



University students' associative knowledge of history of science: Matthew effect in action?

Ismo Koponen and Maija Nousiainen

Department of Physics, University of Helsinki, Helsinki, Finland

For correspondence: ismo.koponen@helsinki.fi

Abstract

We present a method to analyse how pre-service science teachers relate events, ideas, characters and deeds in history of science and in cultural and general history. A group of 25 students presented their views they deemed to be of importance in history of science, culture, society and politics in era between 1550 and 1850. The sample is based on students' study reports and analysed by using network analysis. We show how students' knowledge of history of science and history in general are organised around certain famous characters, ideas, events and institutions, thus revealing the phenomenon of accumulation of fame; the Mathew effect in action.

Keywords: history, science, network approach, degree centrality, betweenness centrality

Introduction

Within science education, educators and researchers have traditionally put forward several motivations and rationalizations for including the history of science as part of the science curriculum, as well as part of a science teacher's education. One major motivation is to improve attitudes towards science and scientists, and the understanding of science as a practice or the role of science in society (Russell, 1981; Höttecke & Silva, 2011, Gooday et al., 2008; Galili & Hazan, 2001). In textbooks of physics, history of science is often presented through individual scientists' achievements and biographies (Leite, 2002), often providing "false idea of independence of science from the rest of the world" (Leite 2002, p. 355). Such a perspective on science history displays the history of science as achievements accomplished by remarkable few individuals and few spectacular breakthroughs. Such conception of history of science is flawed, and does not serve the purpose of providing a broader understanding of science as part of culture.

Here, an attempt is made to investigate to what extent university students conceive the history of science through individuals, ideas and events and how powerfully certain luminous individuals, ideas and events overshadow lesser characters, ideas and events. In this attempt, we first undertake the following simple task: to ask what pre-service teachers know or can find out from various sources of what happened, when and by whom. These simple facts form the basic knowledge from which more involved interpretations and deeper understandings of consequences must be fashioned. To bring some meaning and ordering to the plurality of events, ideas, and deeds in history, one needs to have some associations to start with; to this kind of knowledge we refer here as associative knowledge.

In this report, we study the associative knowledge of a group of about 25 students, as it was revealed through set of preparatory tasks to explore the science history of the three centuries between 1550 and 1850 and how that history of science was embedded in the culture, society, and politics of that same

era. Students were asked to produce pairwise (dyadic) associations, based on their personal judgements, on the relevance of given persons, ideas, inventions, and events during those centuries.

These preparatory tasks produced material which was the basis for deeper discussions in a seven-week course focusing on science history. In this study, we do not consider the course or its education outcomes and impacts, but focus only on the data produced by the preparatory task, which provides a kind of starting point. We analysed the data produced by this preparatory task by using complex network methods, first constructing network of all pairwise associations that emerge from that data, second, performing a network based cartography of that landscape. We show how such networks provide the ability to discern which persons, events, and ideas emerge as the most important and luminous items in that network. We show that the students' judgements of the most important persons, events and ideas are very strongly focused on limited number of items, with a result of so called Matthew effects (Merton, 1968; Price, 1976); fame collects fame and rich get richer. This effect is ubiquitous in science and how scientific fame and merits accumulate (Börner & Scharnhorst, 2009).

Empirical design and sample

The context of this study is a physics history course, which discussed classical physics (mechanics, thermodynamics, electricity and magnetism and electromagnetism), for pre-service physics teachers. The course aim was to introduce physics history as a part of more general science history, part of the history of the humanities and arts, and as part of general history, in expanding circles. The course was attended by 22 students, and it lasted seven weeks and consisted of two three-hour weekly contact teaching sessions.

The sample of data which is examined here comes from a preparatory task for a weekly topic. In these pre-tasks, students were asked to identify those characters, ideas, inventions, and events they thought to be of major interest or importance for the era to be discussed in more detail during the week. The total sample consists of four pre-tasks, focusing on four different historical periods, which are (roughly): 1550–1640, 1640–1730, 1730–1790, and 1790–1850. Students were asked to find information and facts by using easily available sources like textbooks, introductions to physics history (like Holton and Brush 2001 and Simony 2012) and from Wikipedia. The tasks were thus simple, not very deep and on level of very general knowledge. The goal was to collect background knowledge for further discussions. As such, the data reveals what students paid attention to and found relevant. The availability of material supposedly limited not too much the scope and topics, which became featured in samples. For example, the Wikipedia sources alone of e.g. Galileo Galilei, Isaac Newton, Rene Descartes, Scientific Revolution, Reformation etc., which featured as students' targets of interest for period 1550–1730 and to which they could retrieve, is several hundreds of pages in total. Students thus needed to do selections and decisions what to retrieve and what they decided to be relevant. Each of the historical period was treated during one week (two hours for lectures and two for group discussion) and the preparatory task was completed before that, during previous week. Students had thus seven days to complete the preparatory task (which was not graded) but they had had also other tasks (which were graded) to complete (reflective summaries based on lectures and discussions). These other tasks were based on students' preparatory tasks, lectures discussing the same era, and group discussions. Here, however, we analyse only the preparatory tasks.

In the preparatory tasks students were asked to report the connections between historical characters (scientists, philosophers, writers, artists, rulers etc.) and ideas and inventions and historical events they found relevant and important for the given era. The resulting connection they reported in form of pair-wise connections, for example {Newton ↔ gravity}, {Galilei ↔ heliocentricmodel}. On basis of such ordered pairs we constructed a network, where each key word is a node in the network and the dyadic connection a link connecting the two nodes. The resulting network, which includes all the

different keywords and connections reported by the student group of 22 students, consists of 1405 different nodes (keywords) and the 2501 different links between them. The number of links, when all the nodes and links in all the pre-tasks are collated, varies from 70 links for the most extensively connected node to only 1 link for most nodes.

