Central distractors in Force Concept Inventory data

Terry F. Scott and Dániel Schumayer*

Department of Physics, University of Otago, Dunedin 9016, New Zealand

(Received 12 October 2017; published 30 January 2018)

The Force Concept Inventory was designed to poll the Newtonian conception of force. While there are many in-depth studies analyzing response data that look at the structure of the correct answers, we believe that the incorrect answers also carry revealing information about the students' worldview. The inventory was originally designed so that the "distractors" in each question reflected commonly held misconceptions, and thus the rate at which students guess the correct answer is very low. Students select incorrect answers that correspond to the misconception that they hold and there are very few responses which appear obviously wrong to students. A side effect of this approach is that the incorrect responses reflect the non-Newtonian worldviews held by students. These non-Newtonian worldviews are coherent and robust, and this, at least in part, helps to explain why these misconceptions are so resistant to instruction. In this study we focus once more on the misconception data in Force Concept Inventory responses, particularly on the linkages between these misconceptions. We hypothesize that there are distinct groupings of distractor items formed by the strength of the association between these items. The two largest groupings are associated with the "impetus" world view in which the motion of an object is determined by the quantity of impetus which that object contains. We find that certain central items hold particularly important places in these groupings and also that individual groupings may be connected to each other by "connector" items. We finally suggest that, on the basis of this study, that these non-Newtonian worldviews might best be dismantled by addressing these key central and connector items.

DOI: 10.1103/PhysRevPhysEducRes.14.010106

I. INTRODUCTION

Newtonian mechanics is not a natural model of motion in the sense that most people do not arrive at this model of the world on their own, and there are a number of other models of motion which people seem to favor over the Newtonian model. Furthermore, there is a significant body of research [1–4] that indicates that particular non-Newtonian models of motion are extremely resistant to attempts at correction. Rigidity against modifying or replacing models is not unique to physics and has been observed, e.g., in medical sciences [5], economics [6], and engineering [7].

In a previous paper [8] we investigated the structure of the misconceptions that students display in their responses to the Force Concept Inventory (FCI) [9–11]. This research involved performing an exploratory factor analysis on FCI data that had been recorded so that each item in the five option multiple-choice questions in this survey was treated as one of five individual true or false questions. With this coding scheme we were able to retain information about misconceptions in the data.

Factor analysis is a latent trait method and models the response data as being produced by a set of underlying latent traits in the students. This means that our exploratory factor analysis postulated the existence of coherent cognitive structures in the minds of our students. The correlations apparent in the response data are treated as evidence for the existence of these latent traits.

The factor analysis indicated six underlying factors which could explain the correlations between the incorrect options chosen by students. Two of these factors anticorrelated with the correct answers; i.e., the possession of these two underlying cognitive structures would reduce the likelihood of answering questions relating to kinematics and dynamics correctly. Our research also indicated that these cognitive structures were coherent in the sense that the incorrect options were well correlated with each other in specific groups. Thus answering one question incorrectly increased the likelihood of answering several other questions incorrectly with a specific sequence of incorrect answers.

Below we take a different approach. We use network analysis and look at the relationships between the responses themselves without considering the underlying structures that make students susceptible to respond in a particular way.

daniel.schumayer@otago.ac.nz

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Cluster analysis (and perhaps more broadly network analysis) has been previously carried out on misconception data [12]. These analyses attempt to capture the complex situation of classroom learning by focusing on the social aspect of learning and thus analyzed social connections in the network of students [13,14]. Recently Grunspan *et al.* have published [15] a review of the use of network theory within education research, including social networks among students and educators. Our focus below is, however, on the connections between inventory items, rather than amongst the individuals producing them.

Brewe *et al.*'s [12] study introduced a novel method, Module Analysis for Multiple Choice Responses (MAMCR), to identify the structure of response patterns in conceptual inventories. Quoting the authors:

Clearly, many different concepts are possibly represented by each of the non-normative responses. This means that we cannot take any one response to represent a single concept as the student thinks about it. When students select one of these responses, we view this as partial evidence of their thinking [which] has to be linked with other responses to understand students' conceptions. We claim that using MAMCR on the full student non-normative responses will allow us to more fully understand students' conceptions by considering the response pattern from non-normative responses.

We agree with the authors that selecting a particular incorrect (non-normative) item might poll several misconceptions in the student's understanding and, therefore, the item cannot be exclusively labeled by a particular misconception. However, we believe that the web of incorrect items, like a set of questions loading onto a factor in factor analysis, can give a good indication of the dominant or at least the strong misconceptions and their interaction. Our main motivation with this network analysis is not only to deepen the understanding of the structures present in the FCI, but also to find possible weak spots in those cognitive structures that impede student learning of the Newtonian worldview.

We, however, modify this approach by selecting only those items which appear in a given factor. Our approach looks at the structure of relationships between items, and also assumes the existence of an underlying latent trait. We hope to investigate the structure of the underlying traits, rather than simply their relationship with the responses they produce. This modification allows us to investigate the structure of the non-Newtonian worldview underlying the misconceptions and to suggest a particular approach to dismantling these cognitive structures.

As an example of the techniques employed, we show the network of the FCI questions and their relationships in Fig. 1. The strength of connection between two questions is measured by Cramér's V statistics [16] keeping all five



FIG. 1. The network of all FCI questions is shown. The strength of connection between questions is based on Cramér's V statistics, with a minimal threshold of 0.05. The thicker the lines, the stronger the association between questions.

possible choices separate [17]. The network of FCI items is extremely complex, as Fig. 1 indicates, even though here we have only depicted the network of 30 questions and not the network of 150 items. This complexity necessitates some numerical measure of the importance of an item, based on its place in this network. Further simplification of the network is also achieved by focusing on smaller subnetworks associated with the factors found in our earlier factor analysis of the misconception correlation matrix.

