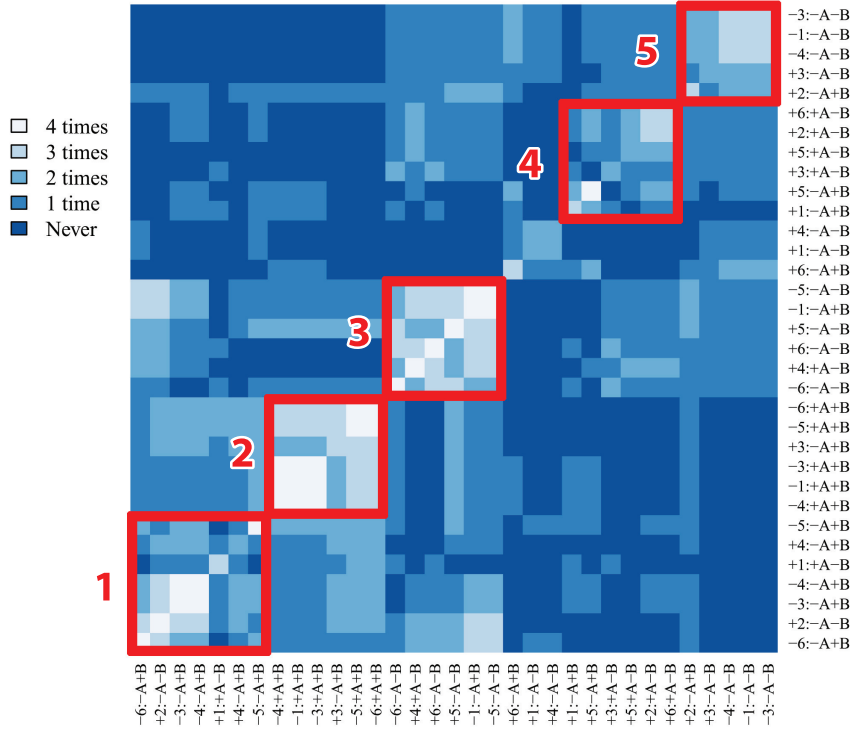


FIG. 3. Heatmap generated from the backbone networks from each of the four courses. The color of each square indicates how many times the two nodes appeared in the same groups across the four courses. The naming scheme is further explained in Tab. II. Some nodes are not present in some of the networks, resulting in an inconsistent maximum occurrence count. The ordering of the nodes was determined by R’s “heatmap” function, which computes a dendrogram from the co-occurrence matrix to determine how dissimilar two nodes are, using the euclidean distance between two rows as a measure of dissimilarity. Nodes that are less dissimilar are closer together on the axes. There appear to be two very strong groupings, groups 2 and 3, as well as three more moderate groupings, groups 1, 4, and 5. It appears plausible that some of these groups are connected in larger modules, but additional data is required to determine this.

problem, they will again get an answer of  $-A+B$ .

TABLE II. The node labelling scheme for 1-D presents four pieces of information: the operation the student was asked to perform, the question number (the set of vectors asked about), and the direction of both original vectors that could be correctly added to yield the student’s chosen resultant.

Operation	Vector Set	Direction of A	Direction of B
+ or -	1-6	+A or -A	+B or -B
addition or subtraction	question number (Q)	positive A or negative A	positive B or negative B

The communities for each course were extracted from 1000 trials of InfoMap. We then create a heatmap from the co-occurrence matrix across all courses, shown in Fig. 3. In this representation we can see how often InfoMap determined that two responses belonged to the same community across all four courses. The ordering of the nodes in Fig. 3 was determined by R’s “heatmap” function, which computes a dendrogram from the co-occurrence matrix to determine how dis-

TABLE III. Cross-course stable memberships for 1-D problems. For each group (highlighted in Fig. 3), the bold responses are those that follow the dominant trend of the group.

Group 1	Group 2	Group 3	Group 4	Group 5
<b>-6:-A+B</b>	<b>-4:+A+B</b>	<b>-6:-A-B</b>	+1:-A+B	+2:-A+B
+2:-A-B	<b>-1:+A+B</b>	+4:+A-B	+5:-A+B	+3:-A-B
<b>-3:-A+B</b>	<b>-3:+A+B</b>	<b>+6:-A-B</b>	<b>+3:+A-B</b>	<b>-4:-A-B</b>
<b>-4:-A+B</b>	+3:-A+B	<b>+5:-A-B</b>	<b>+5:+A-B</b>	<b>-1:-A-B</b>
+1:+A-B	<b>-5:+A+B</b>	-1:-A+B	<b>+2:+A-B</b>	<b>-3:-A-B</b>
+4:-A+B	<b>-6:+A+B</b>	<b>-5:-A-B</b>	<b>+6:+A-B</b>	
<b>-5:-A+B</b>				

similar two nodes are, using the euclidean distance between two rows as a measure of dissimilarity. Nodes that are less dissimilar are closer together on the axes. It should be noted that the center diagonal represents how often a response was present in a backbone network. Some nodes are not present in every backbone - students in a course either never chose that response, or in rare cases, that response was chosen only with

correct answers and therefore was dropped from the backbone network.

