

The application of the microgenetic method to studies of learning in science education: characteristics of published studies, methodological issues and recommendations for future research

Abstract

This paper examines the role of the microgenetic method in science education. The microgenetic method is a technique for exploring the progression of learning in detail through repeated, high frequency observations of a learner's 'performance' in some activity. Existing microgenetic studies in science education (including some studies of science learning published in journals not focused on science education) are analysed. This leads to an examination of five significant methodological issues in microgenetic research. Firstly, the fit of the microgenetic method with qualitative and/or quantitative approaches to data collection and analysis is considered and a case is made for the appropriateness of microgenetic research of a qualitative nature. Secondly, it is argued that researchers may define static intervals, periods within which (for methodological purposes) change is assumed not to occur, when reporting microgenetic studies. Thirdly, researchers should consider providing justifications for their choice of sampling rate with reference to the rate of change of the phenomenon they are studying. Fourthly, the difficulty of distinguishing conceptual change from the existence of multiple understandings is highlighted. Finally, the nature of sequences of repeated measures in microgenetic studies is considered. It is argued that different methodological approaches are suitable for microgenetic studies of different phenomena. The paper concludes with a list of guidelines for the use of the microgenetic method in small-scale, qualitatively analysed studies in science education.

Introduction

Though learning has been conceptualised in a variety of ways, models typically include a description of change over time (Lachman, 1997, p. 479; Marton, 1983, p. 291; Mowrer & Klein, 2001, p. 2). Therefore, assessments of learning, in the context of science education and also in other disciplines, require data collection at multiple points in time in order to develop an understanding of change. In particular, learning in science education is often modelled as conceptual change (Duit & Treagust, 2003; Rusanen & Lappi, 2013, p. 3332), the process by which one concept or set of concepts is altered over time (Chi & Roscoe, 2002, p. 4; Posner, Strike, Hewson, & Gertzog, 1982, p. 211; Rusanen & Lappi, 2013, p. 3332). One common approach to investigating conceptual change is to employ longitudinal studies which are based on reports of students' learning at a small number of points over time (White & Arzi, 2005). Longitudinal studies make use of repeated applications of assessments, for example, concept maps (Novak & Musonda, 1991), concept inventories (Caballero et al., 2012) or interviews (Taber, 2001). Longitudinal approaches, however, may fail to develop a complete representation of change because learning is a 'messy' process (Taber, 2013, p. 216), that is, change may occur over a number of timescales (Blown & Bryce, 2006; Clement, 2008; Gilbert & Watts, 1983; Thornton, 1997), involve changes to multiple co-existing understandings (Harrison & Treagust, 2000; Mortimer, 1995; Taber, 2000b) and fail to reach a stable state even during extended periods of observation (Shuell, 1990, p. 531). In studies

that seek to represent change, a researcher's choice of the frequency with which data is sampled, relative to the rate of change of the phenomenon being studied, will influence how a data set is constructed. This paper examines a particular approach towards the collection of data related to change, the microgenetic method.

In order to develop a faithful representation of change, the microgenetic method, which will be defined in detail below, involves the collection of data at a frequency which is considered to be high compared to the rate of change of the phenomenon being studied (Siegler & Crowley, 1991, p. 606). The method therefore offers a way of exploring learning that goes beyond the pre-test/post-test paradigm that simply asks whether there is an observable change that might be taken as evidence of learning between two assessments, and investigates the patterns of change over time as a learner engages in some activity.

Consider the case of a researcher investigating conceptual change in a particular topic. They might probe a student's understanding at two points in time, several months apart, and draw inferences about learning from differences between the assessments. Alternatively, a researcher adopting the microgenetic approach might seek to sample the student's understanding at a frequency that is high compared to the rate at which they assume conceptual change occurs. For example, if they assumed conceptual change occurred over many months, they might argue that weekly assessment, over a period of several months, would represent a high density of sampling relative to changes in understanding. Though the longitudinal study can describe the start and endpoints of conceptual change, it cannot shed light on the patterns of change. The microgenetic study therefore, as well as providing information on what students know, can also describe 'how they got there' (Goldin-Meadow & Wagner Alibali, 2002, p. 94). A more traditional longitudinal approach is also unable to distinguish genuine changes in understanding from the situation where a learner holds manifold conceptions and different responses are elicited some time apart from a wider repertoire of available and (to the learner) feasible responses. The detailed data developed by the approach, which has been described as 'untidy' (Flynn, Pine, & Lewis, 2007, p. 4), can be both a benefit and present a challenge to the researcher as it may contain moments of sudden change, multiple advances and regressions, or evidence of multiple co-existing positions. This paper sets out to examine the application of the microgenetic method to studies in science education and considers the manner in which the approach constructs data.

The word microgenetic was coined by Werner (1956, p. 347) to describe the 'unfolding' of a 'human activity such as perceiving, thinking, acting etc.' over timescales ranging from seconds to days. The term genetic here refers to the origin or genesis of ideas, rather than sections of chromosomes (Siegler, 2006, p. 471). This usage is familiar from the genetic epistemology programme of Jean Piaget (1970), that has been very influential in science education (Bliss, 1995). Werner's work arose from ideas developed by the *Ganzheitspsychologie* group led by Friedrich Sander (Rosenthal, 2004, p. 221). Sander (1930, p. 192) argued that '[o]nly a further and further dissection, and a destructive analysis, can ever arrive at those disconnected pieces which the old psychology designated as elements.' Sander's notion was taken up by Vygotsky (1930/1978, p. 61) who maintained research should include a focus on processes that occur on short timescales and processes of change as '...it is only in movement that a body shows what it is.' As it was adopted by researchers in