As the basis of the data was in the form of pair-wise connections, derived from the sample of 22 students, we formed a network of connections that represents how the students associated the different characters, ideas, inventions, and events in the four periods that were the topics of the pre-tasks. Eventually, such networks are treated as network of associative knowledge. The associative network then provides a broad picture of how the different pieces of facts that the students perceive to be of importance form the big-picture of the given era. From that picture we can then pick out the most central items that emerge from the plethora of details (1405 nodes and 2501 links).

Methods: the cartography of students' knowledge

Historical persons, ideas, inventions, and events, even at the simplest level, are not isolated facts but a connected and contingent set of facts. The importance of notions and facts needs to be approached from a more holistic perspective, by considering how given facts and notions (in what follows referred to as elements of the network) are linked as part of the network. An obvious way is to count the connections of a given element of a network (fact or notion) to other elements. This is an often-used method in analyses of textbooks (Roseman, Stern & Koppal, 2010) or students' concept maps (Ruiz-Primo & Shavelson, 1996; McClure, Sonak & Suen, 1999). Such counting, however, ignore the global structure of the network. For this, more advanced methods of analysis are needed.

The structure of knowledge, its organization and development, can be seen from the viewpoint of complex networks (Börner & Scharnhorst, 2009, Chen et al., 2009). Such network-based approaches pay due attention to the global features of interlinked knowledge structures. To analyse the importance of a given element in the network, we introduced one local measure and one global measure to gauge the importance of the given element in the network $G(V;E)$ which has V nodes (vertexes) and E links (edges). The local measure is the Degree Centrality (DC), which is based on the number of local connections that the element has to a neighbouring element. The global measure is the Betweenness Centrality (BC), which counts how different nodes in a network can be reached through a given element (Boccaletti et al., 2006; Estrada, 2012).

The Degree Centrality DC (Boccaletti et al., 2006; Estrada, 2012) is simply the number of links (out- and ingoing) $Deg(v)$ attached to a given node v .

$$DC(v) = Deg(v)$$

The Betweenness Centrality BC (Boccaletti et al., 2006; Estrada, 2012) is the number $\sigma_{pq}(v)$ of shortest paths (smallest number of links) between nodes p and q so that node v belongs to the path, normalized with the number of all shortest paths between p and q .

$$BC(v) = \sum_{p \neq q \neq v \in V} \frac{\sigma_{pq}(v)}{\sigma_{pq}}$$

Both DC and BC are important, and to pick out the pieces of knowledge that are important in both respects, the Importance $I(v)$ of node v in the network is defined as a geometric mean.

$$I(v) = \sqrt{BC DC}$$

This kind of importance ranking allows us to identify the key nodes in each individual network (Chen et al., 2009). These methods are in principle like methods introduced earlier for the analysis of conceptual knowledge, but are different in respect to centralities of interest (Koponen & Nousiainen, 2014).

Results

Historical persons, ideas, inventions, and events are not isolated facts, but always embedded as part of a larger picture. The broadness of a particular perspective determines what details stand out from the plethora of other details, and that perspective changes when the time span under inspection is changed. We first examine the era from the Reformation (ca 1550) to the first half of the Enlightenment period (ca 1730), and then from Enlightenment to industrialization and liberalism (ca 1850). Both periods are divided into two sub-periods. The sub-periods for the first era are 1550–1642 and 1642–1730, and for the second era 1730–1789 and 1789–1850.

The era from the Reformation (ca 1550) to the Enlightenment (1730)

The first era spans roughly 180 years, starting with the middle years of the Protestant Reformation (1550) and the birth of a new science with a heliocentric worldview, and ending with the years when the era of scientific revolution had already changed enough to deserve a new periodization, often referred to as the era of Enlightenment, with the dividing line being Newton's death in 1726/7. This period was divided into two pre-task periods of 1550–1642 (the year of Galilei's death and Newton's birth) and 1642–1730.

The students' conception of the most central characters, ideas, inventions, and events in the period 1550–1730 are shown in network form in Figure 1. In that figure, the size of the nodes is proportional to the Betweenness Centrality (BC). The corresponding Importance I as the geometric mean of BC and Degree Centrality (DC), and the values of BC and DC, are all given in Table 1. We have chosen to display the BC by the size of the nodes because it is most directly related to the global connectedness of the network, while the Importance I is the most relevant measure to concentrate on when both global and local connectedness are of interest.

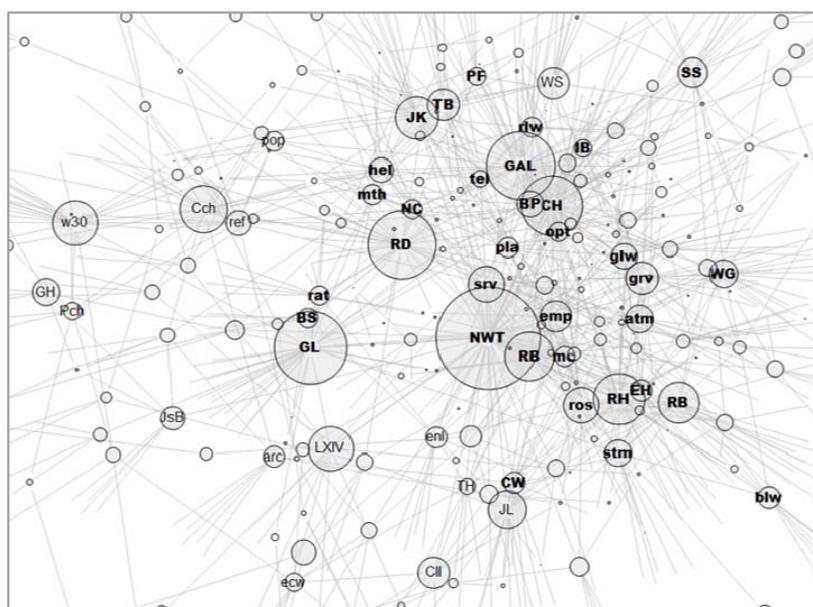


Figure 1. The network of connections between characters, ideas, innovations, and events for the years 1550–1730. The size of the node is proportional to the Betweenness Centrality. The acronyms are given in Table 1.