It is apparent in Fig. 1 that there are "central" questions which have more and stronger connections than others, despite the nonvanishing connection between any two questions. Although Fig. 1 is indeed expressive, it is difficult to interpret as the correct and incorrect items all contribute to the strengths of connections. As we wish to focus on misconceptions, we apply the network analysis to individual items (i.e., each option in a question) and usually remove the correct, Newtonian items. This will be discussed in more detail in Sec. II.

With the help of some quantitative measures (see centrality measures in Sec. III A) we hope to show that in the students' worldview one may find some concepts deeply embedded and others loosely attached to the network of misconceptions. If this hypothesis can be confirmed then one might focus on weakening these firmly set, but incorrect notions. Another possibility is that there is a "cognitive patchwork" in students' minds where a patch represents a firm non-Newtonian belief, but these patches are loosely connected. If one can break any linkages between the patches, it would help fragmenting the incorrect worldview and thereby help confronting the student with their incoherent worldview.

In this study we have combined techniques from network analysis with our earlier factor analysis of FCI distractors. We calculate network metrics for the incorrect items in the two largest factors found in this earlier analysis.

II. STATISTICAL ANALYSES OF DISTRACTOR DATA

A. Data collection and coding

In an earlier work we have carried out a factor analysis [18] of FCI data gathered in 2008 and 2009. Briefly, that data set was collected over two years from a large (about 1500 students per year), algebra-based, physics service course, after delivering the mechanics section of the course. The students in this course have wide variety of academic backgrounds and abilities. The FCI was presented to them via the Blackboard online course management system [19] and students were required to view the FCI as part of their internal assessment. Careful consideration was given to the quality of data collected in this way and obviously frivolous attempts at the survey were removed from the sample. A more detailed description of the data collection may be found in earlier papers [8,18,20,21].

The data set described above was recoded to preserve information relating to misconceptions. The idea is that instead of treating the FCI as containing thirty questions, each of which has five options, we recast the FCI as containing 150 items which may or may not be selected by students. In other words, we treat the FCI as containing 150 true or false questions. Naturally these 150 items are organized into groups of 5 corresponding to each question, only one of which may be selected. This fact leads to two major problems. First, the five items in each question are naturally perfectly anticorrelated with each other. This problem is mitigated by removing these inherent anticorrelations from the correlation matrix. The second problem is more difficult to resolve. The five items in each question are also linearly dependent. This results in a correlation matrix which is not positive definite and is thus not well formed and cannot be used for factor analysis. The solutions to these two problems are presented in earlier work [8] and will not be discussed here. For comparison, however, we use the same correlation matrix in this analysis as was used in our previous work.

The factor analysis showed six factors in the entire data set. Two of these factors represent general cognitive biases in students, in that they display the tendency of students to pick either option A or option C if they do not know the answer to a question. The two smallest factors were entirely composed of incorrect items and appeared to be related to cognitive structures which formed early in the learning process. This was indicated by incorrect responses to a particularly straightforward and well-known situation. We will not investigate these factors in this paper. The two largest factors contained a large number of incorrect items and all of the correct items. One factor, which we will refer to as *factor 1*, contained 21 of the correct items and 25 of the incorrect items. The second factor, which we will refer to as *factor 2*, contained 9 correct and 33 incorrect items. We note here that there were 3 correct items which did not appear in our analysis [22].

Broadly speaking, these two factors represented a coherent non-Newtonian worldview. They both contained a large number of items relating to an impetus conception of dynamics and we therefore consider that these items are the response of students with a coherent, robust impetus based worldview. Ultimately the strong anticorrelation between these factors and the correct Newtonian items indicates that it is these factors which are impeding the development of a Newtonian worldview in the early stages of a student's study of physics.

Here we look into the detailed correlation structure of these two factors. For ease of reference the items (both correct and incorrect) which make up these two factors, as found in Ref. [21], are listed in Table I. Our purpose in investigating this structure is ultimately to propose a teaching methodology which effectively moves students from an impetus view to the Newtonian view. Before discussing this analysis, we will briefly outline the main characteristics of the impetus worldview.

B. The impetus worldview

The factors that we analyze in this section indicate possession of an impetus worldview. This assertion is argued for in previous work and is based on the categories provided by the original authors of the FCI [23]. Before proceeding with the analysis of these factors we will briefly outline the main defining characteristics of this worldview.

TABLE I. The table lists the membership of items in factors 1 and 2 as found in Ref. [21]. The items are sorted here in the order of their appearance in the FCI test. Correct items are indicated with an asterisk (*).

	Fact	or 1		Factor 2					
*3C	*11D	18D	24E	1B	*8B	15D	24D		
4A	*12B	18E	25C	*1C	8D	15E	26C		
*4E	12C	19D	25D	3D	*10A	*16A	27D		
*5B	13B	*19E	*25E	4B	10B	16B	27E		
6A	13C	20C	27A	4D	10E	16D	28B		
7A	*13D	*20D	*27C	*6B	11B	17C	29A		
8A	14A	*21E	28D	6C	*12B	18C	*29B		
9B	*14D	*22B	*28E	6D	13E	19B	29C		
*9E	16C	22D	*30C	*7B	14E	20B	30D		
*10A	17A	*23B	30E	7D	15B	*24A			
10D	*17B	23D		7E	15C	24B			
11C	*18B	*24A							

The history of the development of the Newtonian worldview extends from the ancient Greek natural philosophers up until Newton. This development was an integral part of the evolution of the scientific perspective. A good introduction to this history is provided by a number of monographs devoted to this subject, for example, Refs. [4,24]. We will here do little more than point out that the development of the impetus worldview occurred towards the end of this process, first appearing in the western intellectual tradition in the writings of Jean Buridan in the mid-14th century. The details of the particular version of the impetus view held by various philosophers in this period will not interest us. The point of noting the historical place of this worldview is to indicate that students who hold this view are in no way backward. This is a sophisticated, late development which required significant insight on the part of the top tier of intellectual thought in Europe in the late middle ages and early modern period.