developmental psychology, the usage of the term microgenetic has shifted, to refer to studies of the variation in the use of strategies or skills over time (Siegler & Crowley, 1991, p. 606). For example, microgenetic studies have sought to investigate the development of: number conservation (Siegler, 1995), analogical reasoning (Cheshire et al. 2007) and motor skills (Adolph et al., 2008). Flynn, Pine and Lewis' (2006) paper, provides a good introduction to the method in the context of developmental psychology. A number of studies published in the developmental psychology literature focus on the development of skills and strategies related to learning about science and may be of interest to researchers in science education (For example, Opfer & Siegler, 2004; Van Der Steen, Steenbeek, Van Dijk, & Van Geert, 2014). In addition to developing a set of recommendations for researchers describing change in science education, this paper also seeks to highlight a number of significant studies describing changes in students' learning of science that have been published outside science education research journals.

The microgenetic method may be relevant to research in science education for three reasons. Firstly, investigating conceptual change is a central research programme of science education (Treagust & Duit, 2009, p. 90) and the microgenetic method is well suited to observing change as it occurs (Kuhn, 1995, p. 133). Secondly, investigating cognitive structure (i.e. the hypothetical underlying structure of representations of the world that facilitates a person's active conceptualisations) is challenging as learners may hold manifold conceptions of a topic (Harrison & Treagust, 2000; Mortimer, 1995; Taber, 2000b). Microgenetic investigations of a person's conceptualisations, over extended time periods, may be able to differentiate between substantive conceptual change and the existence of multiple conceptions (see the discussion of the concept of the static interval, below). Finally, though high-density sampling may produce data that appear highly variable and 'untidy' (Flynn et al., 2007, p. 4), such data may allow researchers to begin to distinguish 'mental flotsam and jetsam' (Taber, 1995, p. 5) from more stable constructions.

Following a search of the literature (the search procedure is described in detail in the microgenetic method in science education section, below), we identified around thirty studies, which focus on contexts related to science education and whose authors claim to have used the microgenetic method. As will be discussed later in the article, these studies are mostly small-scale in terms of number of participants (N). An examination of these studies leads to five questions, which arise from an examination of learning using high-density sampling. These questions will act as the basis of the discussion in the body of this article.

- a) Can the microgenetic approach be used within research designs collecting and analysing either, or both, qualitative and quantitative data?
- b) When examining data that describe changes in students' representations of their thinking how can researchers define intervals over which change in underlying cognitive does and does not occur?
- c) What is an appropriate sampling rate for a microgenetic study of learning?
- d) How is it possible to manage the high variability that may occur in data sampled at a high frequency?

- e) What kinds of sequences of probes are appropriate within the microgenetic approach?

The microgenetic approach is well established for the study of strategy acquisition, that is learners' development of potentially conscious and controllable activities that achieve cognitive purposes (Pressley, Forrest-Pressley, Elliott-Faust, & Miller, 1985). For example, a strategy when approaching tasks that require the addition of two relatively small numbers is to begin at the lower number and count up the number of integers that matches the larger number. However, the use of the microgenetic method to study the kind of conceptual learning that is often the focus in science education, is less well established and requires alternative approaches (Parnafes & diSessa, 2013, p. 15). This review leads to a series of recommendation for researchers applying the microgenetic approach to learning in science education. The next section examines the definition of the microgenetic approach and examines what criteria may be appropriate for the application of the term in science education.

The microgenetic method

The microgenetic method has been described as 'an observation of change whilst it occurs' (Kuhn, 1995, p. 133) which allows researchers to investigate not only students' knowledge, but the processes of change. The microgenetic approach has been defined by three characteristics:

- (a) Observations span the entire period from the beginning of the change to the time at which it reaches a relatively stable state.
- (b) The density of observations is high relative to the rate of change of the phenomenon.
- (c) Observed behaviour is subjected to intensive trial-by-trial analysis, with the goal of inferring the processes that give rise to both quantitative and qualitative aspects of change. (Siegler & Crowley, 1991, p. 606)

However, in practice, the categorisation of studies as microgenetic is 'somewhat subjective' (Siegler, 2006, p. 479). A range of different types of study fit the approach: Miller and Coyle (1999, p. 210) argue that both a single thirty-minute session, containing multiple probes, and multiple sessions, spanning several months, may be considered microgenetic. To give a sense of the microgenetic approach, there follows a brief description of two contrasting studies in science education that are described by the authors as microgenetic.

Garcia-Mila and Andersen (2007) investigated how children and adults made notes whilst undertaking scientific enquiry. Whilst completing a ten-week programme of practical work, the participants kept notes in a journal and were interviewed, twice a week, to understand their approach to note-taking. This method, the authors argue, allowed them to 'observe progress from repeated practice' and so develop 'a fine-grained assessment of strategy use' (Garcia-Mila & Andersen, 2007, p. 1040). Van Der Steen and colleagues (2014) conceptualised their study of one 4-year-old boy's developing understanding of air pressure as both longitudinal and microgenetic (for a discussion of the relationship between these methods, see the section on sampling rate, below). The boy was allowed to interact with three different pieces of apparatus consisting of syringes filled with air, and questioned about their functioning, over

three sessions spaced three months apart. The transcribed interviews were coded for the level of complexity of the boy's answers, and graphs representing the second-by-second fluctuation in inferred complexity in each of the three sessions were plotted. The authors claim this method allowed insight into both the short-scale variability and longer-term development of understanding of air pressure. These two examples illustrate how differing approaches may be utilised within a microgenetic framework and show that researchers may select different timescales for both observations made within sessions, and the schedule of data collection sessions.

