Research article

Characterising curricular goals for colligations in students' causal arguments

James Edward Carroll[1],[*] [ID]

[1] Department of Education, University of Oxford, UK
* Correspondence: james.carroll@education.ox.ac.uk

How to cite

Peer review

This article has been peer-reviewed through the journal's standard double-anonymous peer-review process, where both the reviewers and authors are anonymised during review.

Copyright

Open access

History Education Research Journal is a peer-reviewed open-access journal.

Abstract

In the 'historical thinking' tradition of curriculum design, the philosopher of history W.H. Walsh's concept of colligation has mostly been adopted to enable students to construct coherent, powerful and usable big pictures of the past. Less attention has been paid to the potential of colligation in enabling students to construct causal arguments at meso- and micro-levels, despite Walsh's arguments emerging from twentieth-century debates regarding the status of historical explanation. A theory-building case study was conducted with a class of 17- and 18-year-olds at a sixth-form college in England to identify possible curricular goals for colligation in students' causal arguments at higher resolutions. To characterise the status of disciplinary colligation, analytic philosophies by Walsh and others, as well as authentic historical explanations from one historiography – the Salem witch trials – were analysed by reference to one another. The students' work suggested that some were capable not only of constructing their own causal colligations, but also of appreciating the disciplinary framework that underpinned those constructions. Curricular goals for historical causal colligation are identified: individuation and historical contextualisation; reification of underlying explanatory models; and clarity regarding colligation's status in relation to disciplinary and substantive concepts. Finally,

recommendations are made to those operating in the historical thinking tradition on how they may achieve more empirical warrant for their claims regarding the essential nature of historical explanation.

Keywords historical thinking; colligation; causation; explanation; history-teacher research

Introduction

In the English tradition of history-education research into curriculum design, a predominant focus has been on 'historical thinking' (Smith, 2024; Smith and Jackson, 2021). Participants in this discourse have tended to argue that the purpose of school history is to equip students – at developmentally appropriate levels – with academic historians' expertise. Consequently, epistemic connections between 'school history' and 'disciplinary history' – defined as academic historical knowledge predominantly produced in universities – are considered desirable. To achieve this end, research has often centred on how to organise curricula around demarcated 'substantive' and 'disciplinary knowledge', and their interplay. For researchers such as Peter Lee (2005: 61–5; 2017: 60), the former represents the 'substance' of history and is exemplified by concepts such as 'peasant' or 'democracy'. The latter is encapsulated by syntactic 'metahistorical' or 'second order' concepts such as 'causation', 'change and continuity' and 'similarity and difference', which are designed to induct students into historians' disciplined modes of thinking (Lee and Ashby, 2000; Shemilt, 1980). These concepts seek to capture the types of questions that historians conventionally ask, historical methodologies, how historical claims are made and challenged and how substantive content regarding the past is shaped into historical knowledge.

This history-education research conducted by academics privileging historical thinking has also influenced history-teacher research in England, which itself amounts to a sustained, coherent and codified research discourse of professional knowledge (Counsell, 2011; Fordham, 2016). Unlike the university-based researchers, these teachers have not conducted large-scale studies. Rather than investigating the extent to which students currently think historically or the efficacy of approaches in achieving that thinking, these practitioner-researchers have predominantly focused on curriculum construction and evaluation, theorising about the curricular 'what', rather than the pedagogical 'how' (Counsell, 2016). While the history-teacher researchers have been influenced by the aforementioned history-education research by Lee and others – for example, by adopting substantive and disciplinary concepts as an interpretative framework – they have been more concerned with addressing immediate problems of practice.

A further appreciable difference between the two discourses is that history-teacher researchers have tended to seek disciplinary warrant for their theorisation in methodologies by practising historians or authentic historiography, as opposed to philosophies of history (Fordham, 2016; Lee, 1983). An example of these overlapping discourses and their differing loci of disciplinary authority has been the co-option of the analytic philosopher of history W.H. Walsh's (1942, 1974) concept of 'colligation' (for example, Cercadillo et al., 2017; Lee, 2004; Lee and Howson, 2009). For Walsh (1942: 133–4), colligation rendered the 'confused jumble' of raw historical data intelligible by reducing them to revealing, pervasive themes. While history-teacher researchers have drawn on such theorisation, they have only been influenced by Walsh as mediated by Lee and other university-based researchers (for example, Carroll, 2016; Ellis, 2020; Fearn, 2018; Michalaki, 2021; Navey, 2018; Rodker, 2019; Valentine, 2017). Instead, the history teachers' claims about disciplinary colligation have been validated by reference to narrative histories by authors such as Ian Kershaw (1999), Richard Evans (2004) and John Hatcher (2008). In either relying on select works of philosophy of history or on historiography to provide disciplinary warrant about colligation, participants in both discourses have tended to avoid corroborating philosophers' prescriptions with the praxis of historians. Consequently, claims about 'disciplinary' colligation risk being overgeneralised, oversimplified and unsubstantiated, which in turn compromises the epistemic connections between 'academic' and 'school' history desired by advocates of historical thinking.

Thus far, history-teacher researchers have not systematically investigated the proximate-practical, ontological and epistemological considerations for the curriculum designer when trying to induct students into expert historical thinking regarding colligation. In keeping with the research tradition of

history teachers in England, a theory-building case study was conducted with one of my own classes, seeking not only to solve a problem of my own practice, but also to contribute to professional curricular theorisation. In particular, I aimed to identify achievable curricular goals for colligation in students' causal arguments at the meso- and micro-levels (Counsell, 2011; Fordham, 2016; Thomas, 2013). In order to characterise the status of colligation in disciplinary history to inform the planning of my lessons, I surveyed explanations in an authentic historiography – the Salem witch trials – and also in philosophies of history by Walsh and those who have written in his wake, interrelating the two.

The students' work suggested that at least some 17- and 18-year-old students are capable not only of constructing their own causal colligations, but also of appreciating the underpinning disciplinary framework that those constructions symbolise. I identify curricular goals for historical causal colligation which capture domain specificity: individuation and historical contextualisation; reification of underlying explanatory models; and clarification of colligation's status in relation to disciplinary and substantive concepts. Finally, I recommend that those operating in the historical thinking tradition(s) may provide more warrant for their claims about the essential nature of historical explanation by employing more systematicity, diversity and representativeness in their analyses of 'the' historical discipline.

Literature review

Defining colligation

The term 'colligation' appears to have been first coined by the philosopher William Whewell (1840, 1847) in the mid-nineteenth century, and it derives from the Latin *colligere* – meaning 'to bring things together' (McCullagh, 2011: 152). In 1942, Walsh (1942) influentially appropriated the concept, arguing that colligations such as the 'Industrial Revolution' or 'Enlightenment' had an organisational function in historical argument (see also Dray, 1959; Kuukkanen, 2015; McCullagh, 1978). Such colligated themes have no direct experiential referents, and consequently they cannot be explained by referring solely to a concrete object, person or event. Higher-level colligations are themselves comprised of increasingly lower-level colligations or individual actions, which are couched less theoretically, and about which historians are less likely to dispute. Ever more encompassing colligations enable the historian to cover an entire period in a colligatory pyramid (Walsh, 1974). For example, the historian E.W. Baker's (2015: 7, 12–13) colligation 'the Salem witch trials' (*Q*) is, on the one hand, incorporated into higher-level colligations such as the 'great age of witch hunts in Europe and America', while, on the other hand, encompassing lower-level colligations or events such as:

- (*x*) 'a reservoir of pent-up local complaints and differences'
- (*y*) 'a serious spiritual and political crisis in Massachusetts'
- (*z*) 'growing war panic' (Baker, 2015: 54, 64–7).

By bringing together a collection of happenings *x*, *y* and *z* and colligating them as *Q*, we can resultantly speak pithily about *x*, *y* and *z*.

Colligation as organisation is how Walsh's arguments have been most enthusiastically adopted by history-education researchers, perhaps most notably by Lee (Retz, 2018). For Lee, the value of colligation to history-curriculum designers lies predominantly in enabling students to construct coherent, powerful and usable big pictures of the past which encapsulate states of affairs, processes and changes. Colligations useful to such ends include 'The Renaissance', 'The Reformation', 'The American Constitution', 'The Holocaust' and 'The Cold War' (for example, Cercadillo et al., 2017; Lee, 1983, 2004, 2017; Lee and Howson, 2009). In the wake of Lee's (1983) recommendation that researchers devote 'more systematic' attention to this facility of colligation, numerous history-education researchers in a variety of international contexts have explored how colligation enables students to, for instance, contextualise atomised information about the past, as well as to apply the methodological apparatus of the discipline over broad timescales (for example, Havekes et al., 2012, 2017; Lévesque, 2008; Lévesque and Clark, 2018; Shemilt, 2000; Van Boxtel and Van Drie, 2012, 2017, 2018; Van Drie and Van Boxtel, 2008). In establishing the focus on colligation and macro-scale narratives, Lee (1983) simultaneously helped divert attention from colligation's role in students' construction of causal arguments at the meso- and micro-levels. According to Lee (1983: 31–5), 'colligation under appropriate conceptions is more interesting, and its importance for history extends much wider, than the area of explanation'. (While

acknowledging that 'explanation' can have multiple meanings in history, from this point I am defining it as 'explain why', which is how most historians employ the term when making causal arguments [Førland, 2004; Megill, 2007; Smith, 2011].) With certain notable exceptions (for example, Halldén, 1997, 1998), the researchers following Lee have been similarly indifferent to colligation's role in enabling students to construct mid- and lower-scale causal arguments (for example, Lévesque, 2008).