It is important to note that the impetus worldview presented by students is not informed by the historical idea or the history of physics. The students who display this idea are forming a worldview that is based on the cultural background in which they grew up. The impetus worldview appears to be a natural first model of motion for students. This fact is, to our knowledge, currently unexplained and would seem to warrant further investigation.

With these caveats in place, we will now identify two of the central components of the impetus worldview. We will compare these characteristics with the corresponding components of the Newtonian view. First, the motion of an object is caused by a property of that object, which we call for convenience, the impetus. Impetus is a property of an object in the same way that the mass or color of an object is a property. However, the impetus is not seen as a static, fixed property of an object. Impetus may change, the quantity of impetus which an object possesses may increase or decrease and the exact amount of impetus that an object contains indicates the magnitude of motion that the object will display. Thus the impetus of an object is similar in structure to modern ideas like momentum or energy. It is important, particularly for students, to recognise the difference between the impetus idea and these modern concepts. Energy and momentum are descriptive notions, momentum, in particular, was introduced by Newton to measure the "quantity of motion" of an object. On the other hand, the impetus of an object was supposed to be causative; i.e., the motion of an object is caused by the impetus that object contains, and the impetus is depleted as the object moves.

Second, in the impetus worldview, motion in and of itself requires a cause and thus an explanation. This is the premise that leads to the notion of impetus as the cause of motion.

The most obvious difference between the Newtonian view and the impetus view is that the Newtonian view posits the existence of forces, which are interactions between objects rather than a property of the objects themselves. This idea is encapsulated in Newton's third law, which along with the second law, states that momentum is transferred between objects via Newtonian forces. The second law relates the concept of force to the rate of change of momentum and the third law shows that the momentum lost by one object is gained by the other; i.e., momentum is conserved in these interactions. The Newtonian conception of force is analogous (and closely related to) the modern concept of work which quantifies the transfer of energy between two objects.

In the impetus worldview, impetus may be gained by an object by interaction with another object with a large quantity of impetus. This transfer of impetus will then cause an increase in the motion of the receiving object. This is not the only way that an object is able to gain impetus, for example, vertical impetus may be gained simply due to a change in height. The impetus view also implies that an object will lose impetus simply by moving and thus all moving objects will slow down—even in the absence of any interactions with any other objects. It is also clear that an object with no impetus will not move at all.

The central underpinning premise of the impetus worldview is the Aristotelian view that the natural state of an object is to be stationary. Thus the raw fact of motion is an indication that an object contains impetus. In the Newtonian worldview, motion in and of itself does not require an explanation. In other words, the natural state of an object is constant, rectilinear motion. Newton (among others) recognized that this was the central consequence of Galilean relativity.

III. NETWORK ANALYSIS

Network analysis is a technique for investigating and presenting the correlation structure of a group of items. The correlation between two items is a measure of association, meaning that strongly correlated items are grouped together. Naturally each item is correlated with a number of other items, each of which is correlated with a number of other items, etc. In the normal procedure this structure of association between the items is represented as a twodimensional array of numbers, i.e., as a correlation matrix. While this representation of connectedness is traditional and is a perfectly acceptable way of representing this structure, it is not necessarily the most useful way to see patterns of relationship.

In factor analysis, the interpretation of factors is based on the magnitude of the factor loading of the items. In an orthogonal factor analysis, the factors are not correlated with each other and the factor loading gives the correlation between that item and the underlying factor. Thus the larger the factor loading the more strongly that item is correlated with the underlying factor. This means that the interpretation of the meaning of a factor should depend most strongly on those items with the highest factor loadings. It was on this basis that we interpreted factors 1 and 2 as both representing the impetus worldview.

In the network analysis approach we have a different mechanism at our disposal. We are no longer able to talk about correlations between items and the underlying cognitive structures which cause that response. Instead, we are confined to discussing the correlations between the items themselves.

We note here a significant difference between our approach and that of Brewe *et al.* [12]. Our assumption is that there are global misconceptions, i.e., misconceptions held by an entire group of students, such as the impetus worldview, and that these misconceptions may form a selfsupporting structure within the students' thinking about physics problems. For this reason we analyzed either the network of the 150 items, or the 120 incorrect items, or the 30 correct items. Brewe *et al.*, looked at the set of responses as a bipartite network, where one of the subgraphs corresponds to the responses, while the other contains the students.

While we find this flexibility attractive, we were unable to use this approach due to the nature of our data collection procedures and, therefore, our data set.

A. Centrality of items

In the present study we rely on ideas from network theory which we introduce here very briefly. These ideas enable us to determine which items are the most important and provide the most structural support for the impetus worldview.

The core concept of graph theory is *relationship* between the elements of a predefined set, e.g., the set of items in the FCI test. These elements and their relationships form a graph.

Traditionally a graph is represented by listing a pair of elements which are related to each other. This representation is simple as it describes the entire graph via binary relationships. A graph can also be represented visually by depicting the elements as points (*nodes*) and relationships as *edges* between nodes. A third representation of a network is via its *adjacency matrix*, A, whose (i, j)th element is either 1 or 0 depending on whether the nodes *i* and *j* are connected directly or not. However, the matrix elements (*weights*) may represent some other quality than the existence of the connections, e.g., strength. The weights can be arbitrary as long as the value zero represents the absence of a connection.