The heatmap in Fig. 3 shows two strong groups, brighter white at middle (group 3), and just below and to the left of middle (group 2), and three slightly weaker groups, two at top right (groups 4 and 5) and one at bottom left (group 1). The membership of each group is listed in Tab. III, where the groups are numbered from bottom-left to top-right. Some of these groups are possible subgroups of larger modules, but we would need more data sets in order to determine this.

#### IV. DISCUSSION

Based on the literature, we attempted to classify these resulting modules into possible error types. Each group indicated in Fig. 3 is addressed individually, from bottom-left to top-right, to present possible arguments or misconceptions for the bold responses in Tab. III.

Group 1 - The dominant mistake in this module, present in four of the the seven nodes, is  $-Q:-A+B$ . This mistake is consistent with two plausible solution methods. The first method is for students to perform subtraction backwards, reversing the direction of  $\vec{A}$  instead of  $\vec{B}$ . Alternatively, students could be doing the set up correctly, aligning  $\vec{A}$  and  $-\vec{B}$  tip to tail, but then closing the loop and going from the tip of  $\vec{B}$  to the tail of  $\vec{A}$ , rather than from the tail of  $\vec{A}$  to the tip of  $\vec{B}$ .

Group 2 - There appears to be a subset of students who will consistently do addition instead of subtraction. Group 2 contains all of the responses that follow this pattern (recall that we had to drop question “-2”). It is interesting that in each class section the largest nodes are  $-1:+A+B$ ,  $-5:+A+B$ , and  $-6:+A+B$ . Since questions were presented sequentially, “-1” was the first combination of vectors that students were asked to subtract. It is possible that some of the weight of the  $-1:+A+B$  node simply came from the students being on autopilot - to further examine this, we plan to shuffle the order of the questions in future iterations of the quiz. The prevalence of  $-5:+A+B$  and  $-6:+A+B$  is interesting because these are the two subtraction questions with anti-aligned vectors.

Group 3 - This module includes the same answer,  $-A-B$ , for both the addition and subtraction versions of problems 5 and 6. These questions have anti-aligned vectors, and that these nodes often end up in the same group together suggests that students are making the same mistake for both addition and subtraction. This could stem from students correctly adding  $-\vec{A}$  and  $-\vec{B}$ , or closing the loop when combining  $\vec{A}$  with  $\vec{B}$ .

Group 4 - Intriguingly, there seems to be a subset of students that gave the answer consistent with doing subtraction when prompted to do addition. It is interesting to note that the three nodes that seem to have the strongest connections within this module are  $+2:+A-B$ ,  $+5:+A-B$ , and  $+6:+A-B$ , which are the three pairs of anti-aligned vectors. If students were to reverse the direction of  $\vec{B}$  and then add it to  $\vec{A}$ , they would

come out to this result. This suggests that students are conflating vectors pointing to the right with a positive vector, and therefore addition.

Group 5 - In this last module, there are three responses that follow the  $-Q:-A-B$  pattern. Interestingly, these three responses are the three subtraction problems with aligned vectors, while the other two  $-Q:-A-B$  responses (from anti-aligned vectors) were in Group 3. This suggests that different students make this same mistake depending on the alignment of the vectors. There are again two plausible errors that we consider. First, these answers are consistent with adding  $-\vec{A}$  to  $-\vec{B}$ . Students may be inclined to do this if they associate vectors pointing to the left with negative - and therefore subtraction as well. The other possible mistake is students may have chosen the negative of the result from vector addition, again consistent with closing the loop.

#### V. FUTURE WORK

Moving forward, collecting handwritten work on the same set of questions would validate or refute our interpretations of each group. We are currently analyzing the results of the two-dimensional questions, not discussed here, for which we have both multiple-choice responses and handwritten work. A modified version of the 1-D questions could also be useful, with some questions having vector  $\vec{A}$  pointing to the left. This would aid in our interpretation of Groups 4 and 5.

#### VI. CONCLUSIONS

The results of this work support earlier findings that students are consistent in how they approach vector addition and subtraction in an “arrows-on-a-grid” format. We find evidence for five broad groupings of mistakes that students consistently make across multiple problems: performing subtraction in the wrong order, performing addition instead of subtraction, closing the loop regardless if asked to perform addition or subtraction for anti-aligned vectors, performing subtraction instead of addition for anti-aligned vectors, and closing the loop for aligned vectors. We again emphasize that we cannot be certain without written work that students used these specific methods, but the consistency of their choices and the groupings found strongly suggest this to be the case.