Colligation as a problem of practice

The existing history-education research surrounding colligation has clearly been theoretically powerful and empirically generative, and participants in the discourse are convincing when arguing that the uses of colligation need not be reduced to discussions of causation (Lee, 1983; Lévesque, 2008). It is possible, however, that they have been somewhat occlusive towards colligation's functions in students' lower-level causal arguments. While the majority of history-education researchers concentrating on colligation have proposed broader structural curriculum reforms, as a practising history teacher in a sixth-form college (ages 16–18) in England teaching the Salem witch trials, my concerns were more proximate and pragmatic: ensuring that my students achieved success in national Advanced-level (A-level) examinations. Similar to what researchers have noted in other international contexts, while the awarding body Pearson Edexcel make no specific mention of colligation in their specifications, assessment materials and endorsed textbooks, all these resources implicitly presume that successful candidates will be able to navigate and construct their own mid- and lower-level colligations in their causal arguments (Lévesque, 2008; Retz, 2017; VanSledright, 2010). (As well as the sample assessment question that the students in the case study ultimately answered, which requires an explanation of the relative causal importance of the colligated cause 'the unusual political conditions operating in Massachusetts in 1692', Pearson Edexcel's [2014: 364] indicative content for answering this question suggests that students could be rewarded by referring to further colligated causes such as 'a political vacuum in Massachusetts', 'real concern about Indian attacks on the Indian frontier' and 'social tensions within the community'. Furthermore, Pearson Edexcel's endorsed textbooks ask students to argue the causal importance of colligations such as 'the weakened authority of the Massachusetts government after the Glorious Revolution' [Bullock, 2016: 171].) In order to successfully read and write explanations of events such as the Salem witch trials – where 'success' is determined not only in terms of 'writing like a historian', but also in terms of national examination achievement – the student needs to scale-shift, navigating and constructing colligations of differing generality, and appreciating their interrelationships.

Characterising disciplinary causal colligation

Further theorisation appears necessary regarding colligation's ontological status in history-curriculum design, particularly if disciplinary warrant is said to derive from analytic philosophy of history. Philosophers of history such as Walsh tend not to substantiate their propositions regarding colligation by reference to authentic historical explanations, and none have conducted thoroughgoing historiographical surveys to this end. These lacunae have been reflected in the history-education research which has sought disciplinary warrant in such philosophies. While analytic philosophy of history provides consistent, coherent and general conceptions of historical epistemology based on abstract argument, such arguments have been criticised for being generalised, oversimplified, prescriptive and divorced from historians' praxis (for example, Froeyman, 2009; Tucker, 2009). Amid prolonged and fierce twentieth-century debates in Anglophone analytic philosophy of history about the status of historical explanation, Elazar Weinryb (1975: 33) stated that the 'original sin of analytic philosophy is its unsystematically causal use of historical examples. Philosophers have based pretentious theories on a deplorably small number of examples.' Walsh (1951: 62), who participated in these debates, claimed that colligation was a process 'which historians do use, and therefore any account of historical explanation should find a place for it', but he did not provide any accompanying analysis of authentic historical explanations to substantiate such claims. Described models or theories derived from authentic historiographical explanations might therefore provide empirical warrant or contradict the general claims made by philosophers, thereby linking the descriptive and the normative. A survey of explanations from the historiography of the Salem witch trials to inform such curricular theorisation is apt, because 'explaining why' has been one of the predominant exercises of these historians. As Benjamin C. Ray

(2015: 2) notes, 'historians have repeatedly asked why Salem's witch hunt became so widespread, lasted so long, and spiralled so dangerously out of control'.

Walsh (1942, 1951) initially argued that colligation itself represented a specifically historical explanation-type. This argument can be seen as a response to – but also in some senses as a compromise between – the strict idealist and positivist explanatory models preponderating at that time. Walsh was partly critiquing the positivists' 'unificationist' view of historical explanation. For unificationists, science provided the explanatory ideal because the logic, syntax and semantic rules of 'explanation' were the same, invariant of discipline (Tucker, 2009; for example, Hempel, 1942, 1962). For Walsh (1951), however, historical colligation differed from scientific generalisation. While the scientist generally relates events of the same class – identifying general rules across units of comparison – the historian, by contrast, explains a unique historical event not by reference to general laws, but instead 'by tracing its intrinsic relations to other events and locating it in its historical context' (Walsh, 1951: 59). For example, when explaining how King William's War (1688–97) contributed to the Salem witch crisis, Mary Beth Norton (2002) did not begin by relating it to other wars outside colonial New England. Instead, she associated it with other subordinated entities from its own period, such as the status of the Wabanaki in the Puritan colonists' belief system; the first 'afflicted' accusing the Native American Tituba; and some accusers' fits appearing to resemble Wabanaki raids. In part prompted by a thoroughgoing critique by Marvin Levich (1965), Walsh (1974) later tempered his claim that colligation itself represented a specifically historical form of explanation, suggesting instead that it represented a broader interpretative undertaking which facilitates historical understanding in a more general sense (Stanford, 1998). Regardless, Walsh's (1974: 136–7) later accommodations did not preclude the investigation of colligation's relationship to causal argument; he still maintained that 'we have recourse to colligatory concepts in the interest of explaining something'.

Colligation, then, is not necessarily a separate exercise from causal argument. In a historiography such as the Salem witch trials, many of the participating historians have aimed to explain a short-term, regional event and, consequently, causes have often doubled as pervasive themes. Such colligations enable the construction of 'significant' narratives which simultaneously describe and explain (Walsh, 1942). Accordingly, colligation is not synonymous with 'generalisation' or 'grouping'; instead, it constitutes 'an identifiable whole which is more than just the sum of its parts' (McCullagh, 1978: 283). For L.B. Cebik (1969: 45):

> the colligation of events (and/or conditions) *x*, *y*, and *z* as a *Q* allows one to see *x*, *y*, and *z* as one could not seem them before, i.e., logically prior to the colligation. Colligation adds something, but not new empirical information. Rather, it adds (or perhaps better, it changes) a conceptual framework, a kind of discourse.

In short, colligation 'illuminates' the colligated facts, making them 'intelligible' in a way they were not prior to colligation (Walsh, 1942, 1974). In the case of causal argument, the built-in relevance relation of the question being answered will affect such alteration of the conceptual framework, providing a descriptive generalisation which simultaneously points towards a causal explanation. For example, Ray (2015: 46, 66–7) colligated events such as:

- (*x*) the believing of the afflicted
- (*y*) the failure to demand bonds
- (*z*) the holding of examinations in public

    as:

- (*Q*) the 'zeal of the magistrates'.

Here, Ray (2015) employed a colligation rather than a grouping such as 'the magistrates' actions' because 'the zeal of the magistrates' more obviously explains why a witch crisis occurred – by implying a propensity for witch-hunting. By doing so, Ray 'both sums the individual events and tells us how to take them' (Walsh, 1974: 136–7). Furthermore, historians explain the causal relationships between colligations, which act as both the subjects and the objects in causal claims (Dray, 2006; McCullagh, 2011). For example, when Ray (2015: 66–7) argues that 'the rampant confessions fuelled more accusations as confessors named more and more suspects, which, in a vicious judicial circle, continued to legitimate the court's arrests and convictions', colligations such as 'rampant confessions' and 'a vicious judicial circle' adopt agentive properties such as 'fuelling' and 'legitimating'. Colligation therefore does not simply precede the explanation: both are linked reflexively, co-constituting one another.

Students' understanding of causal colligations

In terms of epistemology, additional research appears necessary regarding students' understanding of the colligations they encounter, as well as their ability to construct their own colligations. Colligation moves beyond strictly configurational understandings, where individual events are selected and placed in narrative order. Instead, categoreal understandings are required, as the events are placed into a category under a conceptual framework (Dray, 2006). Consequently, colligation is highly interpretative (Walsh, 1974). Colligation's status as interpretation is also a consequence of the fact that 'the known data' are not necessarily agreed upon by all colligators, who in any case will make value judgements regarding the relative relevance and importance of any constitutive data (White, 1965). Given the interpretative and argumentative essence of colligation, some of historiography's 'fiercest controversies have centred on how given groups of events are most appropriately colligated' (Dray, 2000: 229).