Naturally, we wish to find some simple structure or organizing principle for a graph and the matrix representation is exceptionally useful for extracting such features of a network. As in most cases in network analysis we are interested in nodes which are, in some sense, central to the network, e.g., they have the highest number of connections, or they have the strongest connections, or their position is vital to traverse from one side of a graph to the other, etc. In the current study we focus on the most central items as these represent misconceptions which are critical in preventing the formation of a coherent Newtonian worldview.

In order to classify nodes we employ some widely accepted centrality measures: degree, closeness, betweenness centrality. These measures rely on two fundamental concepts of graphs, distance and shortest path. In a nonweighted graph the distance of a given path is simply the number of edges participating in the path. Between two nodes there could be several different paths and the one with minimal distance will be the shortest path between the nodes and the corresponding distance is called the distance of these nodes. In weighted graphs each edge has a weight which measures its importance or strength, and then the distance is the minimal sum of weights of edges which leads from one node to the other. The two measures agree for graphs where the nodal strengths are either 1 or 0, however, they usually lead to different results in general cases.

In our analysis we use the tetrachoric correlation matrix, C, and as such its entries, c_{ij} , measure the strength of association of two FCI items. Here we need to note three consequences of this choice.

- (a) A valid weight is given for any pair of items (n = 150), thus the FCI graph is a complete graph, i.e., every node is connected to every other node, the number of connections is $\frac{1}{2}n(n-1) = 11175$.
- (b) Any entry, c_{ij} , can be either negative or positive. The absolute value determines the strength of relationship between two items and not the sign of the correlation.
- (c) Traditionally weights are considered to be the strength of edges and not the association of its end nodes. High correlation values represent strong association of items, thus we expect these items to be "close" to each other, i.e., their distance to be small whereas items with low correlation values are weakly associated with each other and we wish to represent them being "far" from each other. Moreover, a correlation of zero between two items would require that these items were connected by an infinitely long edge. Therefore, it seems natural to define the distance of two adjacent nodes *i* and *j* as $d(i \rightarrow j) = d(j \rightarrow i) = 1/c_{ij}$. Indeed, taking the reciprocal of edge weights has been proposed in the literature of network analysis [25,26].

We illustrate the following centrality measures in a small, artificial network shown in Fig. 2 and we list the numerical values of these measures in Table II. In this network each edge has unit distance, i.e., the edges are not weighted.

Degree centrality is perhaps the simplest measure, as it measures the average distance from a fixed node to its neighbors:

$$\nu_d(i) = \frac{1}{n-1} \sum_{j \in \text{neighbors of } i} d(i \to j).$$



FIG. 2. A small, undirected, nonweighted graph for demonstrating selected centrality measures. It is apparent that node 4 must score high on several centrality measures as it has the most connections. However, nodes 6 and 7 are also important as they form the single link connecting two subnetworks $(\{1, 2, 3, 4, 5, 6\}$ and $\{7, 8, 9\})$ together. Their betweenness score should capture this property.

The sum represents the total distance of node *i* from its neighbors and the prefactor is used for normalization as it simply divides by the maximum number of neighbors, (n-1). In Fig. 2 the total number of nodes is 9. Node 1 has only a single neighbor, thus its degree centrality is $\frac{1}{8} = 0.125$. Nodes 4 and 6, however, each have four neighbors, thus their degree centrality is $\frac{4}{8} = 0.5$.

Closeness centrality measures how close, to a particular node, other nodes are. The formal definition of closeness is



FIG. 4. The network structure of factor 2 items is plotted in a circular distribution. The clockwise ordering of items represents their natural order in the inventory. The stronger the correlation the thicker the connecting line, while the colors black and blue express positive and negative correlations, respectively. For sake of transparency connections with absolute value less than 0.35 are omitted. Item 7D seems to be the central item as it is connected to several other items. Some items, e.g., 8D, seem to hang alone meaning that their connection to factor 2 is quite weak.



FIG. 3. The network structure of factor 1 items is plotted in a circular distribution. The stronger the correlation the thicker the connecting line, and the colors black and blue represent positive and negative correlations, respectively. For sake of transparency, connections with absolute value less than 0.1 are omitted. While there are connections with strong correlation they only seem to form binary groups, e.g., 6A and 7A, rather than larger cliques. There are, perhaps, two items (10D and 30E), which could play some central role and provide cohesive force for the other items. Otherwise this circular graph seems to be connected in a balanced manner, i.e., lots of items are connected relatively weakly to lots of other items.



FIG. 5. The network structure of all thirty correct, Newtonian items is plotted in a circular distribution. The clockwise ordering of items represents their natural order in the inventory. The stronger the correlation the thicker the connecting line. For sake of transparency, connections with absolute value less than 0.1 are omitted.

TABLE II. Four measures of centrality, $(\nu_d, \nu_c, \text{ and } \nu_b)$, are listed for the graph in Fig. 2. Highest values are underlined.

Node	ν_d	ν_c	ν_b	Node	ν_d	ν_c	ν_b
1	0.125	0.333	0.000	6	0.500	0.615	0.548
2	0.500	0.471	0.262	7	0.375	0.500	0.429
3	0.375	0.533	0.095	8	0.250	0.364	0.000
4	0.500	0.571	0.107	9	0.250	0.364	0.000
5	0.375	0.533	0.095				

$$\nu_c(i) = \frac{n-1}{\sum_{j \in \text{all nodes}} d(i \to j)}.$$

This difference between degree and closeness centrality also highlights that degree centrality is a local measure while closeness is a global measure of a node's importance in the graph, since one needs to calculate a node's distance from *all* other nodes. For example, node 1 in Fig. 2 has unit distance from node 2, it is at distance 2 from nodes 3, 4, and 5, distance 3 from node 6, distance 4 from node 7, and distance 5 from node 8 and 9. Therefore, the closeness centrality measure of node 1 is

$$\nu_c(1) = \frac{9-1}{1+2+2+2+3+4+5+5} = \frac{8}{24} = \frac{1}{3}.$$

On the other hand, if one calculates the same measure for node 6, one finds

$$\nu_c(6) = \frac{9-1}{3+2+1+1+1+1+2+2} = \frac{8}{13} \approx 0.615.$$

These two centrality measures emphasize the distance between two nodes, and neglect the topology of the graph. A node can be important in a network by simply connecting two subnetworks together weakly, even though this particular node can be on the periphery on both subnetworks. *Betweenness* centrality captures this topological influence exactly [27–29]:

$$\nu_b(i) = \frac{2}{(n-1)(n-2)} \sum_{j,k \in \text{ all nodes}} \frac{p(j \to i \to k)}{p(j \to k)}$$