The top ten characters, ideas, inventions, or events for the whole era 1550–1730 produced quite a similar list, on the basis of Importance I, Betweenness Centrality BC, and Degree Centrality DC, with some reordering in rank. On the basis of Importance I, the top ten list is (see Table 1): 1. Newton, 2. Galilei, 3. Leibniz, 4. Descartes, 5. Huygens, 6. Hooke, 7. Bacon, 8. Kepler, 9. Thirty Years' War, 10. Boyle. On the basis of Betweenness Centrality BC, the top ten are: 1. Newton, 2. Leibniz, 3. Descartes, 4. Galilei, 5. Huygens, 6. Hooke, 7. Bacon, 8. Catholic Church, 9. Ludwig XIV, 10. Thirty Year's War. On the basis of Degree Centrality DC, the list reads: 1. Newton, 2. Galilei, 3. Huygens, 3. Descartes, 4. Hooke, 5. Leibniz, 6. Kepler, 7. Gravity, 8. Thirty Years' War, 9. Boyle, 10. Bacon. As expected, Isaac Newton emerges as a central figure of that era, but also Galileo Galilei, Gottfried Leibniz, and Rene Descartes were also ascribed high values of Importance. In general, the students' historical landscape is biased to persons, which are mentioned in about 70% of cases among all top ten concepts. This probably reflects the heavily biographical conception of science history, also noted in some studies by the way in which the history of science tends to be introduced in science textbooks (see e.g. Leite, 2002). Of the ideas and conceptual inventions, the most important are connected to gravitation and the law of universal gravitation, to the heliocentric model, and to pressure and steam engines. These topics are mentioned in about 30% of cases and thus gain lesser attention in comparison to individual scientists.

Table 1. The most important characters, ideas, innovations, and events for the years 1550–1730 and their Importance I, Betweenness Centrality BC, and Degree Centrality DC. Note that the last ten items 21–30 were picked up from low ranking elements and exemplify how literature, arts, and music are ranked. The acronyms used in the Figures are also provided below.

Item			I	BC	D C		Item			I	BC	DC
I. Newton	NWT		18.7	538	65		B. Pascal	BP		2.7	54	14
G. Galilei	GAL		11.5	269	49		Planetary movement	pla		2.7	35	20
G. Leibniz	GL		10.7	303	38		Mechanics	mc		2.5	35	18
R. Descartes	RD		10.4	270	40		C. Wren	CW		2.4	39	15
C. Huygens	CH		10.1	221	46		Optics	op		2.4	33	17
R. Hooke	RH		8.1	167	39		E. Halley	EH		2.4	40	14
F. Bacon	RB		6.1	155	24		N. Copernicus	NC		2.3	33	16
J. Kepler	JK		6.0	124	29		Refraction law	rlw		2.2	34	15
Thirty Years' War	w30		5.6	133	24		Boyle's law	blw		2.2	41	24
R. Boyle	RB		5.3	116	24		Mathematics	mth		2.2	36	13
Catholic church	Cch		5.2	148	18		Rationalism	rat		2.1	33	14
Gravitation	grv		4.6	79	27		Enlightenment	enl		2.1	37	12
J. Locke	JL		4.4	103	19		G.F. Handel	GH		2.0	59	7
Ludvig XIV	LXIV		4.3	133	14		I. Beeckman	IB		2.0	28	14
Royal Society	Ros		4.2	91	19		B. Spinoza	BS		1.9	34	11
Scientific revolution	Srv		4.1	94	18		London Fire	lof		1.9	30	12
T. Brahe	TB		4.0	75	22		Cambridge Univ	cun		1.9	39	9
Empiricism	emp		4.0	73	22		J.S. Bach	JsB		1.9	44	8
S. Stevin	SS		3.5	66	19		Popes	pop		1.8	34	10
W. Shakespeare	WS		3.5	76	16		Architecture	arcr		1.8	38	9
W. Gilbert	WG		3.3	61	18		Telescope	tel		1.8	26	13
Gravitation law	GrI		3.2	51	20		P. Fermat	PF		1.8	30	11
Heliocentricity	hel		3.2	50	20		T. Hobbes	THo		1.7	25	11
Reformation	Ref		3.0	51	18		English civil war	ecw		1.6	30	9
Atmospheric pressure	atm		2.9	60	14		Protestant chc.	Pch		1.6	27	9
Steam engine	stm		2.9	55	15		Restoration	res		1.4	34	6
Charles II	CII		2.8	73	11		Opera	opr		1.3	32	5

Some general historical events are also placed quite high in importance. For example, the Thirty Years' War of the Reformation era is of high importance, as is the Catholic Church and the Reformation itself. Interestingly, this is probably a reflection of the central role of these topic in the Finnish upper secondary school curriculum. In addition, "Le Roi Soleil", Ludvig XIV, is of high

importance in the students' historical landscape of 1550–1730. Some institutions are also noted as central, among them the Royal Society.