Colligations' construction therefore requires conditions which warrant belief in their truth as judged by the probative community of historians (McCullagh, 1969, 2011). However, while historians often have tacit working criteria regarding when to accept colligations, these principles are not universally agreed, and many historians might struggle to articulate their own principles (Walsh, 1974). Such criteria will almost certainly include displaying a demonstrable relationship between the colligation and the supporting evidence. This evidence is framed, comparatively, in untheoretical terms, and it will generally be considered relatively inarguable by historians. A further criterion is that the historian should avoid manipulating the data into Procrustean structures (Mandelbaum, 1977). Instead, according to C. Behan McCullagh (2011: 156), 'conscientious historians will carefully consider which patterns best fit the data available to them'. Additionally, although counterexamples are inevitable, the historian should be alert to exceptions to their colligation. If exceptions become too numerous, or if the omitted data are too salient, then the generalisation becomes untenable. In sum, a colligation should successfully cover the known detail while not offering a misleading impression of the whole. To do otherwise, the historian 'would forfeit the sympathy' of their peers (Walsh, 1974: 138–9).

In the historiography of the Salem witch trials, we encounter historians justifying their colligations – even for or against the ostensibly inarguable 'Salem witch trials'. For example, Baker (2015: 12–13) argues:

> In 1692 witchcraft spread across Essex County and as far away as Boston and southern Maine. More people from Andover were accused than from any other town, and it is even possible that these events triggered trials in Connecticut. It might be most appropriate to refer to the subject of this book as 'Essex County witchcraft' or even 'the New England witchcraft crisis of 1692'. For simplicity I have elected to stay with 'Salem witch trials', but I acknowledge that much of New England was involved. In 1692, people would have said the trials took place in Salem, but I will follow the convention, used by many historians, of referring to it in the colonial era as Salem Town, to differentiate the urban core on the waterfront from Salem Village. Like many modern scholars, I will at times use the term *outbreak* to describe the events at Salem and other large witch hunts. It is a useful descriptor for a phenomenon that spread rapidly, like an infectious disease. However, in doing so, I do not mean to endorse a biological explanation for the events of 1692.

Here, Baker (2015) operates with partially communicated criteria justifying his colligations, for example, by directly referencing his peers (for example, 'like many modern scholars') to validate his choices. Additionally, Baker (2015) interweaves a series of relevant and connected, non-theoretical, empirical statements that most historians would generally accept (for example, 'more people from Andover were accused than from any other town'). He does so to suggest colligations such as 'the New England witchcraft crisis of 1692', which might more satisfactorily cover the known data than *a priori* alternatives typically employed. Finally, Baker (2015) takes care to provide his working definition of the potentially misleading metaphorical term 'outbreak' to his readers, recognising that his colligatory choices might imply a non-historical explanation that he does not advocate.

From a disciplinary perspective, therefore, it would be inappropriate for students to complete their compulsory history education with the view that colligations are 'fixed'. Starting with Lee in 1983, there has been recognition among history-education researchers that further research is required regarding what teachers' curricular goals for colligation might be, including how to make colligation-as-argument explicit to students. For instance, Lee and Howson (2009) note that students tend to be presented with

pre-colligated 'conventional' colligations in history lessons which hide the colligations' constructedness. Lis Cercadillo et al. (2017) note such dangers, which can result in students having a historical ontology where colligations are conflated with pre-given structures that are simply discovered. As Lee (2004) notes, this danger is particularly pronounced if teachers provide students with large-scale frameworks driven by pre-colligations (see also Shemilt, 2000). A focus on colligation at higher resolutions may therefore better enable students to interrogate their status as an interpretation. A further issue identified in the literature is that students are often unlikely to employ their own colligations, leading Stéphane Lévesque (2008: 80–5) to encourage further consideration of how to 'allow students to elaborate, with the teacher's assistance, their own (general or specific) colligatory concepts' (Havekes et al., 2012). Research on the extent to which students can understand how colligations are constructed and the suppositions underlying those constructions therefore appears timely. Similarly, the extent to which students can colligate themselves, and under which circumstances, invites further investigation.

Research design

Research site and participants

The research was conducted in a sixth-form college where I was working as a history teacher. The college, which has approximately 2,000 students, is in the south-east of England, and it is high achieving in terms of national averages. For example, in 2019, approximately 65 per cent of students achieved grades A*–B. Given the study's aim of identifying curricular goals, the site presented an opportunity to gain access to participants approaching the end of their school careers who might ordinarily be expected to produce work of a high standard as judged by external examiners. These students' work, therefore, could be used to characterise the goals that younger students and/or students with lower prior attainment might aspire to, maximising the potential of what could be learnt (Stake, 1995).

The case study was bound to one Year 13 class (students aged 17–18 years) taught by the author, with 11 consenting students. The students had a range of prior-attainment levels, as judged by their mean General Certificate of Secondary Education (GCSE) grade when aged 15–16 years (Table 1). The class was studying the Pearson Edexcel (2017: 86–7) History A-level module 'The witch craze in Britain, Europe, and North America, c.1580–c.1750', which included the depth study 'Cotton Mather and the Salem Witch Hunt 1692–93'. The students studied the Salem depth study for five weeks for a total of nine 65-minute lessons. (For a more detailed overview of the lessons, see Carroll, 2018, 2022.) After the final lesson, students were asked to answer the awarding body's sample assessment question at home: '"It was the unusual political conditions operating in Massachusetts in 1692 that explain the extraordinary events in Salem". How far do you agree with this explanation of the Salem witch hunt of 1692?' (Pearson Edexcel, 2014: 355).

Table 1. Overview of the participating students' prior attainment

Pseudonym	Prior attainment (mean GCSE grade/9)
Jason	7.6
Madeleine	7.2
Sophie	6.9
Stella	6.9
Ava	6.8
Abigail	6.7
Caroline	6.1
Agnes	5.7
Naomi	5.1
Isabella	5.0
Elena	4.9

The lesson sequence

In England, awarding bodies author and manage 'high-stakes' assessments. These exigencies can result in students and teachers changing behaviours, including the narrowing of taught material and superficial learning of subjects (Baird et al., 2017; Dickens, 2021). For ethical reasons, I had to ensure that the students were given the optimum opportunity to succeed in their A-level studies. Accordingly, when making decisions in my planning I had two primary considerations. First, I had to ensure that I adhered to the awarding body's specification. Second, I drew on my professional judgement as an experienced A-level teacher regarding what was feasible for this cohort of students, who had a range of prior-attainment levels, to achieve. Beyond this uppermost level of priority, in my planning and data analysis, I aimed to ensure data-source triangulation by cross-referencing the specificities of this particular historiography with the generalised claims of philosophers of history regarding colligation (Haydn, 2019). In other words, I subdivided the 'discipline' of history into two sub-loci of authority: philosophies of history and authentic explanations by academic historians operating in the historiography of the Salem witch trials, giving the sub-loci equal weight and cross-referencing them against one another.

Each lesson involved the students reading extracts of academic historians' arguments, which included colligations. The extracts were: Baker (2015: 31–2, 41–2, 123–5, 127, 183–6, 195–6, 201–2); Boyer and Nissenbaum (1974: 31, 50–2, 68–9); Hansen (1970: 84–5, 145–6, 204); Norton (2002: 8, 34–6, 72, 77–8, 279–82, 288, 295–7); Ray (2015: 3–5, 7–8, 29, 33–5, 40–1, 46, 66–7, 89–90, 94–6, 116–17, 136–7, 144–5, 149–50, 188–201); Rosenthal (1993: 3–7, 151–2, 175, 178–82, 193–5) and Starkey (1949: 183). Lesson 9 was especially designed to highlight the constructedness of colligatory generalisation to the students. The students were asked to complete a 'card-sort' activity, in which they organised cards summarising information they had learnt over the previous eight lessons into pervasive themes (Carroll, 2016; Evans and Pate, 2007). In answering a question explaining the causes of 'the extraordinary events in Salem', the colligations that students needed to construct were necessarily lower level, further down the colligatory pyramid than the consequence specified in the question. Besides the colligated cause in the question, about which the students were obliged to argue the relative importance ('the unusual political conditions operating in Massachusetts in 1692'), the onus was put on the students to construct their own causal colligations to emphasise that colligatory generalisations are not conventional. To further make clear the constructed quality of colligation, the students were asked to present and defend their colligations to their classmates to expose them to scrutiny from a probative community (Carroll, 2017). The students were then asked to summarise their thinking using a teacher-made resource which encouraged them to display a demonstrable relationship between their colligation and their supporting empirical evidence (Figure 1).