Here, $p(j \rightarrow k)$ is the number of shortest paths between nodes *j* and $k(j, k \neq i)$, while $p(j \rightarrow i \rightarrow k)$ is the number of shortest paths from *j* to *k* that also contain *i*. We can think of this measure as the extent to which a node lies between any other nodes in the graph. The normalizing prefactor contains (n-1)(n-2), which is the number of possible edges that do not include node *i*, while the numerator 2 expresses the fact that in an undirected graph if a path exists between two nodes then the path can be traversed backwards as well. For betweenness centrality we do not include an explicit calculation here due to the relatively large number of terms involved in the double summation.

B. Factor 1 items

In this section we investigate the network properties of the items that formed factor 1 in our earlier factor analysis. Thus, we are looking at the correlation structure between the responses produced by a particular underlying trait. In our earlier paper we interpreted this factor as representing the "impetus worldview" discussed above.

In Table III we list the incorrect items in factor 1 and depict the network of items in Fig. 3. The loading of these items into the underlying factor, and several centrality measures for each item.

As we would expect there is a noticeable relationship between the network analysis measures and the factor loadings, as both analyses are attempts to organize and understand the correlation structure of the FCI data.

In factor 1 the item with the largest closeness is item 11C, this item also has the highest betweenness and the largest factor loading. The large factor loading means that possession of the underlying latent trait (i.e., the impetus worldview) makes this item the most likely response by the student. The closeness rating of this item, while not surprising, does give us extra information. It tells us that this item is also very tightly tied to a number of other items. If we were to convince a student that item 11C was incorrect, then it is natural to suppose that they would be likely to change their response to a number of other items which are correlated with item 11C. The centrality of item 11C gives an indication of the number of other items effected by such a change. The fact that item 11C has the highest centrality indicates that the number of items effected in this way is higher than for any other item, and thus that this item is the most important structural support of the impetus worldview.

TABLE III. Factor loading (L) and degree, closeness, and betweenness centrality measures for factor 1 items. The values of ν_c are multiplied by a constant factor of 10^3 .

Item	L	ν_d	ν_c	ν_b	Item	L	$ u_d$	ν_c	ν_b
4A	0.412	3.64	6.37	0	18D	0.346	1.99	5.62	0
6A	0.361	4.05	7.29	11	18E	0.370	3.42	6.42	0
7A	0.304	3.08	6.26	0	19D	0.328	2.69	5.50	1
8A	0.320	3.14	6.06	6	20C	0.293	2.36	5.53	0
9B	0.333	3.16	6.24	7	22D	0.342	3.20	6.24	0
10D	0.428	4.63	8.11	9	23D	0.308	3.12	6.30	0
11C	0.522	5.50	9.24	48	24E	0.315	3.47	6.87	2
12C	0.378	3.56	6.47	0	25D	0.314	2.07	5.29	0
13B	0.384	3.58	6.70	0	25E	0.336	2.80	6.18	0
13C	0.349	2.56	6.10	0	27A	0.328	1.91	5.06	0
14A	0.346	4.29	7.46	9	28D	0.432	3.84	6.83	7
16C	0.384	3.90	7.13	6	30E	0.484	4.48	7.55	12
17A	0.411	3.70	6.52	9					

The betweenness of this item is also very high, in fact it is 4 times larger than the next highest betweenness. The betweenness can be thought of as quantifying the degree to which an item connects subgroups of items within the factor together. In other words, if this item is removed from the factor by convincing a student that the item is incorrect, then the structural integrity of the latent trait is significantly reduced and the factor may even split into two or more smaller factors.

The fact that the betweenness and closeness of item 11C are both high indicates that this is a particularly important item as it not only has a large number of other items associated with it, it is also important in connecting separate subgroups of items in this factor. Briefly, item 11C refers to a situation in which a hockey puck is sliding along a frictionless horizontal surface at constant velocity. The puck receives a "swift horizontal kick" in a direction perpendicular to its velocity. Students are then asked a number of questions related to the motion of the puck after the kick, question 11 asks them to identify the forces acting on the puck after the kick has been received. Item 11C states that these forces are "a downward force of gravity, an upward force exerted by the surface, and a horizontal force in the direction of motion."

Item 11C is clearly affirming the impetus worldview. It is non-Newtonian in that it asserts the existence of a force in the direction of motion when there is no such force. This amounts to the assertion that there is a force in the direction of motion *because* the object is moving in that direction. This item is a direct contradiction of Newton's first law and affirms one of the defining characteristics of the impetus worldview, namely, that velocity requires a cause.

The fact that item 11C is the item which is structurally at the center of this factor tells us that this idea underpins the impetus worldview. This, in turn, suggests that the belief that motion requires a cause, and thus an explanation, is the structural centre of the impetus worldview. It is the denial of Newton's first law, rather than the rejection of the second and third laws, that ultimately leads to the impetus worldview.