A number of approaches to sampling students' understanding of science at multiple occasions over time have been described and should be distinguished from microgenetic research. Learning process studies are a type of research which develops representations of a sequence of conceptions of a particular topic over time (Niedderer, Budde, Givry, Psillos, & Tiberghien, 2007). A related type of studies, learning pathway studies, reports the development of a particular conception over time. For example, Petri and Niedderer (1998) described a sequence of four conceptions of the atom a student transitioned through over the course of 16 weeks of instruction. Alternatively, an approach in which a student was interviewed 23 times over the course of two years (Taber, 2001) was conceptualised as a case study of learning. Such approaches can develop models of how learners' understandings transition through a series of different models, referred to variously as conceptual trajectories (Driver, Leach, Scott, & Wood-Robinson, 1994), learning progressions (Berland & McNeill, 2010) and developmental trajectories or corridors (Brown, 2004, pp. 84–85). However, none of these approaches specify a sampling rate; the unique feature of microgenetic research is that it requires more of the researcher than the collection of data at multiple points in time. The method entails the additional constraint that the rate of sampling is high compared to the rate of change of the phenomenon of interest (Siegler & Crowley, 1991, p. 606) in order to develop a representation of the processes of change, rather than snapshots of the progression (Flynn et al., 2006). Microgenetic studies may take many forms, but they must aim to collect data with a sufficiently high frequency to develop a construction of the trajectory of change.

The term microgenetic has been used to describe a number of different phenomena or processes. Wertsch (1991, p. 33) observed that Vygotsky used microgenetic to refer to both the short-term unfolding of a single psychological act and also to a series of developmental changes that occur over the course of an experimental session. Here the term microgenetic is used to label a methodological approach for investigating learning, rather than for what is being investigated (so the potential ambiguity of the term in Vygotsky's work is not an issue). A separate distinction separates the usage of microgenetic to refer to psychological models of short timescale perceptual or cognitive processes, from its use as description of a methodological technique for deliberately intervening in learning which accelerates the course of a developmental sequence in a clinical setting (Rosenthal, 2004, p. 221). For example, the process of learning the skill of tying shoelaces, which may take several weeks in typical development, might be 'miniaturised' into sessions over a few days of intensive trials. For Catán (1986, p. 260), such acceleration remains a defining feature of microgenetic research; however, we will argue below (see section d, the noisiness and stability of data) that it is not a necessary component for studies in science education. To avoid such ambiguity, in this paper, microgenetic will be

used to refer to an experimental technique, which meets, or attempts to meet, Siegler and Crowley's (1991, p. 606) criteria. The next section examines some of the assumptions that may be made when data relating to cognition are constructed as a series of events over time.

The assumption of a serial flow of data: some epistemological considerations

Before proceeding, we wish to briefly outline a key assumption that informs the present review. We assume that this area of research needs to distinguish between the thinking a person is experiencing at some particular time (and which may be directly represented in their responses in research interviews for example) and the available conceptual resources that support that thinking (Taber, 2013). That is, we consider that explicit thinking (something that is of limited focus and in constant flux) is resourced by something more extensive, and - by comparison - relatively stable. This is what has been called cognitive structure (White, 1985), and might be related to the notion of the contents of long-term memory. Components or aspects of that underlying structure are accessed ('brought to mind') or activated during reflective thinking, and sometimes there may be multiple alternative potential resources that could be accessed to think about a particular issue, problem or question. In these cases it is quite possible only one of the viable alternatives will be cued on any particular occasion.

We use the term conceptual change to refer to some substantive change (e.g., in contents and/or structuring) in this underlying substrate of mental resources. Even a difference in response to precisely the same stimulus question will not necessarily imply conceptual change, as the internal mental context in which particular resources are activated or not may depend upon many factors outside the control and knowledge of a researcher. Research into conceptual change is therefore complicated because a change in what a person is thinking (which is what a researcher can hope to directly infer by interpreting data elicited at any one time) from one time to another, may, or may not, reflect a substantive change in the underlying cognitive structure (which is only partially and less directly reflected in research data).

Data used in science education research to explore cognition are constructed representations, and imperfectly reflect the nature of cognition (Taber, 2013). For example, though cognition may be parallel (Rumelhart & McClelland, 1986), many forms of data have an apparent sequential nature, for example the utterances in interviews and even concept maps are produced as, and might be read as, serial chains of concepts interlinked among themselves. The serial nature of data can be seen as arising from processes occurring at three levels, shown in figure 1.