Data analysis

Due to the case-study methodology, it was inappropriate to make claims about cause-and-effect relationships as a result of my teaching (Hammersley, 2012). Consequently, I deliberately limited my goals and instead searched for 'manifestations'. Manifestations allowed me to investigate possible examples of the students' prior experiences in the lessons revealing themselves in their subsequent writing, while concurrently acknowledging that such examples cannot conclusively be attributed to the teaching (Stake, 1995). My supporting data – which revealed the students' prior classroom experiences – were the resources my students had read in their lessons: the extracts written by academic historians, as well as materials from the awarding body. In a tradition common in English history-teacher research, the primary data were the students' essays (Counsell, 2009). Distinguishing 'supporting' and 'primary' data allowed me to systematically analyse manifestations of the students' classroom experience in their writing.

The data were also analysed to provide insight into the extent to which the students had participated in a 'discourse community' of those who produce academic knowledge regarding the Salem witch trials (Paltridge, 2012; Swales, 1990). A discourse community tends to have shared common goals; particular ways of communicating its own genres; beliefs, norms and views regarding appropriateness; and a high level of expertise in their domain, which includes a threshold of proficiency before one is admitted to membership. Manifestations were therefore subdivided into 'direct' and 'indirect' instances. The former referred to when the student appeared to draw on the exact wording of the colligations in

their resources, implying – while they were employing the discourse community's specialised terminology – that the student viewed colligations as 'fixed', rather than as their own constructions. 'Indirect' manifestations were where I identified colligations in a student's work that were similar but different to the language they had encountered in lessons. This difference may indicate that the students were moving beyond mimicry, for example, by employing synonyms of the historian's original phrasing when arguing regarding the efficacy of a particular cause or re-appropriating language they had encountered in discussion of a different cause.

Figure 1. Completed example of a teacher-made resource to scaffold the students' colligations

Findings

Colligations at differing levels of generality

Individuating colligations

Temporal and spatial delimitation

Philosophers of history have argued that a characteristic that distinguishes historical colligation from scientific generalisation is that the former is timebound in a unique historical context (Kuukkanen, 2015; Walsh, 1951). This contextualisation does not apply to the experimental scientist, who can generate their own data by restoring conditions and repeating experiments in different times and places. For example, when Ray (2015: 66–7) argued regarding the 'zeal of the Salem magistrates in 1692', the colligation was delimited both spatially ('Salem') and temporally ('1692'). Additionally, a colligation might also be individuating by indicating unusualness or uniqueness. Baker (2015: 183–6), for instance, colligated 'the legal irregularities in the proceedings'. If we accept that the discipline of history is concerned with

explaining the unique event (for example, Antonova, 2020; Mabbett, 2007), then it follows that the causes of those events also need to be unusual. In this sense, individuating colligations help the historian explain why events such as the Salem witch trials originated in Essex County, but not somewhere else, or in 1692, but not before.

Formality implying change over time

A further way of explaining why an event happened when it did, but not before, is by using colligations which incorporate change over time. Walsh (1974: 128) stressed that colligation often requires awareness of 'the importance of process'. McCullagh (1978: 268) developed this notion, identifying 'formal' colligations where what gives unity to the historical process is the form of historical change involved in that development. McCullagh (2011: 153) later added:

> historical events can be colligated as constituting a change of a certain kind, such as an instance of economic growth, or political decline, or a scientific revolution. Such phrases describe the kind of change brought about by a collection of historical events, admittedly in very general, and therefore rather vague, terms.

McCullagh (1978, 2011) suggested terms indicative of formal colligations, such as 'conflict'. Examples are evidenced in the historiography of the Salem witch trials, such as when Norton (2002: 298) argued regarding 'ongoing conflict within Salem Village'. Some colligations – particularly of the 'formal' variety – are often at least partly metaphorical (Cebik, 1969; Dray, 2000). The philosopher of history William H. Dray (2006: 777) accepted that the metaphorical quality of colligations added a 'problematic element to their logical status', but he considered judicious use acceptable because 'it draws attention to an intelligible pattern in the subject matter' (Dray, 1959: 407).

The students' essays displayed both direct and indirect manifestations of spatial and temporal limitation, as well as formality invoking change over time. For example, Madeleine wrote (direct manifestations in **bold**, indirect manifestations <u>underlined</u>):

> … unstable political conditions had a significant part to play. There was <u>uncertainty surrounding the lack of a Charter from 1684</u>, <u>unpopularity with</u> **the new Charter in 1691** and the legal instability which led to **unresolved conflicts** among the people of Massachusetts.

Madeleine's colligations indicating temporal delimitation ('the new Charter in 1691') and formality over time ('unresolved conflicts') are both direct manifestations of her reading, suggesting mimicry (Baker, 2015, Ray, 2015). In both cases, however, she adapted the historians' originals. In the case of the former, she included reference to societal disposition – 'unpopularity' – which hints towards an explanation as to why the colonists might have been increasingly inclined towards witch-hunting. In the latter, Madeleine incorporated specific spatial delimitation ('Massachusetts'), helping to explain why the crisis occurred in that particular geographic area. Furthermore, Madeleine introduced a temporally delimited colligation not derived from her reading ('uncertainty surrounding the lack of a Charter from 1684'), helping to explain why the Salem witch trials became more likely from the mid-1680s onwards.

Students also individuated colligations through allusion to unusualness. Elena directly manifested her reading of Baker (2015) when arguing about the consequences of the accusation and ultimate execution of Salem Village's former minister, George Burroughs. According to Elena, George Burroughs's accusation 'led to the belief that if … Burroughs for example, could be a witch, anyone could, creating room for **illegal irregularities** to be considered without public opposition'. When expounding on her argument, Elena more indirectly manifested her reading by employing synonyms: '<u>the distinctive legal processes</u> in Salem explain why so many more people [were convicted] of being witches than in cases before the Salem 1692 witch trials'. In both cases, Elena's intimation of peculiarity provides insight into why a unique consequence might have occurred.

Finally, some students constructed formal colligations indicating change over time which had no obvious analogues in their reading. For example, Jason wrote:

> The Salem witch craze of 1692 erupted from <u>a growing atmosphere of unease and anxiety</u> stemming not only from the unusual political conditions, but also as a consequence of Native American raids, and <u>growing tensions</u> between the pre-modern puritan demographic and the newer, less religious <u>centred and more</u> economically motivated settlers.

Jason's colligation, where the quality uniting the colligation was a metaphorical process, helps to provide explanation of why a witch crisis did not occur prior to 1692; antagonisms between the factions in Salem Village had to reach a tipping point, enabling the disaster.

General colligations

While colligations are historically specific, in some cases they can draw on transhistorical substantive concepts. McCullagh (1978) accepted that some colligations did not refer to more than one historical process and therefore could be termed 'singular' colligations. McCullagh (2011) deviated from Walsh, however, by arguing that some colligations – although still explaining specific historical events – employ general terms and, as such, classify by referring to wholes that have common features. Such general colligations might include the 'Puritan Revolution' (Demos, 1982: 381), the 'Glorious Revolution' (Hill, 1996: 15) and the 'scientific revolution' (Norton, 2002: 6). While all these colligations have distinguishing characteristics, they all appeal to certain commonly held understandings of 'revolutions' – such as upheaval or long-term change – to warrant the inclusion of the general term. Certainly, the semantic content of the word 'revolution' is equivocal, and it can operate almost as a 'catchword'. It has diverse applications, from political upheavals to social convulsions, to intellectual innovations. Nonetheless, since the French Revolution, *révolution* – the same in many languages – has gained an ambivalent yet ubiquitous semantic quality so as to operate as a flexible 'collective singular' met in many international contexts with a certain initial comprehension (Koselleck, 1985).

While the Glorious Revolution was not explicitly referred to in the historians' extracts which the students had read, it was mentioned in the awarding body's resources (Pearson Edexcel, 2017: 87). Stella directly manifested this general colligation, displaying imitation, when she wrote:

> There was [*sic*] a number of root causes that allowed the spread of the witch hunt all over New England; such as the threat of witchcraft and the devil to the perfect theocratic society and the **weakened authority after the 'Glorious Revolution' of 1688.** [Some students' mischaracterisation of Massachusetts Bay as a 'theocracy' was possibly a result of their reading of the awarding-body-endorsed textbook (Bullock, 2016: 158). In fact, late seventeenth-century Massachusetts ministers were not allowed to serve in elected office.]

Elsewhere in her essay, however, Stella developed the general, conventional example that she had encountered in her reading to create her own, modified colligation when writing 'both the fear of the devil on their perfect society, and political instability since **the Glorious Revolution**, caused mass hysteria and fear – which in turn, caused a witch craze'. In this example, Stella appeals to widely held conceptions of 'revolution' – in this case, political disruption – which, in turn, implies an explanation for colonists' witch-hunting.