The next highest factor by closeness is item 10D, the closeness of this item is only slightly lower than the closeness of item 11C. However, the betweenness of item 10D is significantly lower than the betweenness of item 11C. This means that this item is closely associated with a reasonably large number of other items but it is not so important as a connection between subgroups. Item 10D is closely associated with item 11C, it is part of the same question group, so we would prima facie expect that it would be of similar importance as item 11C. It also asserts an impetus worldview, but it is not so clearly a rejection of Newton's first law. Rather it appears to indicate a belief about the way that impetus works to produce motion. In this case it asserts that the delivery of impetus to an object will cause the object's velocity to increase and then decrease.

Item 30E is the next highest item by closeness and is the second highest item by betweenness. But note that the betweenness of this item is still less than a quarter of the betweenness of item 11C. This item is an important structural element in this factor both because it is closely connected to a large number of other items and also because it is important in connecting subgroups of items. Item 30E is quite similar to item 11C in that it asserts that there is a force present which is in the direction of motion, in this case the force supplied by a hit. While not so clearly a denial of the first law it is clearly an impetus view.

Factor 1 was shown in a previous paper to be the largest factor, by this we mean that factor 1 explains the greatest proportion in the variation of the responses in the FCI and it, therefore, has the largest eigenvalue of any of the factors. Thus the analysis of factor 1 given above is the most important part of this paper in terms of applications to teaching. The greatest improvement in the understanding of students will be achieved by convincing them that constant velocity does not need explanation. On the basis of the analysis above we encourage instructors to emphasize the philosophical basis of Newton's first law.

C. Factor 2 items

We now analyze the second largest factor identified in our earlier work (see Fig. 4). While factor 2 is perhaps of secondary importance from the perspective of instruction, it is nonetheless interesting from the perspective of the central focus of this paper. We list the items in factor 2, along with their factor loadings and network centrality measures in Table IV.

Item 7D has the highest closeness and betweenness in this factor. This item relates to circular impetus and indicates a particular belief about the effect of circular impetus over time, specifically it implies that circular impetus will take effect after a short time lag. It is not particularly surprising that this item appears in factor 2 as it is clearly a derived notion. It is not central to the impetus

TABLE IV. Degree, closeness, and betweenness centrality measures for factor 2 items. The values of ν_c are multiplied by a constant factor of 10^3 .

Item	ν_d	ν_c	ν_b	Item	ν_d	ν_c	ν_b	Item	ν_d	ν_c	ν_b
1B	1.98	7.60	0	11B	0.40	6.36	0	19B	1.70	7.65	0
3D	4.34	8.74	1	13E	4.40	8.75	9	20B	2.91	8.73	1
4B	5.95	9.71	18	14E	7.30	10.37	25	24B	2.55	8.40	3
4D	6.48	9.62	16	15B	3.68	8.44	1	24D	1.62	7.31	1
6C	3.60	8.36	0	15C	1.69	7.52	0	26C	5.97	9.47	3
6D	3.67	8.68	5	15D	3.57	8.24	0	27D	0.91	7.12	0
7D	9.56	11.25	37	15E	2.68	7.71	0	27E	4.74	9.43	10
7E	0.82	7.35	0	16B	4.42	9.07	3	28B	7.58	10.53	5
8D	0.00	5.78	0	16D	3.00	8.90	2	29A	3.51	8.65	2
10B	0.00	6.78	0	17C	5.94	9.73	7	29C	6.51	9.64	8
10E	0.36	6.37	0	18C	6.54	9.61	9	30D	5.37	9.03	8

worldview but is, in a sense, a consequence of a the belief in impetus and the dynamics of impetus applied to circular motion. What is interesting is that this is structurally the most important item in this factor in terms of closeness and betweenness. It has a high factor loading, but the item with the highest factor loading, 28B, is not nearly as important structurally. Item 28B has a high closeness rating, indicating that there are a lot of other items clustered around it, but its betweenness is quite low, meaning that it does not connect further subgroups within factor 2. It would be extremely surprising if the item with the highest factor loading did not also have a very high closeness.

Item 28B is an item from a question about Newton's third law and asserts that the more active agent supplies the greatest force. The active agent items are often treated as though they are a more sophisticated model than the impetus view. We can consider them compatible: students are thinking about the amount of impetus supplied by a hit and the more active agent is capable of supplying a bigger hit. Notice that the item is asserting something about the force applied to the other agent, i.e., there is no problem in thinking about this being an impetus force being supplied rather than a Newtonian force.

Both of these items relate to the application of the impetus idea, whereas factor 1 contains items which refer to the central ideas of the impetus worldview. In this sense our analysis indicates that factor 2 is dependent on factor 1. Thus, dismantling factor 1 by developing Newton's first law, should, at the same time, destabilize factor 2.

However, the importance of item 7D suggests that the impetus worldview may be attacked empirically using the scenario outlined in this question. Students with a robust impetus worldview are likely to expect an incorrect trajectory in this situation. Showing students the situation and highlighting the conflict with their expectations could motivate a reconsideration of their worldview.

D. The Newtonian items

As the final piece of analysis we repeat the network analysis used above on the Newtonian (i.e., correct) items. Their network structure is shown in Fig. 5. The measures of centrality for the correct items are listed in Table V. In this case, however, we have not listed the factor loadings as these were loadings onto a non-Newtonian underlying trait. One of our earlier studies [18] gives a factor analysis of the Newtonian worldview. Table V shows quite clearly that there is a single item, item 13D, which has a very high betweenness and also a high centrality. We also note that there is a group of five more items which have high betweenness and centrality and there is a significant drop in the size of these measures outside this group. These items are items 28E, 25C, 7B, 22B, and 11D. We investigate these items in this section.

We should raise a note of caution here. Earlier we have stated that an item with high centrality and betweenness is

TABLE V. Degree, closeness, and betweenness centrality measures for all thirty correct items. The values of ν_c are multiplied by a constant factor of 10^3 .