The cluster of philosophies and philosophers is also interesting, since in addition to Leibniz (who can also be counted as a philosopher) and Descartes, Francis Bacon and John Locke also receive substantial attention. Connected to these philosophers, empiricism and rationalism are also well represented in the students' historical landscape. Somewhat disappointingly, however, the visibility of literature and arts in the students' history landscape for 1550–1730 is quite limited. In practice, only William Shakespeare in literature and Johan Sebastian Bach and Georg Friedrich Händel in music stand out.

The details of the history landscape change, of course, if the time span is changed. If the time span of 1550–1730 is divided roughly into the periods 1550–1640 (ending with Galilei and the Reformation era) and 1640–1730 (the Baroque period, ending with Newton), the details of the picture are somewhat different. For 1550–1640, the key characters in terms of Importance are now: Galilei (8.8), Descartes, (6.6), Bacon (4.8), Kepler (4.6) and Brahe (3.7). The most important historical events are rather high in importance: the Thirty Year's War (3.4) and the Reformation (2.8). In this period, Shakespeare (3.3) competes in importance with the scientists Huygens (3.3), Stevin (2.9), and Gilbert (2.7). Of institutions, the Royal Society (2.4) and the Catholic Church (2.8) receive lot of attention.

For the period 1640–1730 the most important scientists, according to students' history landscape, are: Newton (13.3), Leibniz (8.2), Hooke (7.1), Huygens (5.2) and Boyle (3.6). In importance, Boyle is immediately followed by Ludwig XIV (3.4). The scientific ideas and themes that clearly outweigh any other are gravity (3.2) and gravitation law (2.5). Interestingly, at the top of the list one finds Locke (2.4) and Wren (2.4), both competing for attention with the steam engine (2.4), while next on the list one finds Charles II (2.0), Halley, and the London Fire (both 1.9).

The ordering of ranks in the sub-periods is thus different in its details, but in line with the analysis based on the whole period from 1550–1730. The visualization of the connection within the sub-periods with nodes scaled to the size of BC is shown in Figure 2, to be compared with Figure 1. Note that the positions of the nodes change because of changes in connections, and the visualization algorithm tends to cluster the nodes according to their connectivity. The acronyms are as in Table 1.

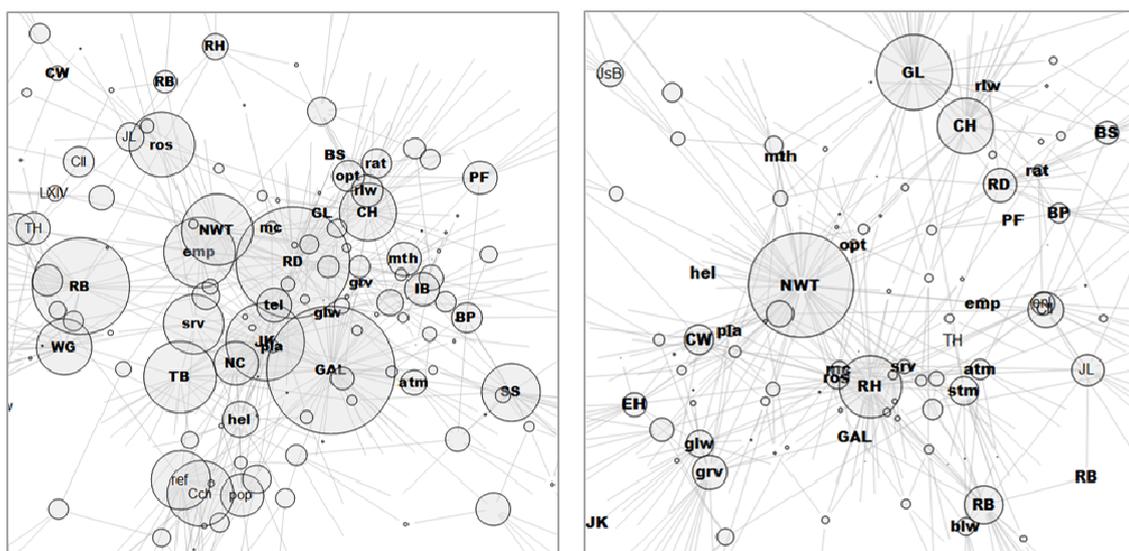


Figure 2. The network of connections between characters, ideas, innovations, and events for the sub-periods 1550–1640 (left) and 1640–1730 (right) for the years 1550–1730. The size of the node is proportional to the Betweenness Centrality.