Colligations indicating explanatory models

Intentionalist colligations

The framing of a colligation can be indicative of the historian's explanatory model. Although reluctant to self-identify as an idealist, Walsh's (1942) earlier views of colligation privileged this tradition, particularly R.G. Collingwood's view of explanation (Cebik, 1969; Dray, 1989). At this point, Walsh stressed that history, unlike natural science, is the study of rational beings who exercise free will. Accordingly, Walsh adopted an ontology where the historian's object of study becomes individuals' outside actions, which are understood by empathising with that person's thoughts, intentions and reasons. In this view, the intrinsic relations unifying the colligation are strictly collective actions carried out by a person or small group of people consciously acting in concert, and where their goals are ultimately achieved. Intentionalist colligations, therefore, imply teleology. We occasionally encounter intentionalist colligations in the historiography of the Salem witch trials. For example, Bernard Rosenthal (1993: 195) argued regarding the colligated intentions and actions of the individual Chief Justice William Stoughton by referring to 'the almost monomaniacal role of Stoughton'. Similarly, Baker (2015: 185) argued that the small group of Salem magistrates possessed an 'overriding need to find and convict witches'.

Certain students indirectly manifested colligations inflected with intentionalism in their explanations, elaborating on their reading. The students had encountered colligations such as

'particular decisions and motives of individuals and groups at various levels of society' (Ray, 2015: 3–4), and some elaborated and instantiated these, constructing even lower-level colligations. For example, Isabella wrote, 'when the mostly female accusers got the attention and credibility for accusing, more accusations occurred suggesting <u>the motive for young girls in a Patriarchal [*sic*] society to accuse</u> was getting attention from the male political figures'.   Similarly, Elena wrote of the causal importance of '<u>Key individuals [*sic*] decisions to allow spectral evidence</u>'.   Elena also constructed intentionalist colligations which were not obviously derived from her reading. For example, she wrote:

> in a state where people were so desperate to maintain the puritan theocracy the older citizens had once created, these conditions of fear and uncertainty led to people in Massachusetts using witchcraft as <u>a strategy to maintain the very foundations of the colony</u>.

As these two examples demonstrate, of all the students, Elena adopted the most explicitly intentionalist explanatory model, assigning responsibility for the crisis to the actions of a small group of individuals' rational, teleological decision-making.

Colligations of non-intentional action

Walsh's intentionalist view of colligation was open to many of the criticisms applicable to Collingwoodian idealism more generally.   It was reductionist and divorced from the praxis of historians, thereby undermining Walsh's own claim that colligation had more empirical warrant that other explanatory models (Cebik, 1969; McCullagh, 1998; Paul, 2015; Stanford, 1998). For example, even if a historian seeks to emphasise the causal role of human action in their explanations, Walsh's initial model did not allow for actions resulting from emotion, unconscious decision-making, irrationality, unintended consequences, chance, or purposive actions that are not easily attributable to individuals or small groups consciously acting together. In sum, explanation reliant solely on intentionalist colligation would fail to conform to most authentic historiographical causal arguments.

Furthermore, the intentionalist model seemingly disables further colligations found in the historiography of the Salem witch trials which encapsulate the actions of individuals and small groups. Bryan Le Beau (1998: ix), for instance, emphasised the consequences of omission of expected action when ascribing responsibility to 'the failures of authorities'.   In one of the most explicit examples of mimicry, Abigail adopted Ray's (2015) argument for why George Burroughs was accused of, and executed for, witchcraft. Abigail included Ray's (2015: 198–201) colligation summarising the significance of Burroughs's non-fulfilment of his role as Salem Village's minister in the view of his former congregation:

> Very few accused witches in Salem were linked to the impact of the India Raids [*sic*]. Even George Burroughs, the individual that was most connected to the Native Americans, the charges against him was [*sic*] not centred around his involvement with the Indian raids, but rather with **his failure to perform sufficiently as a Salem Pastor**.

Abigail imitatively employed a colligation encapsulating non-intentional individual actions.   She explained George Burroughs's fate as an unintended consequence of his failure to meet the expectations of his church members, thus making him susceptible to accusation.

Non-purposive, dispositional colligations and complex colligations

Considering the aforementioned and other criticisms, Walsh loosened his definition to include non-intentionalist colligations (Paul, 2015).   Perhaps most crucial in the rejection of Walsh's initial, exclusively intentionalist model of colligation was its elimination of the role that conditions play in the shaping of human consciousness and action. Regardless of whether they *should*, historians – such as those in the historiography of the Salem witch trials – include colligations describing supra-personal conditions (Cebik, 1969; McCullagh, 1978; Paul, 2015; Walsh, 1974). The admission of 'formal' colligations allows historians to construct colligations characterising social, economic and cultural changes, even if they did not have a clearly demonstrable human agent driving a teleological change, which Walsh's initial model had neglected (Dray, 1989).

Even in his earlier view, Walsh (1951) conceded that the intentionalist model was only an ideal. In practice, Walsh acknowledged that historians often resorted to 'semi-teleological' explanations where

participating agents' collective action might be driven by similar ideas, although those agents might not be consciously collectively acting on them. For McCullagh (1978), colligations can unite complementary ideas or values held by wider groups, even if those groups do not collectively recognise a preconceived goal (Dray, 1989; McCullagh, 2011). We find examples of such non-purposive, dispositional colligations in the historiography of the Salem witch trials. When historians write colligations such as a 'situation of instability and anxiety' (Rosenthal, 1993: 3–4), 'political and spiritual tensions' (Baker, 2015: 125–6) or 'anger and disappointment at the failure of the militia leadership' (Baker, 2015: 146), they are not suggesting that these groups consciously and collectively set themselves the goals of being anxious, tense or angry; rather, groups of peoples' similar dispositions helped create conditions conducive to similar actions.

Some students both directly and indirectly manifested non-purposive, dispositional colligations that they had encountered in their reading. For example, Madeleine drew heavily on Rosenthal (1993: 3–4) when arguing:

> Puritan authorities in Salem used the **political instability** caused by the absence and introduction of a new Charter [*sic*], and the **social anxiety** as a consequence of the Indian raids, as examples of Gods [*sic*] punishment. This is supported by numerous jeremiad sermons conducted by Puritan ministers such as Increase and Cotton Mather. In these sermons, ministers would preach about how the people in Salem were to blame for **the political and social instability** and that the Devil was at work in Salem. This led to an atmosphere of fear and **anxiety** amongst the community which developed into a witch-hunt in order to please God by preventing evil.

Dispositional and formal colligations which account for change over time are not mutually exclusive. Some colligations are 'complex' in displaying both dispositional and formal features (McCullagh, 1978). For example, 'an intensified campaign for moral reformation' (Baker, 2015: 185) not only indicates formality ('intensified'), but it also alludes to the dispositions underpinning the change ('moral'). Other students constructed non-purposive, dispositional colligations which were not obviously drawn from their reading, while also increasing their complexity by incorporating formality. For example, Jason argued:

> The period of political limbo created a build-up of resentment and unease, as no legal structures allowed small scale local disputes to fester. This can be seen in the amount of accusations that correspond to unresolved disputes involving the Putnam Family, a point argued by historians Boyer and Nissenbaum.

In these examples, the students employed dispositional colligations to describe conditions relating to affect – 'anxiety', 'resentment' and 'unease' – which help explain why a witch crisis became possible. In Jason's case, the identification of 'build-up' evokes a change over time, helping the reader appreciate how such emotional dispositions reached a critical mass in terms of enabling witchcraft allegations in 1692.

Conclusion

The students' experiences of causal colligation in the classroom appeared to both directly and indirectly manifest themselves in the students' subsequent work. In epistemological terms, this study provides some initial insight into the extent to which reading, understanding and writing causal colligations 'like a historian' is achievable for 17- and 18-year-olds. As Lee (2011: 184) cautioned, students from ages 14 and above are capable of using specialist terminology in 'superficially convincing ways without grasping the nature or significance of the conceptual apparatus to which this terminology pertained'. Some students' final products certainly suggested such a superficiality; the degree of mimicry implied that they continued to view their colligations as pre-given, fixed and conventional (Cercadillo et al., 2017; Lee and Howson, 2009). It is therefore questionable as to whether those students will be able to replicate a similar level of performance when asked to construct similar but different causal arguments without comparable levels of scaffolding (Havekes et al., 2012). In this sense, one short sequence of lessons – albeit explicitly focusing on the constructedness of colligations at higher resolutions – was insufficient in enabling all students to construct their own colligations. Nonetheless, the indirect manifestations in other students' work from

across the prior-attainment range suggests that at least some A-level students – under conditions which need to be more precisely determined – are capable not only of writing convincing causal colligations, but also of appreciating the underpinning disciplinary framework that those constructions symbolise.