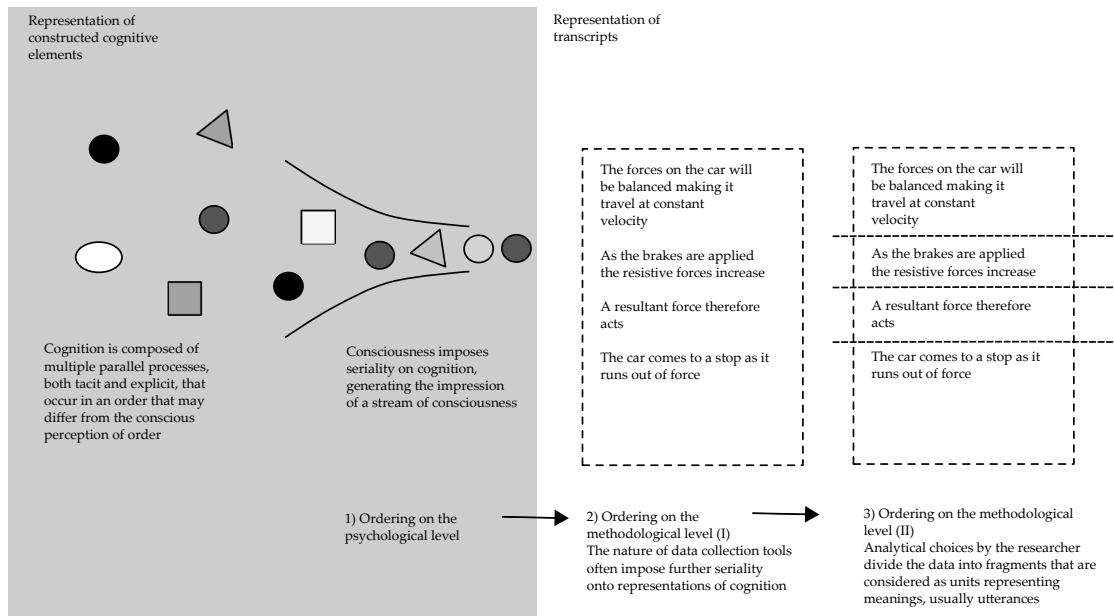


Figure 1: The channelling of cognition into serial-in-time reports

Cognition is complex (O'Brien & Opie, 1998, p. 13): processing may be parallel (Rumelhart & McClelland, 1986), involve tacit elements (Brock, 2015; diSessa, 2000; Taber, 2014) and the conscious experience of the order of cognitive processes over short timescales may be misleading (Dennett, 1991, pp. 168–169). The stream of consciousness experienced in the lifeworld is more than a single stream of information (Blackmore, 2002): Dennett (1991, pp. 253–254) argues conscious experience does not consist of a unitary flow of data, rather different systems create multiple drafts of information from stimuli, which are processed by different regions of the brain in quick succession. Consciousness imposes a sense of seriality onto the parallel nature of cognition (Baars & Franklin, 2003, p. 167), as illustrated in the left hand section of figure 1.

In addition to psychological ordering, research instruments cause a further layer of sequencing (right hand side of figure 1) in the sense that the particular stimuli presented, and the order and form in which they are posed, act as cues for what a respondent 'brings to mind'. However, what a respondent 'brings to mind', which knowledge representations are activated, and in what order, is also in part determined by an internal mental context (put simply, what they have been recently thinking about). Probes of consciousness at different times will then produce different narratives: cognition at a given time is partly (but not wholly) dependent on the nature of the stimuli encountered (Dennett, 1991, pp. 135–136). This presents a methodological difficulty: it is assumed a learner may possess simultaneously existing understandings (Mortimer, 1995), but reports of their existence will be sequential.

In the final stage of ordering, a researcher may divide the data into sections representing meanings, often at the level of utterances. A number of such units of meaning, for example a subsection of an interview related to a particular probe or an entire interview, may be considered to represent a single moment in time in the development of a student's thinking, a static interval in the sequence of data. We define a static interval as a section of data over which it is assumed change is not taking place. In some studies researchers have (explicitly or implicitly) assumed a

single interview, involving multiple probes, can be considered as a single static interval (e.g. Taber, 2008b), whilst others (e.g. Parnafes & diSessa, 2013) have chosen to define the utterances related to a series of probes within a single interview as separate static intervals and sought to observe changes between them (a detailed discussion of these differing constructions of change occurs below). In practice, the nature of the human cognitive system is such that each activation or 'recall' of a stored representation in memory is intrinsically also a modification of that representation (Alberini, 2011; Nader, Schafe, & LeDoux, 2000; Taber, 2013), so although it is methodologically necessary to make a judgement about what counts as a static interval in a particular study prior to analysis, the processes that underpin conceptual change are active during any period of data collection that requires a participant to engage in relevant cognition. The three levels of temporal ordering are accessible and controllable by the researcher to varying degrees (see table 1).

Level	Description
Student's cognitive processing	Cognition is perceived as a series of mental states, but may have involved parallel processing. There is no direct access to the processes at this level.
Student's metacognitive processing	In response to probes, the student will make explicit and tacit decisions about the ordering of ideas. A student's metacognitive comments may allow the researcher to construct some understanding of this processing. For example, a student may be able to explain consciously available aspects of their choice of a particular conception from multiple alternatives they possess.
Researcher's choice of probes and interpretations of student's conceptions	The researcher makes reasoned decisions regarding the nature and sequencing of probes and the division of data, which influence the constructed sequence of the representation of the student's cognition.

Table 1: Different levels of temporal ordering

In this paper, we will focus on the third level of temporal ordering, which is explicitly controlled by the researcher. On this level, the term conception refers to a researcher's construction of a student's personal expressed understanding at a given time and concept refers to formal meanings in public knowledge (Gilbert & Watts, 1983, p. 69; Taber, 2013, p. 283). The seriality of reports of cognition has received little attention, perhaps because the notion becomes increasingly significant when data is sampled at relatively high frequencies, as occurs in the microgenetic method, and so seriality has not tended to be a concern in studies that have explored cognition at one point in time, or longitudinally with a few temporally well-separated measurement points. In order to examine the ideas raised in this section in the context of science education, the following section examines those studies in science education that have claimed to make use of the microgenetic method.