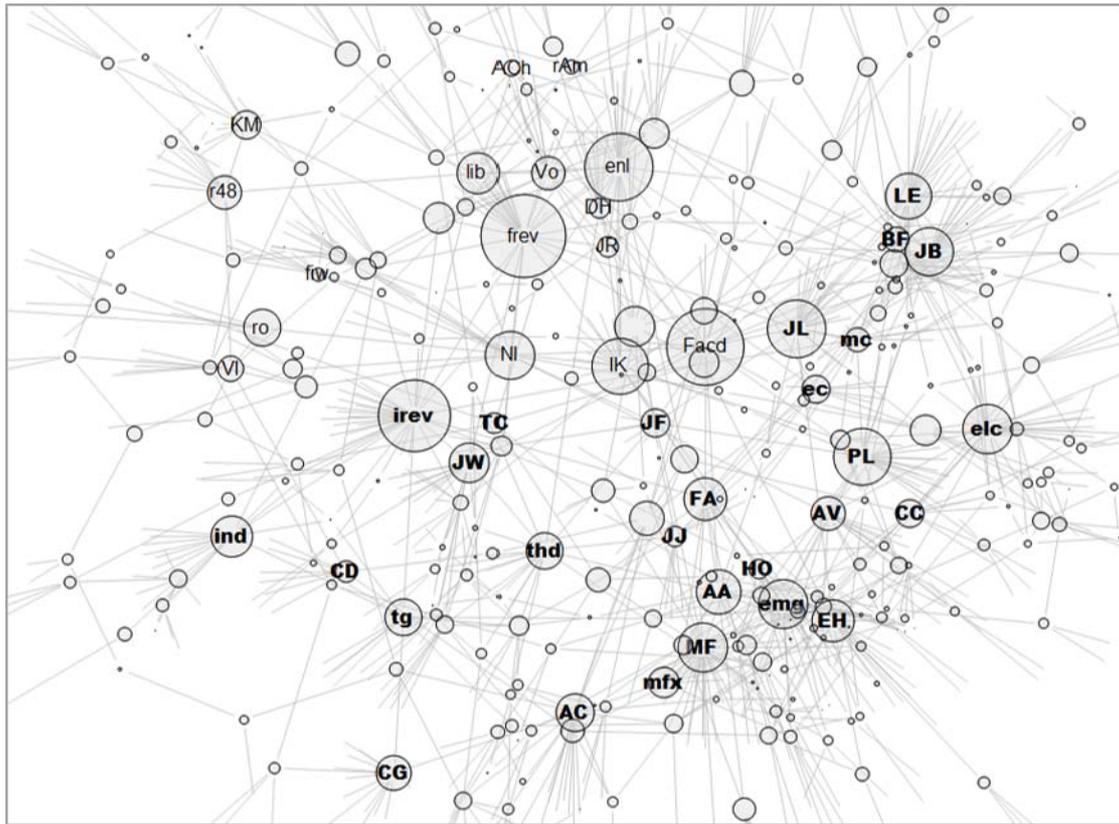


Figure 3. The network of connections between characters, ideas, innovations, and events for the years 1730-1848. The size of the node is proportional to the Betweenness Centrality. The acronyms are given in Table 2.

The era from the Enlightenment (ca 1730) to Liberalism (ca 1850)

The era starting in the beginning of 18th century is often described as an era of maturing science, rational thought, and reason. This period is thus often referred to as the Enlightenment. Quite commonly, it is thought that a scientific revolution preceded the era of Enlightenment, although some scholars prefer to see the scientific revolution and Enlightenment overlapping. Here, in the form of the pre-task, the periodization was chosen so that the year of Newton's death (1726/7, depending on calendar) set the dividing line. The end of the period was chosen to be the year of European revolutions, 1848. This period of roughly 120 years was also divided between two pre-tasks, with division line marked by the outbreak of the French Revolution in 1789.

The historical landscape constructed on basis of the students' responses to the two pre-tasks covering this period of 120 years, roughly from 1730 to 1850, is shown in Figure 3. The explanations of the acronyms for the characters, ideas, inventions, and events that emerge as central on the basis of the students' responses are provided in Table 2, where the Importance I, Betweenness Centrality BC, and Degree Centrality DC are also given for each character, idea, innovation, or event. Again, in Figure 3 the size of the nodes is proportional to the Betweenness Centrality BC, not directly to the Importance I. Comparing the Importance I and BC reported in Table 2, it is again seen that most nodes having high BC also have high I, but not all. In more detail, the relative ranking on basis of BC would be somewhat different in comparison to I. In general, the Importance seems to favor institutions (such as the Royal Society and the French Academy of Sciences) and movements (like the French Revolution and the Industrial Revolution) than individual persons. This is as expected, since institutions and movements tend to have connecting roles between individuals.

For example, based on Importance the top ten are (see Table 2): 1. French Revolution, 2. Industrial Revolution, 3. Enlightenment, 4. French Academy of Sciences, 5. Lagrange, 6. Faraday, 7. Laplace, 8. Electricity, 9. Kant, 10. Bernoulli. On the basis of BC the top ten would be: 1. French Revolution, 2. French Academy of Sciences, 3. Industrial Revolution, 4. Enlightenment, 5. Lagrange, 6. Laplace, 7. Kant, 8. Electricity, 9. Faraday, 10. Electromagnetism. Finally, on the basis of Degree Centrality DC, the top ten would be: 1. Industrial Revolution, 2. Faraday, 3. French Revolution, 4. Lagrange 5. Enlightenment, 6. Bernoulli, 7. Electricity, 8. Laplace, 9. Euler, 10. Electromagnetism. The differences among the top ten are not dramatic, but as mentioned previously it is safe to assume that for important nodes both DC and BC need to have high values.