This study suggests that causal colligations at higher resolutions are currently underappreciated by history-education researchers, and not only because teachers in England are required to support students in constructing geographically and temporally delimited causal arguments. If the historical thinking traditions valued in curriculum design in England are to continue legitimating their claims regarding what should be taught by allusion to academic history – be that to philosophers of history by history-education researchers or individual works of narrative history by history-teacher researchers – more justification, systematicity and representativeness may be required to plausibly claim that the essential features of the highly pluralistic practice of historical explanation have been captured (Førland, 2004; Froeyman, 2009; Gaddis, 2002; Hammer, 2008; Jordanova, 2000; Maza, 2017; Smith, 2011; Tosh, 2006; Van Bouwel and Weber, 2008).

Overreliance on claims made by philosophers of history – especially if such philosophers' arguments are historically decontextualised and selectively appropriated – risk overgeneralising and oversimplifying 'the' historical discipline. Furthermore, such an approach potentially lacks empirical warrant, privileging the prescriptive and ignoring the praxis of historians. Described models or theories derived from authentic explanations in authentic historiography may therefore be beneficial in corroborating, complicating or contradicting the general claims of philosophers of history, thereby linking the descriptive and the normative.

Similarly, extrapolating claims of historical explanation writ large by reference to any one historiography, let alone to a single work, has pitfalls because it represents too small a sample to make such general claims (Graham, 1983; Martin, 1989; Megill, 2007). Approaches common in the discourse of history-teacher researchers where claims are based on the work of a single author risk an unfettered relativism – especially if that author has been selected solely by virtue of being 'a historian'. By implying the sacrosanctity of practice, ineffective or non-representative examples of historical explanation may be misleadingly held to be emblematic (Førland, 2004; Skinner, 2002; Van Bouwel and Weber, 2008). Cross-referencing specificities in a historiography with normative philosophies might enable the curriculum designer to make more general claims regarding the typical patterns of colligation in explanations that historians employ.

In the view of some philosophers of history, it is feasible and useful to create a taxonomy of colligations (for example, McCullagh, 1978). Cebik (1969: 49), for instance, noted 'one may (and often does for pedagogical purposes) distinguish classes of colligatory concepts'. This study has contributed to identifying, honing and systematising achievable curricular goals for colligation in meso- and micro-level causal arguments. As well as providing some clarity in teachers' planning in the immediate term, such goals might help frame larger-scale investigations. These studies might determine with more certainty the extent to which such manifestations can be attributed to the students' experiences in lessons and the efficacy of these approaches to teaching colligation.

First, historical colligations individuate and historically contextualise, through temporal or geographical delimitation, or incorporation of change over time, or allusions to unusualness. The 'unificationist' view of explanation has elicited little sympathy since its zenith in the mid-twentieth century, with anti-unificationists emphasising history's distinctive objects of study, epistemology, metaphysics and praxis, which necessitate disciplinarily distinctive logical structures of explanation (for example, Graham, 1983; Hammer, 2008; Ricoeur, 1984; Stanford, 1994; Tucker, 2009). While the borders of the historical discipline are clearly permeable, certain philosophers of history and practising historians find the distinction between 'explanations in history' (non-historical explanations we might find in historical writing) and 'historical explanation' (specifically historical explanation) fruitful.

For example, social scientific explanations tend to identify general rules across units of comparison, while historical explanations focus on the unique event, with Bevir (2007: 308) noting the 'philosophical collapse of positivism that informs social science history with its attempts to explain historical particulars by reference to mid-level or even universal generalities' (Antonova, 2020; Hewitson, 2015; Mabbett, 2007). Such trends have been reflected in the historiography of the Salem witch trials. Historians who incorporated social scientific methods from functionalist anthropology, sociology and Freudian psychohistory to place the events of 1692 into the wider context of New England witchcraft were highly influential in the 1970s and 1980s. More recently, however, critics have argued that such works 'do not adequately explain why Salem Village's outbreak of accusations, unlike those that had come before,

turned into a thorough witch hunt' (Gragg, 1992: 213; for example, Boyer and Nissenbaum, 1974; Demos, 1970, 1982; Karlsen, 1989). This distinction is reified in colligations. For example, Walsh (1942, 1951) did acknowledge that historians occasionally included science-like 'covering laws' in their explanations, but that they characterised these as an explanation in history rather than as a specifically historical explanation. When Chadwick Hansen (1970: 145–6) argues regarding the causal importance of 'an outbreak of epidemic hysteria in Salem Village', this does not qualify as a *historical* colligation, despite existing in historical writing. While the generalisation is spatially bound, the indefinite article suggests a generally applicable, biological explanation extending beyond the specific event.

Second, in some senses, colligation treads the penumbra between the disciplinary and the substantive. Currently in history-education literature, there appears to be a lack of consensus regarding what qualifies as a 'colligatory' concept, particularly vis-à-vis a 'substantive concept' and their interrelationships. Haenen and Schrijnemakers (2000) usefully subdivided substantive concepts, distinguishing between the 'unique' – which apply exclusively to the period in question (for example, 'D-Day', 'The Peace of Westphalia' or 'Napoleon') – and 'the transhistorical' – which requires students to appreciate the characteristics that unify the concept across time, even if those meanings are subject to change over that time (for example, 'king', 'parliament' or 'depression'). The distinction between unique substantive concepts and colligations may be one of degree rather than order. A concept such as 'The Peace of Westphalia' cannot be explained by reference to a single concrete object, person or event, so it can perhaps also be characterised as a lower level colligation, albeit one that is conveyed less theoretically.

Furthermore, other researchers categorise concepts such as 'slavery', 'freedom', 'progress' and 'multiculturalism' as colligations, arguing that such terms are 'transhistorical' (Lévesque, 2008: 70–6, 104). If claims are legitimated by appeals to academia, such as the philosophy of history of Walsh, then a 'transhistorical colligation' in a *historical* explanation – contrasted with an explanation in history – is a contradiction in terms, because it provides little opportunity to explain the unique event. 'Slavery', for example, is better categorised as a transhistorical substantive concept because it is a term that relates events of the same class across time without specifying a particular historical context. In other words, such concepts are certainly potentially constitutive of colligation – particularly of the 'general' variety – but in themselves they fail to qualify as colligatory concepts because they do not provide complex particularity, and, as such, they cannot achieve colligation's purported pathway to domain specificity (Havekes et al., 2017).

Such findings have broader implications for how colligations are characterised in history curriculum design. Some history-education researchers have classified colligation as a discrete second-order concept (for example, VanSledright, 2010). Furthermore, others have claimed that citing causes is a different exercise from colligating pervasive themes to events (for example, Lévesque, 2008). A danger possibly exists in conceptualising a dichotomy between colligation and other second-order concepts, in this case, causation. Such an approach appropriates Walsh's analytic philosophy of history as a locus of disciplinary authority while simultaneously decontextualising Walsh's arguments, which were strongly situated in the mid-twentieth-century Anglophone debate about the status of historical explanation (Dray, 1989). By committing to answering a causal question, the historian is concurrently obliged to colligate the supra-personal as 'causes', and will therefore impose categoreal understandings on to the configurational (Dray, 2006). In other words, if colligation is to achieve its putative status as a second-order concept, this should not dismiss the possibility that the construction of individual colligations are themselves inflected by conceptual thinking; in this specific case, the intrinsic relations linking the isolable 'facts' are causal (Chapman and Hale, 2017). Like all models describing complex phenomena, a risk is that the categorisation of colligation as a second-order concept leads to simplification, such as by downplaying interactivity with other concepts.

Acknowledgements

Data availability statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request

Declarations and conflicts of interest

Research ethics statement

The author declares that research ethics approval for this article was provided by the UCL IOE Research Ethics Committee.

Consent for publication statement

The author declares that research participants' informed consent to publication of findings – including photos, videos and any personal or identifiable information – was secured prior to publication.

Conflicts of interest statement

Professor Arthur Chapman, current Editor-in-Chief of *History Education Research Journal*, was the author's PhD supervisor at the time the work reported in this article was undertaken; this article was not managed or edited by Professor Chapman and he was removed from the decision-making process for publication. All efforts to sufficiently anonymise the author during peer review of this article have been made. The author declares no further conflicts with this article.

References

Antonova, K.P. (2020) *The Essential Guide to Writing History Essays*. Oxford Oxford: University Press.

Baird, J.A., Andrich, D., Hopfenbeck, T.N. and Stobart, G. (2017) 'Assessment and learning: Fields apart? Assessment in Education: Principles'. *Policy & Practice*, 24 (3), 317–50. [CrossRef]

Baker, E.W. (2015) *A Storm of Witchcraft: The Salem Trials and the American Experience*. Oxford: Oxford University Press.

Bevir, M. (2007) 'National histories: Prospects for critique and narrative'. *Journal of the Philosophy of History*, 1, 293–317. [CrossRef]

Boyer, P. and Nissenbaum, S. (1974) *Salem Possessed*. Cambridge, MA: Harvard University Press.