The microgenetic method in science education

In creating a catalogue of microgenetic studies in science education (table 2) the following criteria were used:

- Studies that self-define as microgenetic were included, regardless of the absolute value of the frequency of observations. Some studies that are defined as longitudinal may have relatively short intervals between observations (e.g. Pearsall, Skipper and Mintzes (1997) used a spacing of four weeks between observations): such studies have not been included where their authors do not describe them as microgenetic. A characteristic of microgenetic research is a claim for a high density of observations relative to the rate of change of the phenomenon being investigated, rather than an absolute claim of high-density sampling. This study aims to examine different authors' conceptualisations of microgenetic research in science education and therefore the principal criterion for inclusion is an assertion that the research was carried out using the approach, rather than a specific frequency of sampling.
- The selection strategy proposed by Cheng and Wan (2015) was used to find articles. Nine science education journals with high impact factors (*The Journal of Research in Science Teaching*, *Science Education*, *Chemistry Education Research and Practice*, *The International Journal of Science Education*, *Research in Science Education*, *Science & Education*, *The International Journal of Science and Mathematics Education*, *The Journal of Science Education and Technology* and *Research in Science & Technological Education*) were searched for the key-word "microgenetic." Studies that claimed to make use of data collection methods based on microgenetic principles were included in the analysis. Cheng and Wan (2015) recommend that studies referred to in the initial set of papers are then examined to discover additional publications. This step was undertaken and several microgenetic studies in the context of science education, but not published in journals primarily focused on science education, were included in the sample. For example, Parnafes and diSessa's (2013) research is a microgenetic study of the development of understanding of simple harmonic motion, but appeared in the journal *Human Development*. These studies are included as they met our selection criteria for the present review. We also hope that by including such studies they might be brought to the attention of colleagues in science education who would find them relevant but do not normally read the journals in which they were published.
- Only studies related to science education in school or university contexts were included. For example, the context of a research laboratory in a study by Roth (2014) was not deemed relevant to this catalogue.
- Apart from one PhD thesis (Chinn (1997)) that has not been published elsewhere, the catalogue of articles consists of peer-reviewed studies.

Table 2: Microgenetic studies in science education. Where details are unclear in the original study, this is indicated in the table. In cases where authors have described change as occurring over multiple tasks in one session (E.g. Opfer & Siegler, 2004), those tasks are defined as separate observations. The papers are listed by year of publication.

Author(s)	Number and age of Participants	Phenomenon being studied and context	Length of Study	Number of observations	Spacing of observations	Length of individual observations	Type of observation
Kuhn & Phelps, 1982	15 students aged 9-11 years old	Problem solving in chemistry	12 weeks (including 1 week of vacation)	11	1 week	Unclear	Questions based on interactions with practical equipment
	15 students aged 9-11 years old	Problem solving in chemistry	14 weeks	Up to 13 sessions but participants left study after 8 sessions if stability in strategy use achieved	1 week	Unclear	Questions based on interactions with practical equipment
Kuhn & O'Loughlin, 1988 (Study 5)	20 students aged 8-12 years old	Evaluating scientific evidence	9 weeks	9 sessions	1 session per week	30-45 minutes	Task involving evaluating evidence related to different types of balls, participants interviewed whilst completing task
Kuhn, Schauble, & Garcia-Mila, 1992	12, 10-year old, fourth graders	Theory and strategy revision about factors affecting the speed of cars and boats	9 weeks	19	2 sessions in a week	20-30 minutes for car domain 30-45 minutes for boat domain	Problem solving sessions with practical equipment or computer simulation, interview sessions
	20 fifth-sixth graders	Scientific reasoning about factors affecting the speed of cars and balls	7 weeks	7	1 week	30 minutes	Problem solving sessions with practical equipment or computer simulation, interview sessions

Nuthall & Alton-Lee, 1993 (Study 1)	3 students aged 9-10 years old	Learning about conservation, erosion and endangered species	31 days	Transcript divided into 15 second long sections-unclear as to total number or spacing	Transcript divided into 15 second long sections-unclear as to total number or spacing	Total of 129 hours- unclear as to individual duration or spacing	Continuous observation, video and audio recorded. Pre-test then post-test and interviews after observations, and again 12-months after observation
Johnson & Mervis, 1994	16 five-year-old students	Knowledge of shorebirds	17 days	5	Four, one hour long sessions at intervals of 3-5 days, 30 minute session within 2 days of fourth session	Four, one hour sessions and one 30 minute session	Tests of knowledge, triad task, general sorting task
	1 four-year-old boy	Knowledge of shorebirds	17 days	5	Four, one hour long sessions at intervals of 3-5 days, one 30 minute session within 2 days of fourth session	Four, one hour sessions and one 30 minute session	Tests of knowledge, triad task, general sorting task
Zohar, 1995	25 students at community college, mean age, 32 years	Reasoning about variables	10 weeks	20	Two sessions per week	30 minutes	Five tasks on reasoning about variables, recordings of sessions
Magnusson, 1996	8 fourth grade students	Learning about sound	3 months	Unclear	Weekly or biweekly	10 minutes	Classroom activities videotaped and observed, student presentations videotaped, individual interviews and post intervention interviews
Schauble, 1996	10 fifth and sixth graders 10 unrelated adults	Scientific reasoning about objects immersed in fluids and placed on springs	2 weeks	6	Six sessions over 2 weeks	40 minutes	Interviews at start and end of sessions, data record cards from practical sessions
Chinn, 1997	61 sixth and seventh grade students	Knowledge about molecules and chemical	6.5 weeks	13	Two sessions per week	60-80 minutes	Guided interviews with instructors involving engagement with experiments and texts. Think aloud