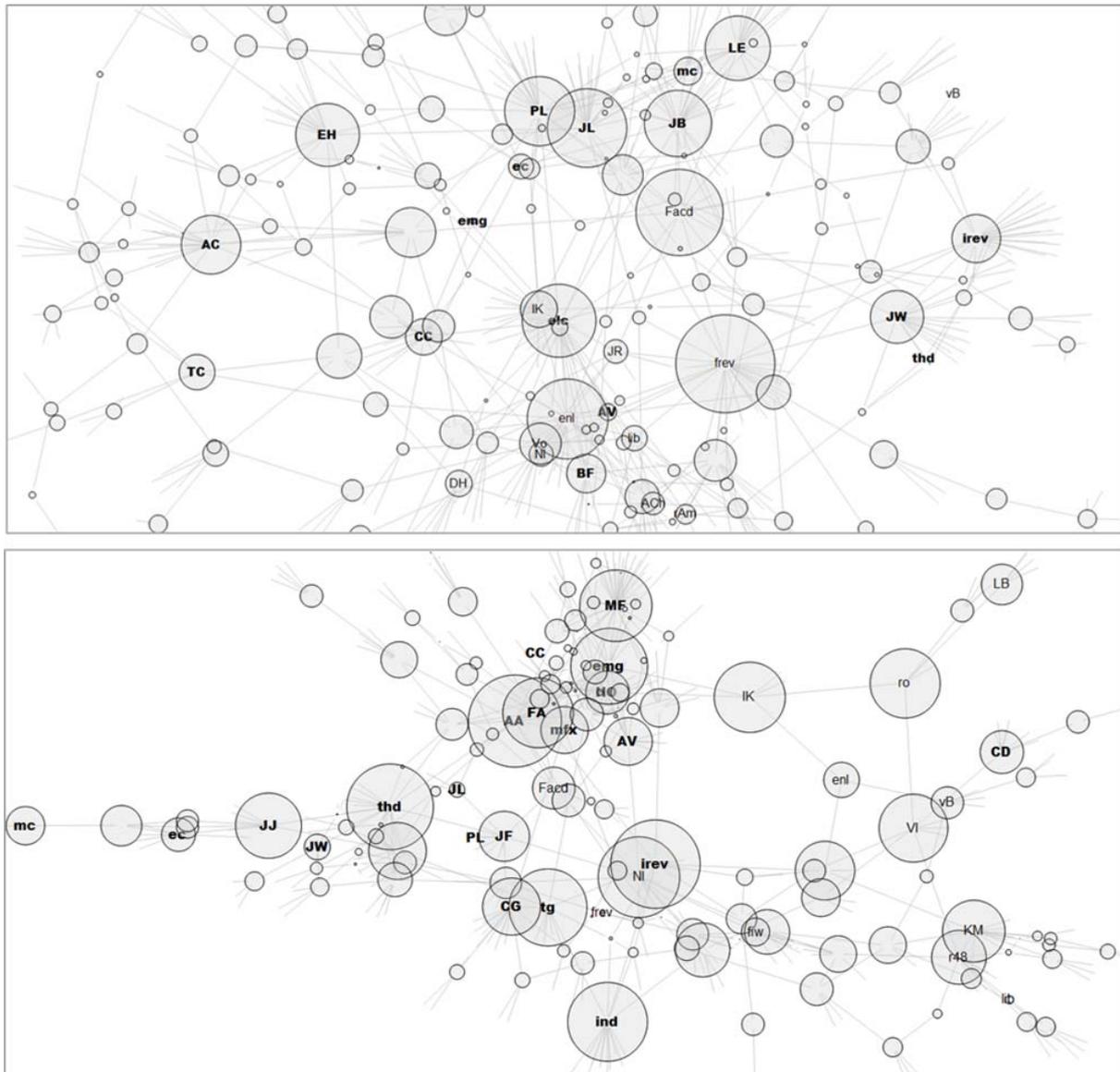


Figure 4. The network of connections between characters, ideas, innovations, and events for the sub-periods 1730–1789 (left) and 1789–1848 (right) for the years 1730–1848. The size of the node is proportional to the Betweenness Centrality.

Table 2. The most important characters, ideas, innovations, and events for the years 1550-1730 and their Importance I, Betweenness Centrality BC, and Degree Centrality DC. Note that the last ten items 21-30 were picked up from low ranking elements and exemplify how literature, arts, and music rank.

The acronyms used in the Figures are provided below.

		I	BC	DC			I	BC	DC	
French Revolution	frev	15.1	713	32		J. d'Alambert	JD	4.1	115	15
Industrial Revolution	irev	14.0	559	35		P. Maupertuis	PM	4.0	162	10
Enlightenment	enl	11.7	508	27		K. Marx	KM	3.9	120	13
French Academy	fa	11.2	625	20		J. Fourier	JF	3.8	118	12
J.L. Lagrange	JL	10.7	392	29		C. Coulomb	CC	3.7	115	12
M. Faraday	MF	10.2	296	35		A. Lavoisier	AL	3.7	133	10
P.S. Laplace	PL	9.9	380	26		Energy convs.	eC	3.5	114	11
Electricity	elc	8.8	299	26		B. Franklin	BF	3.5	95	13
I. Kant	IK	8.8	370	21		Magnetic flux	mfx	3.5	133	9
J. Bernoulli	JB	8.8	285	27		W.A. Mozart	WM	3.4	89	13
Electromagnetism	emg	8.2	295	23		Romanticism	rom	3.3	184	6
L. Euler	LE	7.9	262	24		Mechanics	mc	3.3	90	12
Napoleon I	NI	6.7	285	16		Am Independc	ai	3.3	89	12
A.M. Ampere	AA	6.5	250	17		Napoleon wars	nw	3.1	70	14
J. Watt	JW	6.2	206	19		E. Chatelet	EC	3.1	104	9
E. Halley	EH	6.2	229	17		Electric current	ec	3.0	46	19
Industrialization	ind	6.1	221	17		H.C. Oersted	HO	2.9	64	13
A. Celcius	AC	5.9	191	18		L v Beethoven	vB	2.9	116	7
Thermodynamics	thd	5.9	180	19		Victoria I	VI	2.8	99	8
F. Arago	FA	5.7	236	14		J. Joule	JJ	2.7	69	11
A. Volta	AV	5.2	160	17		C. Darwin	CD	2.6	75	9
C.F. Gauss	CG	5.0	167	15		D. Hume	DH	2.5	66	9
Voltaire	Vo	5.0	154	16		J. Rousseau	JR	2.4	70	8
Liberalism	lib	4.7	218	10		Lord Byron	LB	2.2	96	5
Telegraph	tg	4.7	181	12		T. Cook	TC	2.0	69	6
Revolution 1848	r48	4.5	157	13		A. Chydenius	ACh	2.0	46	9
American Revolution	rAm	4.2	127	14		Finnish War	fiw	1.7	33	9