Bullock, O. (2016) *Paper 3: The witch craze in Britain, Europe, and North America, c. 1580–1750*. London: Pearson.

Carroll, J.E. (2016) 'The whole point of the thing: How nominalisation might develop students' written causal arguments'. *Teaching History*, 162, 16–24.

Carroll, J.E. (2017) "I feel if I say this in my essay it's not going to be as strong": Multi-voicedness, "oral rehearsal" and Year 13 students' written arguments. *Teaching History*, 167, 8–16.

Carroll, J.E. (2018) 'Couching counterfactuals in knowledge when explaining the Salem witch trials with Year 1'. *Teaching History*, 172, 18–29.

Carroll, J.E. (2022) 'Terms and conditions: Using metaphor to highlight causal processes with Year 13'. *Teaching History*, 187, 40–9.

Cebik, L.B. (1969) 'Colligation and the writing of history'. *Philosophy of History*, 53 (1), 40–57. [CrossRef]

Cercadillo, L., Chapman, A. and Lee, P. (2017) 'Organizing the past: Historical accounts, significance and unknown ontologies'. In M. Carretero, S. Berger and M. Grever (Eds.), *Palgrave Handbook of Research in Historical Culture and Education*. London: Palgrave Macmillan, 529–52.

Chapman, A. and Hale, R. (2017) 'Understanding what young people know: Methodological and theoretical challenges in researching young people's knowledge and understanding of the Holocaust'. *Holocaust Studies*, 23 (3), 289–313. [CrossRef]

Counsell, C. (2009) 'Interpretivism: Meeting our selves in research'. In E. Wilson (Ed.), *School-Based Research: A guide for education students*. London: Sage, 251–76.

Counsell, C. (2011) 'History teachers as curriculum makers: Professional problem-solving in secondary school history in England'. In B. Schüllerqvist (Ed.), *Patterns of Research in Civics, History, Geography and Religious Education*. Karlstad: Karlstad University Press, 53–88.

Counsell, C. (2016) 'History teacher publication and the curricular "what?": Mobilizing subject-specific professional knowledge in a culture of genericism'. In C. Counsell, K. Burn and A. Chapman (Eds.), *MasterClass in History Education Transforming Teaching and Learning.* London: Bloomsbury, 243–52.

Demos, J. (1970) 'Underlying themes in the witchcraft of seventeenth-century New England'. *The American Historical Review*, 75 (5), 1311–26. [CrossRef] [PubMed]

Demos, J. (1982) *Entertaining Satan: Witchcraft and the Culture of Early New England.* Oxford: Oxford University Press.

Dickens, S. (2021) 'Intentional, tacit, contingent: Knowledge recontextualization in the official history curriculum–A critical discourse analysis'. *Journal of Curriculum Studies*, 53 (5), 692–710. [CrossRef]

Dray, W.H. (1959) '"Explaining what" in history'. In P. Gardiner (Ed.), *Theories of History*. New York: Free Press, 403–8.

Dray, W.H. (1989) *On History and Philosophers of History.* Leiden, NY: Brill.

Dray, W.H. (2000) Explanation in history. In J.H. Fetzer (Ed.), *Science, Explanation, and Rationality: Aspects of the philosophy of Carl G. Hempel.* Oxford: Oxford University Press, 217–42.

Dray, W.H. (2006) 'Philosophy and historiography'. In M. Bentley (Ed.), *Companion to Historiography*. London: Routledge, 763–82.

Ellis, J. (2020) 'What's in a narrative? Unpicking Year 9 narratives of change in Stalin's Russia'. *Teaching History*, 178, 32–41.

Evans, J. and Pate, G. (2007) 'Does scaffolding make them fall? Reflecting on strategies for developing causal argument in Years 8 and 11'. *Teaching History*, 128, 18–29.

Evans, R. (2004) *The Coming of the Third Reich.* London: Penguin.

Fearn, H. (2018) 'Acquainted or intimate? Towards identifying when and how background knowledge is used in subsequent learning'. *Teaching History*, 173, 41–53.

Fordham, M. (2016) 'Realising and extending Stenhouse's vision of teacher research: The case of English history teachers'. *British Educational Research Journal*, 42 (1), 135–50. [CrossRef]

Førland, T.E. (2004) 'The ideal explanatory text in history: A plea for ecumenism'. *History and Theory*, 43, 321–40. [CrossRef]

Froeyman, A. (2009) 'Concepts of causation in historiography'. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 42 (3), 116–28. [CrossRef]

Gaddis, J.L. (2002) *The Landscape of History: How Historians Map the Past.* Oxford: Oxford University Press.

Gragg, L. (1992) *The Salem Witch Crisis.* London: Prager.

Graham, G. (1983) *Historical Explanation Reconsidered.* Aberdeen: Aberdeen University Press.

Haenen, J.P.P. and Schrijnemakers, H.M.G. (2000) 'Suffrage, feudal, democracy, treaty … history's building blocks: Learning to teach historical concepts'. *Teaching History*, 98, 22–9.

Halldén, O. (1997) 'Conceptual change and the learning of history'. *International Journal of Educational Research*, 27 (3), 201–10. [CrossRef]

Halldén, O. (1998) 'Personalization in historical descriptions and explanations'. *Learning and Instruction*, 8 (2), 131–9. [CrossRef]

Hammer, C. (2008) 'Explication, explanation, and history'. *History and Theory*, 47 (2), 183–99. [CrossRef]

Hammersley, M. (2012) *Methodological Paradigms in Educational Research.* British Educational Research Association online resource. Accessed 16 June 2025. https://www.bera.ac.uk/wp-content/uploads/2014/03/Methodological-Paradigms.pdf.

Hansen, C. (1970) *Witchcraft at Salem.* London: Hutchinson.

Hatcher, J. (2008) *The Black Death: An Intimate History.* London: Weidenfeld.

Havekes, H., Arno-Coppen, P. and Luttenberg, J. (2012) 'Knowing and doing history: A conceptual framework and pedagogy for teaching historical contextualisation'. *History Education Research Journal*, 11 (1), 72–93. [CrossRef]

Havekes, H., Van Boxtel, C., Coppen, P.-A. and Luttenberg, J. (2017) 'Stimulating historical thinking in a classroom discussion: The role of the teacher'. *Historical Encounters: A Journal of Historical Consciousness, Historical Cultures, and History Education*, 4 (2), 71–93. [CrossRef]

Haydn, T. (2019) 'Triangulation in history education research, and its limitations: A view from the UK'. *History Education Research Journal*, 16 (1), 35–49. [CrossRef]

Hempel, C.G. (1942) 'The function of general laws in history'. *Journal of Philosophy*, 39, 35–48. [CrossRef]

Hempel, C.G. (1962) 'Explanation in science and history'. In R.G. Colodny (Ed.), *Frontiers of Science and Philosophy*. Pittsburgh, PA: University of Pittsburgh Press, 9–33.

Hewitson, M. (2015) *History and Causality.* London: Palgrave Macmillan.

Hill, F. (1996) *A Delusion of Satan: The Full Story of the Salem witch Trials*. London: Penguin.

Jordanova, L. (2000) *History in Practice*. London: Arnold.

Karlsen, C.F. (1989) *The Devil in the Shape of a Woman: Witchcraft in Colonial in New England*. New York: Vintage.

Kershaw, I. (1999) *Hitler 1889–1936: Hubris*. London: Penguin.

Koselleck, R. (1985) *On the Semantics of Historical Time*. London: The MIT Press.

Kuukkanen, J.M. (2015) *Postnarrativist Philosophy of Historiography*. New York: Springer.

Le Beau, B.F. (1998) *The Story of the Salem Witch Trials.* Upper Saddle River, NJ: Prentice Hall.

Lee, P. (1983) 'History teaching and philosophy of history'. *History and Theory*, 22 (4), 19–49. [CrossRef]

Lee, P. (2004) '"Walking backwards into tomorrow": Historical consciousness and understanding history'. *International Journal of Historical Learning, Teaching and Research*, 4 (1), 1–46. [CrossRef]

Lee, P. (2005) 'Putting principles into practice: Understanding history'. In M.S. Donavan and J.D. Bransford (Eds.), *How Students Learn: History, Mathematics, and Science in the Classroom*. Washington, DC: The National Academic Press, 31–78.

Lee, P. (2011) 'Historical literacy and transformative history'. In L. Perikleous and D. Shemilt (Eds.), *The Future of the Past: Why History Education Matters*. Nicosia: Association for Historical Dialogue and Research, 129–67.

Lee, P. (2017) 'History education and historical literacy'. In I. Davies (Ed.), *Debates in History Teaching*, 2nd ed. London: Routledge, 55–65.

Lee, P. and Ashby, R. (2000) 'Progression in historical understanding among students ages 7–14'. In P. Stearns, P. Seixas and S.S. Wineburg (Eds.), *Knowing, Teaching, and Learning History: National and International Perspectives*. New York: New York University Press, 199–222.