		reactions					protocols produced
Duit, Roth, Komorek, & Withers, 1998	25 tenth grade students	Conceptual change related to chaotic systems	2 weeks	4	Unclear	90 minutes	Pre-test, students interactions with experiments and simulations videotaped
Hogan, 1999	12 eighth grade students	Personal frameworks related to the nature of matter	12 weeks	Unclear	Classes recorded two or three times a week	Unclear	Interviews before and mid way through course. Classes video and audio taped
Nuthall, 1999 (Study 6)	5 students (average age 11.8 years)	Knowledge of the habitat of Antarctica	6 days	Unclear- 13.4 hours of observation over 6 days	Unclear- 13.4 hours of observation over 6 days	Unclear- 13.4 hours of observation over 6 days	Written multiple choice pre- and post-tests, classroom observation with video-cameras, record of students writing in class, records of students homework, interviews at end of observations and long-term (12-month later) post-test and interview
Chinn, O'Donnell, & Jinks, 2000	105 fifth grade students	Argument structure in group work on electrical circuits	2 sessions over 1 day	2	2 sessions over 1 day	50-60 minutes	Students carried out experiments, wrote conclusions and evaluated pre-written conclusions. Discussions were recorded and transcribed
Izsak, 2000	24 eighth-grade students (students worked in pairs)	Knowledge structures related to the winch	3-4 weeks	3-4	1 week	60 minutes	Problem solving, with physical apparatus, videotaped
Azmitia & Crowley, 2001	24 undergraduates	Scientific thinking in the context of building towers to withstand earthquakes	1 week	6 sessions (4 individual, 2 collaborative. Individual sessions preceded and followed collaborative sessions)	1 week	15 minutes	Questioning, and collaborative sessions videotaped

Roth & Welzel, 2001	8 tenth grade students	Use of gestures and scientific discourse in the context of electrostatics	10 weeks	20	2 per week	45 minutes	Videotaped lessons, pre- and post-interviews, and written tests in week 4, 10 and 15
Wiser & Amin, 2001	4 students just finished eighth grade	Conceptual change in the domain of thermal physics	5 weeks	Unclear	Several mornings a week	2 hours	Teaching-learning sessions audiotaped. Individual interviews were conducted before, during, immediately after, and 6 months after the teaching sessions
Gelman, Romo, & Francis, 2002	22 ninth grade, ESL students	Conceptual and language learning across various science topics studied by ESL students	1 academic term	10 units	Unclear	Unclear	Students' writing and concept maps in notebooks based on experimental sessions
Eichler, Del Pino, & Fagundes, 2004	8 students aged 14-17 years old	Conceptual development linked to air pollution	As many sessions as required to complete the task	Unclear	Unclear	45 minutes	Logfiles from engagement with computer simulation, texts written whilst using simulation, audio transcripts recorded during use of simulation.
Opfer & Siegler, 2004	80 kindergarten students aged 5-6 years old	Conceptual change in the categorisation of living things	1 session	3 tasks in 1 session	3 tasks in 1 session	Unclear	Pre-test, post-test categorisation tasks, questioning
Veal, 2004	2 prospective secondary chemistry teachers	Pedagogic content knowledge	1 year	Unclear	Observations, daily for 1 hour, 5 days a week for first semester. Vignettes shown to participants once every 3 weeks	Unclear	Pre, during and post interviews Observations of teacher, teachers' journals and responses to vignettes
Feldon & Gilmore, 2006	154 students (52 members of sixth grade science class, 42	Scientific problem solving in the context of infectious	Unclear and user dependent	2 Sessions (possible multiple uses of simulations	Unclear and user dependent	Unclear and user dependent	Data from interaction with computer simulations

	free-choice users of online site)	diseases		by participants in each session)			
Pata & Sarapuu, 2006	53 Secondary students aged 15-17 years old	Reasoning processes in the context of genetics	Unclear	Use of collaborative virtual workshop divided into 4 phases	Unclear	Unclear	Pre-essay, post-essay, discussion whilst using collaborative virtual workshop
Garcia-Mila & Andersen, 2007	15 fourth grade students and 16 community college students aged 22-47 years old.	Developmental change during note taking in scientific inquiry	10 weeks	20	2 in a week	30-45 minutes	Students carried out various practical, computer and paper-based tasks. Students' notes in notebooks collected
Soong, 2008	37 students aged 15-16 years old	Computer mediated collaborative physics problem solving	9 weeks	4	Weeks 1, 2 and 7, 8	1.5 hours	Computer based problem solving. Data taken from students' chat logs
Kuhn, 2010	40 sixth grade students	Scientific argumentation skills	7-8 weeks	13	2 in a week over 7-8 week period	Unclear	Students engaged in argumentation via a computer-based system, reflection sheets
Parnafes & diSessa, 2013; Parnafes, 2007, 2010 <i>(Note: Data from one study analysed in three papers)</i>	16 students aged 14-18 years old. Students worked in pairs	Learning, and the development of epistemological complexity related to simple harmonic motion	1 session, 1.5 hours long	3 sections within 1 session	First two sections last 25-40 minutes. Final section 10-15 minutes	First two sections last 25-40 minutes. Final session 10-15 minutes	Students videotaped interacting with physical oscillators, then with computer simulations and finally a discussion with researchers about the main conceptual issues
Garcia-Mila, Andersen, & Rojo, 2011	34 sixth graders aged 11-13 years old	Laboratory record keeping in the context of plant growth	4 weeks	7	Twice a week	30-45 minutes	Tests of content knowledge in first and final sessions, questioning during sessions, students' notes