It is again of interest to examine how the picture changes if the period 1730–1850 is divided in two halves, 1730–1789 and 1789–1848, using the French Revolution as the dividing line. Now, in the first half 1730–1789 of the period, the top ten in importance are: 1. The French Revolution, 2. Lagrange, 3. Enlightenment, 4. Electricity, 5. Bernoulli, 6. Laplace, 7. Euler, 8. French Academy of Sciences, 9. Halley, 10. Celcius. Interestingly, in this period, the arts are also more finely represented through Voltaire (13.) and Mozart (20.). In the period from 1789 to 1848, the top ten nodes by their importance are: 1. Faraday, 2. Ampere, 3. Thermodynamics, 4. Electromagnetism, 5. Industrialization, 6. Napoleon I, 7. Industrial Revolution, 8. Telegraph, 9. Arago, 10. Gauss. The visualizations of these sub-periods are shown in Figure 4.

From the two sub-periods, it is evident that there is a shift towards topics related to electricity and electromagnetism. When the whole period is examined, the institutions and events gain more importance in comparison to individual scientists. This is of course as expected, as institutions have long life spans in comparison to humans.

A historic landscape spanning three centuries 1550–1850

The inspection of the whole time span of three centuries from 1550 to 1850 provides an overall picture of those characters, ideas, innovations, and events that were ascribed the most importance in the students' responses in the pre-tasks. Although this is a picture that agglomerates very different aspects, it tells us what one can expect students to pay attention to when they browse material containing information of history of science (such as Wikipedia) or reading introductory texts of history of science (such as e.g. Holton and Brush, 2001) of this period is of concern. Based on Importance, the top ten elements in the landscape that spans all three centuries is as follows: 1.

persons (Newton, Galilei, Leibniz and Descartes), the highest global centralities, the Betweenness Centralities, are associated with event (French Revolution, Industrial Revolution and Enlightenment). In both cases, however, the luminous items stand clearly apart and above most of the event. This is very clearly revealed by plotting frequency of a given value of centrality against the value of the centrality in log-log plot. This is shown in Figure 6 for Degree and Betweenness centralities DC and BC, respectively, for the data 1550-1850.

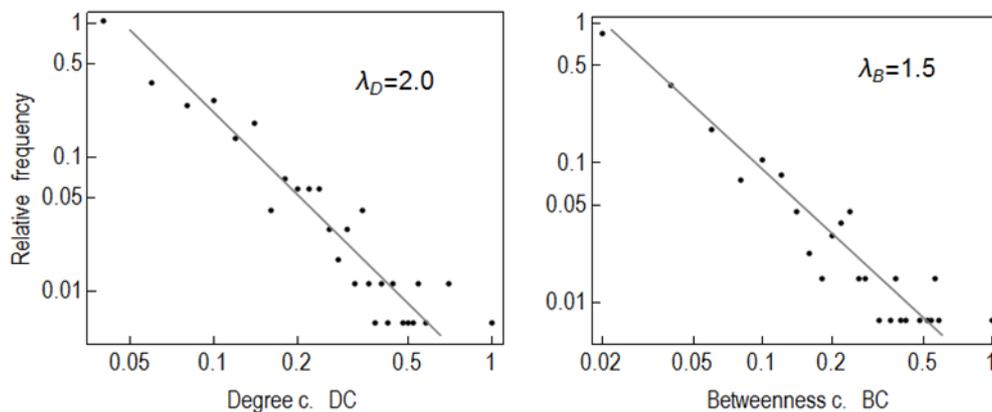


Figure 6. Log-log -plots of distribution of values of Degree and Betweenness Centralities DC and BC. The straight line in log-log scale indicates inverse power law dependence $P(X) \sim X^{-\lambda}$, where $X=DC$ or BC. Note that now normalized values of BC are used.

The log-log plot of centralities DC and BC shows that the centralities are distributed according to the inverse power-law $P(X) \sim X^{-\lambda}$, where $X=DC$ or BC, with power $\lambda \approx 2.0$ for DC and $\lambda \approx 1.7$ for BC. Such an inverse power-law distribution is a hallmark of Matthew-effect, demonstrating that majority of nodes have very few links and very few nodes have very many links, in fact, collecting most of the links in the network (Price 1976; Börner 2009).

From point of view of content, it is of course expected that central figures of science history gather most attention. It is, however, questionable if a balanced picture is possible if the discrepancy between the most luminous characters and events is so large in comparison to all other characters and events. Rather than reflecting any absolute, self-evident ranking, such feature appears to be a fame effect, where some famous characters and events draw unwarranted attention in cost of other events. Such bias probably hinders students in forming a balanced picture of history of science.

Discussion and conclusions

We have here examined how the students' initial ideas of how different characters in the history of science, scientific ideas, inventions, and events in general history are connected. The data used as the starting point consists simply of connected pairs of these items. Despite the apparent simplicity, such pairwise associations are unexpectedly revealing when they are examined as a connected set of associations; the separate pairwise associations form a tightly connected network where certain parts of the network are more tightly connected than others. This connected network forms a kind of associative network of history, on which the students are able locate the events.