Item	ν_d	ν_c	ν_b	Item	ν_d	ν_c	ν_b	Item	ν_d	ν_c	ν_b
1C	1.87	6.65	0	11D	9.25	11.11	10	21E	2.94	7.21	0
2A	0.56	5.82	0	12B	6.06	8.70	0	22B	7.81	9.74	11
3C	0.55	5.68	0	13D	11.10	12.46	102	23B	7.04	9.42	3
4E	3.91	7.39	3	14D	6.42	8.65	0	24A	7.86	9.89	1
5B	7.20	9.55	1	15A	1.16	6.03	0	25C	5.63	8.93	20
6B	7.71	9.55	5	16A	7.11	9.13	7	26E	0.57	5.21	0
7B	7.90	9.32	13	17B	6.07	8.75	1	27C	4.12	7.30	0
8B	5.01	7.82	0	18B	6.91	9.19	0	28E	7.14	9.32	21
9E	5.13	8.22	0	19E	7.24	8.96	0	29B	1.34	6.07	0
10A	7.66	9.78	4	20D	7.88	9.63	0	30C	7.78	9.88	0

structurally important to the cluster of items in which it appears. In that analysis we assumed that the correlation between the most central item and the items around it is due to the fact that the central item has an important function and in some sense leads to the other responses. The centrality is based on the strength of the correlation between items and we must be careful not to equate this correlation with causation, i.e., the selection of item 11C does not *cause* the selection of item 22D, but the selection of the first item by a student makes it more likely that they also select the second item. Both of these selections being caused by some underlying cognitive trait of the students. Underlying traits such as these are proposed as explanations of the response data by various latent trait models, of which factor analysis is an example.

In our analysis of the incorrect items we found that there were a number of underlying traits that explained the data reasonably well and the analysis above presented the network analysis within a single factor. This means that all of the items in these factors are caused by the degree to which a student holds the impetus worldview and the structure of the response data reflects the structure of this trait. Thus an item in factor 1 with high centrality, betweenness, and factor loading, can be interpreted as having an important structural function in the underlying trait.

The factor analysis of the incorrect items does not reveal the structure of the Newtonian trait, however, it merely shows that the impetus trait is strongly anticorrelated with Newtonian responses. The correlation structure between the Newtonian answers cannot, therefore, be related to the structure of an underlying trait. Thus, the analysis of these network measures is qualitatively different to the analysis given above.

An item that has strong connections to a number of other items could be connected in this way because it represents a fundamental concept that leads to a series of other subsidiary concepts. This is the interpretation we have postulated in the interpretation of the network structure of factors 1 and 2 above. This interpretation is supported by the fact that the items in question are relatively simple and are based on relatively self contained ideas, such as "motion requires a cause." In the case of the Newtonian items, on the other hand, we propose a qualitatively different interpretation. It is also possible that an item has a high centrality and betweenness because it is the result of the application of several complementary and mutually consistent ideas. For the Newtonian answers this appears to be the appropriate interpretation.

The most central item is 13D. This is the correct answer to a complex problem that requires the application of several Newtonian ideas, namely, an understanding that a force exists, that the force acts in a single direction, that velocity is constant and perpendicular to this force, and that there is an acceleration in the component of velocity parallel to this force. Thus, this item has high centrality and betweenness because answering this question correctly requires that the student understands all of these ideas and thus also answers other questions relating to these ideas correctly. The centrality of this item does highlight the complexity of this question, whereas without the network analysis this would appear to be a relatively straightforward question.

The other questions in this group, questions 28, 25, 7, 22, and 11 are all similarly complex questions, though perhaps not as complex as question 13. Question 11 has already been discussed. The correct answer requires correctly identifying the forces acting on the hockey puck, and this in turn requires an understanding of the first and second laws as well as the ideas of kinematics. Question 28 is a question which polls the students understanding of Newton's third law. These questions are generally among the hardest in the FCI [21,30-32] as a strong understanding of Newton's third law requires an understanding of both the first and second laws as well as an understanding of the laws of kinematics. While this question does appear at first to be straightforward, this is only because we are able to take shortcuts due to our understanding of the third law, which students may well not have access to. Question 25 is another question that requires students to identify forces.

These questions are complex in that the identification of forces requires a complete understanding of the whole family of ideas that constituted the Newtonian worldview. Question 7 is a question that is often used to identify the impetus worldview; in this question a steel ball attached to a string is swung in a circle. The string suddenly breaks and the student is asked to identify the trajectory of the ball after the break. Clearly, this is a complex question that requires input from a number of Newtonian ideas before the correct answer may be arrived at. Question 22 and 11 are similarly complex, requiring the student to identify the trajectory of an object in the first case under the influence of a continuous force and in the second case after the application of a delta function force.

Clearly these items are at the top of the pyramid of Newtonian ideas rather than supporting this pyramid. This explains their high values of centrality and betweenness.

IV. CONCLUSION

The purpose of this paper is twofold. First, we present a network analysis of the non-Newtonian impetus worldview uncovered in our previous factor analysis of FCI data. We provided another way of analyzing the correlations between the items contained in the impetus worldview. In particular, this analysis allows us to clarify which of the two main characterizing ideas of the impetus worldview is the most foundational. These two ideas are that motion, broadly conceived, requires a cause, that cause being impetus; and that the cause of motion is some quantity contained by the moving object, i.e., impetus is a property of objects. Of these two ideas, we find that the first is the most foundational and thus that the denial of Newton's first law is the most significant barrier to adopting a Newtonian worldview.