Srivastava & Ramadas, 2013	5 students, aged 17-19 years old, in the 1 st year of a bachelor degree course	Understanding of the 3D nature of DNA	9 days	6	Unclear	1-1.5 hours	A clinical interview-cum-teaching sequence was videotaped. The students drew diagrams and engaged with models.
diSessa, 2014	1 grade 8-10 student	Construction of causal schemes related to cooling curves	2 sessions	Change within one session and into start of next	2 days	3 hours	Students carried out experiments, produced a computer model and were videotaped
	Unclear but multiple, grade 8-10 students	Construction of causal schemes related to cooling curves	1 session	Change within 1 session	Change within session	22 minutes	Students were videotaped in discussions around drawing a graph following a practical experiment
Van Der Steen, Steenbeek, Van Dijk, & Van Geert, 2014	1 boy aged 4-years old	Understanding of air pressure	6 months	3	3 Months (Change also assessed within sessions)	15 minutes	Practical tasks involving syringes were videotaped
Berland & Crucet, 2016	2 fifth/sixth grade students	Epistemological sophistication related to plate tectonics	1 session of 90 minutes	Change within session	1 session	Variable	Students' model construction was videotaped, post interview
Ha, Lee, Lim, & Yang, 2016	9 sixth grade students	Understanding of the particle model	6 days	6 sessions	1 day	Unclear	Students were videotaped in semi-structured interviews creating models
Haglund, Jeppsson, & Schönborn, 2016	46 students aged 9-11 years old (Analysis is of subgroup of students)	Understanding of heat	9 minutes 35 seconds	4	Session divided into 4 episodes- length unclear	Unclear	Practical tasks and class presentations and demonstrations videotaped

The next sections address five issues related to microgenetic studies arising from an examination of the studies in table 2: a) appropriate frameworks for microgenetic research; b) claims regarding change and the static interval; c) sampling rate and the rate of change of learning; d) noisiness and stability of data; and e) appropriate forms of repeated measures.

a) Appropriate frameworks for microgenetic research: the case for small-scale qualitative microgenetic research

The papers listed above provide a context in which to examine how the microgenetic method has been understood in science education. One author who has commented on the use of the approach, Chinn (2006, pp. 443–444), made a case that certain kinds of small-scale, qualitative, repeated measures studies are not microgenetic for the following reasons:

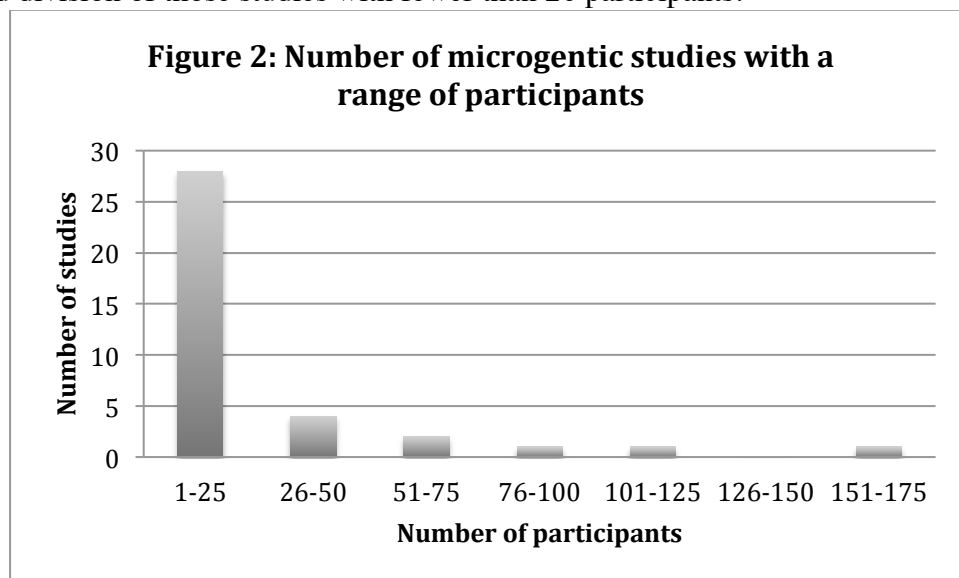
- 1) The sampling frequency in the studies is too coarse to detect change and/or data is not reported between the pre- and post-test.
- 2) Data collected in the studies are not analysed on a moment-by-moment basis.
- 3) Probes in the studies are not counterbalanced (a technique used to reduce practice and task effects by changing the order of the probes presented to individual participants (Gaito, 1961, p. 46)).
- 4) The sample size in the studies is too small for statistical analysis.