#### ACKNOWLEDGMENTS

We thank Eric Brewster for discussions on methodology, and Adrienne Traxler for sharing an implementation of LANS to check ours against. This work was supported by NSF DUE 1560142. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

- 
- [1] R.D. Knight, The Vector Knowledge of Beginning Physics Students, *Phys. Teach.* **33**, 74 (1995).
- [2] N.-L. Nguyen and D.E. Meltzer, Initial understanding of vector concepts among students in introductory physics courses, *Am. J. Phys.* **71** 630 (2003).
- [3] S. Flores, S.E. Kanim and C.H. Kautz, Student use of vectors in introductory mechanics, *Am. J. Phys.* **72** 460 (2004).
- [4] P.S. Shaffer and L.C. McDermott, A research-based approach to improving student understanding of the vector nature of kinematical concepts, *Am. J. Phys.* **73**, 921 (2005).
- [5] J.M. Hawkins, J.R. Thompson, M.C. Whittman, E.C. Sayre and B.W. Frank, Students' Responses to Different Representations of a Vector Addition Question, *AIP Conf. Proc.* **1289**, pp. 169-172.
- [6] P. Barniol and G. Zavala, Vector Addition: Effect of the Context and Position of the Vectors, *AIP Conf. Proc.* **1289**, pp 73-76.
- [7] J.M. Hawkins, J.R. Thompson and M.C. Whittman, Students' Consistency of Graphical Vector Addition Method on 2-D Vector Addition Tasks, *AIP Conf. Proc.* **1179**, pp. 161-164.
- [8] T. Wang and E.C. Sayre, Maximum Likelihood Estimation (MLE) of students' understanding of vector subtraction, *AIP Conf. Proc.* **1289**, pp. 329-332.
- [9] A.F. Heckler and T.M. Scaife, Adding and subtracting vectors: The problem with the arrow representation *Phys. Rev. ST Phys. Educ. Res.* **11**, 010101 (2015).
- [10] J.B. Buncher, Algebra-Based Students and Vector Representations: Arrow vs. *ijk*, 2015 PERC Proceedings, College Park, MD, edited by A.D. Churukian, D.L. Jones, and Lin Ding, pp. 75-78, doi:10.1119/perc.2015.pr.014
- [11] J.B. Buncher, Algebra-based students & vectors: assessing physical understanding in arrow vs *ijk*, 2016 PERC Proceedings, Sacramento, CA, edited by D.L. Jones, L. Ding, and A. Traxler, pp. 56-59, doi:10.1119/perc.2016.pr.009.
- [12] B.D. Mikula and A.F. Heckler, Framework and implementation for improving physics essential skills via computer-based practice: Vector math, *Phys. Rev. Phys. Educ. Res.* **13**, 010122 (2017)
- [13] H.D. Young, P.W. Adams and R.J. Chastain, *College Physics, 10th Ed.* (Pearson, 2016); D.C. Giancoli, *Physics: Principles with Applications 7th Ed.*, (Pearson, 2014); R. Knight, B. Jones and S. Field, *College Physics 2nd Ed.* (Pearson, 2010); N. Giordano *College Physics 2nd Ed.*, (Brooks/Cole Cengage, 2013).
- [14] E. Mazur, *Peer Instruction* (Addison-Wesley, 1997)
- [15] E. Brewe, J. Brunn, and I.G. Bearden, Using module analysis for multiple choice responses: A new method applied to Force Concept Inventory data, *Phys. Rev. PER.* **12**, 020131 (2016)
- [16] T. I. Smith, P. Eaton, S. W. Brahmia, A. Olsho, A. Boudreaux, C. DePalma, V. LaSasso, C. Whitener, and S. Straguzzi, Using psychometric tools as a window into students' quantitative reasoning in introductory physics, 2019 PERC Proceedings [Provo, UT, July 24-25, 2019], edited by Y. Cao, S. Wolf, and M.B. Bennett
- [17] J. Wells, R. Henderson, A. Traxler, P. Miller, and J. Stewart, Exploring the structure of misconceptions in the Force and Motion Conceptual Evaluation with modified module analysis, *Phys. Rev. PER.* **16**, 010121 (2020)
- [18] J. Yang, J. Wells, R. Henderson, E. Christman, G. Stewart, and J. Stewart, Extending modified module analysis to include correct responses: Analysis of the Force Concept Inventory, *Phys. Rev. PER.* **16** 010124 (2020)
- [19] J. Wells, R. Henderson, J. Stewart, G. Stewart, J. Yang, and A. Traxler, Exploring the structure of misconceptions in the Force Concept Inventory with modified module analysis, *Phys. Rev. PER.* **15**, 020122 (2019)
- [20] N.J. Foti, J.M. Hughes, and D.N. Rockmore, Nonparametric Sparsification of Complex Multiscale Networks, *PLOS ONE* **6**, 1, (2011)
- [21] D. Elder, A. Eriksson, and M. Rosvall, The MapEquation software package, available online at <http://www.mapequation.org> (last accessed May 2020)