Lee, P. and Howson, J. (2009) 'Two out of five did not know that Henry VIII had six wives: History education, historical literacy and historical consciousness'. In L. Symcox and A. Wilschut (Eds.), *National History Standards: The Problem of the Canon and the Future of Teaching History*. Charlotte, NC: Information Age, 211–65.

Lévesque, S. (2008) *Thinking Historically: Educating Students for the Twenty-First Century*. Toronto: University of Toronto Press.

Lévesque, S. and Clark, P. (2018) 'Historical thinking: Definitions and educational applications'. In S. Metzger and L. Harris (Eds.), *The Wiley International Handbook of History Teaching and Learning*. Hoboken, NJ: John Wiley & Sons, 119–48.

Levich, M. (1965) 'Review of *Philosophy and History: A Symposium* by Sidney Hook'. *History and Theory*, 4 (3), 328–49. [CrossRef]

Mabbett, I.W. (2007) *Writing History Essays: A Student's Guide*. Basingstoke: Palgrave Macmillan.

Mandelbaum, M. (1977) *The Anatomy of Historical Knowledge*. London: The Johns Hopkins University Press.

Martin, R. (1989) *The Past Within Us: An Empirical Approach to the Philosophy of History*. Princeton, NJ: Princeton University Press.

Maza, S. (2017) *Thinking About History*. London: The University of Chicago Press.

McCullagh, C.B. (1969) 'Narrative and explanation in history'. *Mind*, 78 (310), 256–61. [CrossRef]

McCullagh, C.B. (1978) 'Colligation and classification in history'. *History and Theory*, 17, 267–84. [CrossRef]

McCullagh, C.B. (1998) *The Truth of History*. London: Routledge.

McCullagh, C.B. (2011) 'Colligation'. In A. Tucker (Ed.), *A Companion to the Philosophy of History and Historiography*. Chichester: Wiley-Blackwell, 152–61.

Megill, A. (2007) *Historical Knowledge, Historical Error: A Contemporary Guide to Historical Practice*. London: University of Chicago Press.

Michalaki, M. (2021) 'Deepening Year 9's knowledge for better causation arguments: Using greater detail to push students to reason with given causal categories'. *Teaching History*, 182, 36–47.

Navey, M. (2018) 'Dealing with the consequences: What do we want students to do with consequence in history?' *Teaching History*, 172, 40–9.

Norton, M.B. (2002) *In the Devil's Snare: The Salem Witchcraft Crisis of 1692*. New York: Alfred A. Knopf.

Paltridge, P. (2012) *Discourse Analysis*, 2nd ed. London: Bloomsbury.

Paul, H. (2015) *Key Issues in Historical Theory*. London: Routledge.

Pearson Edexcel. (2014) *Sample Assessment Materials Pearson Edexcel Level 3 Advanced GCE in History (9H10) First Teaching from September 2015 First Certification 2017 Issue 1*. London: Pearson. Accessed 16 June 2025. https://qualifications.pearson.com/content/dam/pdf/A%20Level/History/2015/Specification%20and%20sample%20assessments/History-A-SAM-Collation-WEB-ISBN9781446914373.pdf.

Pearson Edexcel. (2017) *Specification Pearson Edexcel Level 3 Advanced GCE in History (9H10) First Teaching from September 2015 First Certification from 2017 Issue 2*. London: Pearson. Accessed 16 June 2025. https://qualifications.pearson.com/content/dam/pdf/A%20Level/History/2015/Specification%20and%20sample%20assessments/9781446914366-gce-2015-a-hist.pdf.

Ray, B.C. (2015) *Satan and Salem: The Witch-Hunt Crisis of 1692*. London: University of Virginia Press.

Retz, T. (2017) 'The structure of historical inquiry'. *Educational Philosophy and Theory*, 49 (6), 606–17. [CrossRef]

Retz, T. (2018) *Empathy and History: Historical Understanding in Re-Enactment, Hermeneutics and Education*. New York: Berghahn Books.

Ricoeur, P. (1984) *Time and Narrative Vol. 1*. Chicago, IL: University of Chicago Press.

Rodker, A. (2019) 'Austin's narrative: An exploratory case study, with Year 8, into what kinds of feedback helps students produce better historical narratives of the interwar years'. *Teaching History*, 174, 8–15.

Rosenthal, B. (1993) *Salem Story: Reading the Witch Trials of 1692*. Cambridge: Cambridge University Press.

Shemilt, D. (1980) *History 13–16: Evaluation Study*. Edinburgh: Holmes McDougall.

Shemilt, D. (2000) 'The Caliph's coin: The currency of narrative frameworks in history teaching'. In P. Stearns, P. Seixas and S. Wineburg (Eds.), *Knowing, Teaching and Learning History: National and International Perspectives*. New York: New York University Press, 83–101.

Skinner, Q. (2002) *Visions of Politics Volume 1: Regarding Method*. Cambridge: Cambridge University Press.

Smith, J. (2024) 'History in the curriculum'. In E. Rata (Ed.), *Research Handbook on Curriculum and Education*. Cheltenham: Edward Elgar Publishing, 353–66.

Smith, J. and Jackson, D. (2021) 'Two concepts of power: Knowledge (re)production in English history education discourse'. In A. Chapman (Ed.), *Knowing History in Schools: Powerful Knowledge and the Powers of Knowledge*. London: UCL Press, 152–76.

Smith, R.B. (2011) 'Historical explanation: From narrative to causation–and back?' *History of European Ideas*, 37 (3), 382–95. [CrossRef]

Stake, R.E. (1995) *The Art of Case Study Research*. London: Sage.

Stanford, M. (1994) *A Companion to the Study of History*. Oxford: Blackwell.

Stanford, M. (1998) *An Introduction to the Philosophy of History*. Oxford: Blackwell.

Starkey, M.L. (1949) *The Devil in Massachusetts*. New York: Anchor Books.

Swales, J. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Thomas, G. (2013) 'From question to inquiry: Operationalising the case study for research in teaching'. *Journal of Education for Teaching*, 39 (5), 590–601. [CrossRef]

Tosh, J. (2006) *The Pursuit of History: Aims, Methods and New Directions in the Study of Modern History*. London: Pearson.

Tucker, A. (2009) 'Causation in historiography'. In A. Tucker (Ed.), *A Companion to the Philosophy of History and Historiography*. Malden, MA: Blackwell, 98–108.

Valentine, W. (2017) 'From road map to thought map: Helping students theorise the nature of change'. *Teaching History*, 167, 30–5.

Van Bouwel, J. and Weber, E. (2008) 'A pragmatist defense of non-relativistic explanatory pluralism in history and social science'. *History and Theory*, 47 (2), 168–82. [CrossRef]

Van Boxtel, C. and Van Drie, J. (2012) '"That's in the time of the Romans?" Knowledge and strategies students use to contextualize historical images and documents'. *Cognition and Instruction*, 30, 113–45. [CrossRef]

Van Boxtel, C. and Van Drie, J. (2017) 'Engaging students in historical reasoning: The need for dialogic history education'. In M. Carretero, S. Berger and M. Grever (Eds.), *Palgrave Handbook of Research in Historical Culture and Education*. London: Palgrave Macmillan, 573–89.

Van Boxtel, C. and Van Drie, J. (2018) ''Historical reasoning: Conceptualizations and educational applications'. In S.A. Metzger and L. McArthur Harris (Eds.), *The Wiley International Handbook of History Teaching and Learning*. Hoboken, NJ: Wiley-Blackwell, 149–75.

Van Drie, J. and Van Boxtel, C. (2008) 'Historical reasoning: Towards a framework for analyzing students' reasoning about the past'. *Educational Psychology Review*, 20, 87–110. [CrossRef]

VanSledright, B. (2010) *The Challenge of Rethinking History Education: On Practices, Theories, and Policies*. New York: Routledge.

Walsh, W.H. (1942) 'The intelligibility of history'. *Philosophy*, 17 (66), 128–42. [CrossRef]

Walsh, W.H. (1951) *An Introduction to Philosophy of History*. London: Hutchinson.

Walsh, W.H. (1974) 'Colligatory concepts in history'. In P. Gardiner (Ed.), *The Philosophy of History*. Oxford: Oxford University Press, 127–44.

Weinryb, E. (1975) 'The justification of a causal thesis: An analysis of the controversies over the theses of Pirenne, Turner, and Weber'. *History and Theory*, 14 (1), 32–56. [CrossRef]

Whewell, W. (1840) *The Philosophy of the Inductive Sciences*, Vol. 1. London: J.W. Parker & Sons.

Whewell, W. (1847) *The Philosophy of the Inductive Sciences*, Vol 2. London: J.W. Parker & Sons.

White, M.G. (1965) *Foundations of Historical Knowledge*. New York: Harper and Row.