The present study is an exercise in the cartography of students' associative knowledge, based on utilization of graph visualization and certain centrality measures to locate the most important elements in the network. The centrality measures used here are the Degree Centrality, which is simply the number of links the node connects itself with to other nodes, and the Betweenness

Centrality, which measures how many paths between different nodes pass through the node. Degree Centrality is thus proportional to how important the given node is in its local environment, while Betweenness Centrality is related to the importance of the node in whole network. Clearly, nodes that are important should have high Degree and Betweenness Centrality at the same time. In locating such nodes, we have used the Importance I, which is the geometric mean of Degree and Betweenness Centralities.

In the networks that represent the students' associative knowledge of science history, the nodes that have high Degree Centrality are nearly always persons, mostly scientists, next often philosophers, and third, rulers. The overall picture is thus quite person-centred. This is perhaps a reflection of the biography-dominated history perspective that textbooks tend to offer, as for example Leite (2002) has noticed in analysing physics textbooks. As expected, the characters familiar from science, mathematics, and philosophy dominate the hall of fame. Newton stands out as an absolutely dominating figure of the early modern era 1600–1850, followed by Leibniz in second place and Faraday in third place. Somewhat encouragingly, philosophers are also well recognized, and for example Descartes, Spinoza, Locke, and Hobbes have important positions in the network of associative knowledge. When one focuses on the items that have high Betweenness Centrality, the overall picture changes. Over the span of the three centuries, the Reformation and Restoration era are viewed as central, as well as the French Revolution, American Revolution and independence, and then, afterwards, Liberalism. Connected to these, the Industrial Revolution and industrialization are also perceived of as important historical events. This is already a good starting point, although yet far from deeper analysis of such connections. Of the arts, only music seems to have some role, and of that group Bach, Händel, and Mozart are the most important triad. A bit disappointingly, literature and the visual arts are both rather thinly represented. The era of Enlightenment, roughly the period 1730–1789, is different from the other eras over the years 1550–1850 studied here. For this era, the students focused on events and institutions rather than on persons. However, this era is also exceptionally centred on France and French science.

A common and interesting feature of the distribution of centrality values is that they follow an inverse power-laws, which is typical for so called Matthew-effect, where fame catches fame or rich get richer. This effect, however, contains a risk that students' history conception, especially when the Mathew-effect concerns individual scientist, is overwhelmingly dominated by persons, which are superior to most other scientist of the same era. Such bias distracts students' attention from the complex interplay of scientist, ideas and inventions and easily provides a picture, where whole endeavour of science is created by few exceptionally intelligent researchers; a conception giving rise to kinds of hagiology of famous scientists. Clearly, a step to be taken is to guide the students' attention to areas which are underrepresented, and to make them aware of the unwarranted and strong bias towards a biographical and person-centred conception of history.

In summary, we have proposed a network-based approach to the analysis of students' representations of their conceptions of the interlinked nature of science history and general history, as well as cultural history. Many features which remain undetected, or are identified only with difficulty by traditional methods, become accessible using networks to represent how different facts are related. Moreover, using suitably chosen centrality measures it becomes possible to discover the most essential elements in those networks, and to see what kind of changes occur by changing the time-window on history. This demonstrates the advantages of developing analytical methods sophisticated enough for problems in the field of learning and education. This knowledge has direct impact on the design and planning of education and courses which could better address the need to facilitate a deeper understanding of the related nature of science history and history in general.

References

- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D. U. (2006). Complex networks: structure and dynamics. *Physics Reports*, 424, 175–308.
- Börner, K. & Schamhorst, A. (2009). Visual conceptualizations and models of science. *Journal of Informetrics*, 3, 191–209.
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z. & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3, 191–209.
- Estrada, E. (2012). *The structure of complex networks*. Oxford: Oxford University Press.
- Galili, I. & Hazan, A. (2001). Experts' Views on Using History and Philosophy of Science in the Practice of Physics Instruction. *Science & Education*, 10, 345–36
- Gooday, G., Lynch, J. M. Wilson, K. G. & Barsky, C. K. (2008). Does Science Education Need the History of Science? *Isis*, 99, 322–330.
- Holton, G. (2003). What Historians of Science and Science Educators Can Do for One Another. *Science & Education*, 12, 603–616.
- Holton, G. & Brush, S. G. (2001). *Physics, the Human Adventure*. New Jersey: Rutgers University Press.
- Höttecke, D. & C. C. Silva (2011). Why Implementing History and Philosophy in School Science Education is a Challenge: An Analysis of Obstacles. *Science & Education*, 20, 293–316.
- Koponen, I. & Nousiainen, M. (2014). Concept Networks in Learning: Finding Key Concepts in Learners' Representations of the Interlinked Structure of Scientific Knowledge. *Journal of Complex Networks*, 2(2), 187–202.
- Leite, L. (2002). History of Science in Science Education: Development and Validation of a Checklist for Analysing the Historical Content of Science Textbooks. *Science & Education*, 11, 333–359.
- McClure, J. R., Sonak, B. & Suen, H. K. (1999). Concept map assessment of classroom learning: reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36, 475–492.
- Price, D. J. de Solla (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27, 292–306.
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*. 159(3810), 56–63
- Roseman, J. E., Stern, L., & Koppal, M. (2010). A Method for Analyzing the Coherence of High School Biology Textbooks. *Journal of Research in Science Teaching*, 47(1), 47–70.
- Ruiz-Primo, M. A. & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33, 569–600.
- Russell T. L. (1981). What History of Science, How Much, and Why? *Science Education*, 65, 51-64.
- Simonyi, K. (2012). *A Cultural History of Physics*. Boca Raton, FL: CRC Press
- Snow, C. P. (1959). *The Two Cultures and the Scientific Revolution*. The Cambridge University Press: Cambridge.