The second purpose of this paper is to propose a new way to approach the teaching of Newtonian mechanics. This proposal is that teaching Newtonian mechanics should involve the active and explicit dismantling of the impetus worldview. We have identified the most important ideas that support the impetus view so that they may be addressed first. Thus, we propose that the idea of motion should be clarified so that a clear conception of velocity is established. Once this has been achieved, it is then possible to clearly articulate Newton's first law and to show that this idea explains observations that the impetus worldview is unable to accommodate. In particular, instructors should clearly demonstrate that motion, conceptualized as a state of constant velocity, does not require a cause or an explanation. Along with this, it is important to introduce the idea of friction to explain observations that appear to run counter to Newton's first law and that would appear to justify the impetus view. Finally, Newton's second and third laws may be introduced to flesh out the idea of force.

This last proposal is based on our analysis of the structure of the impetus worldview presented in the bulk of this paper. As always, the proposed teaching method should be subjected to empirical investigation and validation.

ACKNOWLEDGMENTS

The authors acknowledge the financial support by the Department of Physics, University of Otago.

- D. Hammer, More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research, Am. J. Phys. 64, 1316 (1996).
- [2] P. Kowalski and A. K. Taylor, The effect of refuting misconceptions in the introductory psychology class, Teaching Psychology 36, 153 (2009).
- [3] A. C. Butler, L. K. Fazio, and E. J. Marsh, The hypercorrection effect persists over a week, but high-confidence errors return, Psychon. Bull. Rev. **18**, 1238 (2011).
- [4] M. Jammer, *Concepts of Force*, Dover Books on Physics (Dover Publications, New York, 2012).
- [5] M. J. Leonard, S. T. Kalinowski, and T. C. Andrews, CBE Life Sci. Educ. 13, 179 (2014).
- [6] I. Busom, C. Lopez-Mayan, and J. Panads, Students' persistent preconceptions and learning economic principles, J. Econ. Educ. 48, 74 (2017).
- [7] J. L. Newcomer, in 2015 IEEE Frontiers in Education Conference (FIE) (IEEE, New York, 2015), pp. 1–7, DOI: 10.1109/FIE.2015.7344352.
- [8] T. F. Scott and D. Schumayer, Conceptual coherence of non-Newtonian worldviews in Force Concept Inventory data, Phys. Rev. Phys. Educ. Res. 13, 010126 (2017).
- [9] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, Phys. Teach. 30, 141 (1992).
- [10] Authorized educators can request the latest FCI test from the contacts appearing on modeling.asu.edu/R&E/Research .html.
- [11] D. Hestenes and I. Halloun, Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller, Phys. Teach. 33, 502 (1995).
- [12] E. Brewe, J. Bruun, and I. G. Bearden, Using module analysis for multiple choice responses: A new method applied to Force Concept Inventory data, Phys. Rev. Phys. Educ. Res. 12, 020131 (2016).
- [13] E. Brewe, L. Kramer, and V. Sawtelle, Investigating student communities with network analysis of interactions in a physics learning center, Phys. Rev. ST Phys. Educ. Res. 8, 010101 (2012).
- [14] R. Dou, E. Brewe, J. P. Zwolak, G. Potvin, E. A. Williams, and L. H. Kramer, Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens, Phys. Rev. Phys. Educ. Res. 12, 020124 (2016).
- [15] D. Z. Grunspan, B. L. Wiggins, and S. M. Goodreau, Understanding classrooms through social network

analysis: A primer for social network analysis in education research, CBE Life Sci. Educ. **13**, 167 (2014).

- [16] H. Cramér, *Mathematical Methods of Statistics (PMS-9)* (Princeton University Press, Princeton, NJ, 1946).
- [17] Cramér's V, also known as Cramér's ϕ , is a measure of association between two nominal variables based on the χ^2 statistics calculated for their contingency table. Further details can be found in M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 3rd ed. (Hafner Publishing Co., New York, NY, 1973), Vol. 2.
- [18] T. F. Scott, D. Schumayer, and A. R. Gray, Exploratory factor analysis of a Force Concept Inventory data set, Phys. Rev. ST Phys. Educ. Res. 8, 020105 (2012).
- [19] Copyright of Blackboard Inc. (1997–2017).
- [20] T. Scott, A. Gray, and P. Yates, A controlled comparison of teaching methods in first-year university physics, Journal of the Royal Society of New Zealand 43, 88 (2013).
- [21] T. F. Scott and D. Schumayer, Students' proficiency scores within multitrait item response theory, Phys. Rev. ST Phys. Educ. Res. 11, 020134 (2015).
- [22] Three correct items loaded on both factor 1 and 2.
- [23] D. Hestenes and M. Wells, A mechanics baseline test, Phys. Teach. 30, 159 (1992).
- [24] K. Simonyi, A Cultural History of Physics (A. K. Peters/ CRC Press, Boca Raton, 2012).
- [25] M. E. J. Newman, Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, Phys. Rev. E 64, 016132 (2001).
- [26] U. Brandes, A faster algorithm for betweenness centrality, J. Math. Sociol. 25, 163 (2001).
- [27] J. M. Anthoisse, Stichting Mathematisch Centrum Technical Report No. BN 9/71 (1971).
- [28] L. C. Freeman, A set of measures of centrality based on betweenness, Sociometry 40, 35 (1977).
- [29] L. C. Freeman, Centrality in social networks conceptual clarification, Soc. Networks 1, 215 (1978).
- [30] J. Wang and L. Bao, Analyzing force concept inventory with item response theory, Am. J. Phys. 78, 1064 (2010).
- [31] G. A. Morris, N. Harshman, L. Branum-Martin, E. Mazur, T. Mzoughi, and S. D. Baker, An item response curves analysis of the Force Concept Inventory, Am. J. Phys. 80, 825 (2012).
- [32] J. Han, L. Bao, L. Chen, T. Cai, Y. Pi, S. Zhou, Y. Tu, and K. Koenig, Dividing the Force Concept Inventory into two equivalent half-length tests, Phys. Rev. ST Phys. Educ. Res. 11, 010112 (2015).