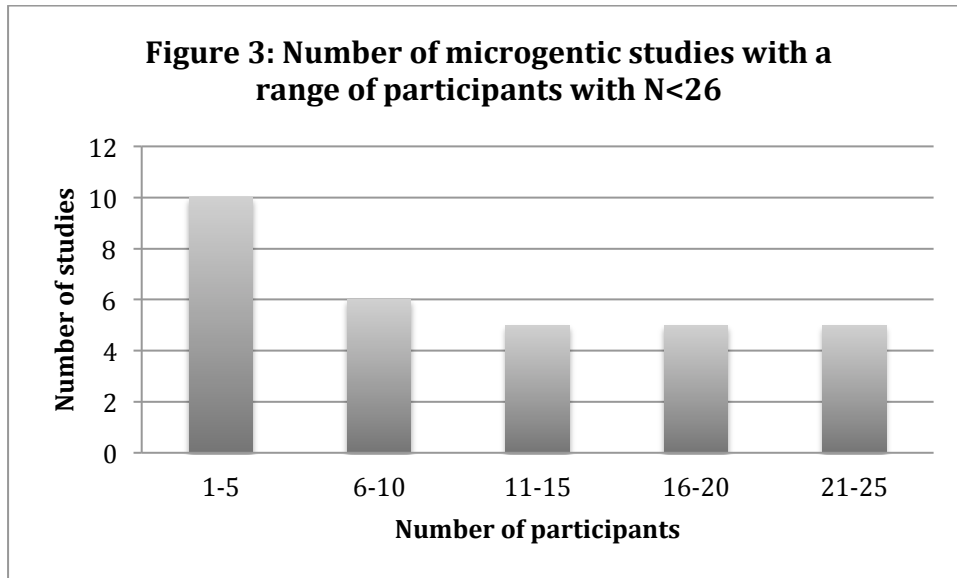
Whilst Chinn has identified some useful indicators for including studies in the microgenetic canon, we argue his indicators should not be considered as absolute criteria and a more nuanced evaluation is appropriate. The criticisms relating to sampling rate, moment-by-moment analysis and counterbalancing of tasks are not inherent issues of small-scale, qualitative studies and may equally be applied to some quantitative, microgenetic studies. These issues are considered further below, after a case is made for the value of the microgenetic approach in studies designed to collect qualitative data from small numbers of participants. The remainder of the paper addresses four themes that become relevant when small-scale, qualitative, microgenetic studies are carried out, including issues relating to Chinn's first three criticisms.

Parnafes and diSessa (2013) have argued that the microgenetic method fits well with a case study approach: it is expected that high-density sampling will lead to '... a high degree of individual and contextual variation' (Parnafes & diSessa, 2013, p. 7) and the case study approach is described as being sensitive to such idiosyncrasy (Yin, 1981, p. 59). Small numbers of participants are usual in microgenetic research, even single participant studies are not uncommon (Siegler, 2006, p. 472). Though small-scale studies have been criticised for their limited generalisability (Feldon & Gilmore, 2006), Parnafes and diSessa (2013, p. 7) argue that microgenetic case studies may have stronger ecological validity and develop greater insights than strictly controlled, laboratory-based studies. A focus on individual variability means '...it makes little sense to average performance over individuals' (Kelso, 1995, p. 161) and, rather than larger sample size, the high frequency of observation in microgenetic research, may 'minimize measurement error' (Lee & Karmiloff-Smith, 2005, p. 257).

Siegler (2006) suggests researchers balance the number of participants against the number of sessions; if the number of participants is low then a high density of observations is required and vice-versa. For example, a study by Robinson and Mervis (1998) in which a single participant was observed for more than 300 sessions, epitomises Merriam's (1995, p. 57) claim that in qualitative research the goal is 'to understand the particular in depth, rather than finding out what is generally true of many.' That is, the research tends to be idiographic rather than nomothetic in nature (Gilbert & Watts, 1983). Arguably, in designing microgenetic research, the number and frequency of observations of an individual should be planned according to the requirements of the method (i.e., sufficient to span likely changes, and frequent enough to document changes occurring, see below), and once this is established the numbers of participants may be considered based on the purposes of, and the resources available, for the study.

The mean number of participants in the studies listed in table 2 is 25 and the mode is 15. The frequency distributions of participants in these studies are shown in figures 2 and 3. Note that Kuhn and Phelps (1982), Kuhn, Schauble, and Garcia-Mila (1992) and Johnson and Mervis (1994) consist of two discrete studies, involving different numbers of participants and those studies have been counted separately. Kuhn (2010) and one of the cases in diSessa (2014) could not be included in these graphs, as the number of participants is not stated. Note that figure 3 represents a fine-grained division of those studies with fewer than 26 participants.





The number of participants in microgenetic studies leads to some distinctions in analytical approach: in studies with high N (for example, Chinn et al., 2000; Feldon & Gilmore, 2006; Opfer & Siegler, 2004) analysis tends to be quantitative. In smaller scale studies, a greater variety of analysis types are found: largely qualitative analysis (Wiser & Amin, 2001), mainly quantitative approaches (Van Der Steen et al., 2014) or a mixture of both (Nuthall, 1999). Clearly, different approaches are appropriate for different kinds of investigation, depending upon the precise research questions posed in particular studies. Chinn's (2006, p. 444) argument against qualitative microgenetic studies might be valid if statistical generalisability were the only goal of research. However, qualitative studies are typically not aiming for statistical generalisability but aim instead to provide in-depth detail about particular cases (Taber, 2000a). Provided researchers undertaking small-N, microgenetic studies follow the recommendations of qualitative methodologists to address issues of reliability and validity (see Conclusions, below), the emphasis in-depth studies place on the richness of individual experience may be an excellent fit for the microgenetic approach.

b) The static interval in microgenetic research

The microgenetic method's focus on the processes of change requires researchers to describe the interval over which change is considered to occur. It is common in analysing qualitative data, such as transcripts of interviews, classroom dialogue, think aloud protocols and the like, for the analyst to divide up the transcript into what are perceived (by the analyst) as a sequence of somewhat distinct episodes in the data that can each be coded in their own terms and then considered as parts of a sequence (see figure 4). This fragmentation of the data set into more manageable sections is usually undertaken based on the analyst's interpretation of shifts in the narrative being constructed from the data. Often, the underlying assumption is that the sequence of episodes refers to one point in time in the sense of a particular stage in the ongoing history of a learner or group. In microgenetic research, however, a researcher may assume that significant change may occur during data collection such that a transcript may be a record of several distinct phases in the development of the

phenomenon being explored. This raises a key issue about how data, perhaps a transcript with a long sequence of utterances alternating between an interviewer and participant (labelled as I and P in figure 4), can be fragmented during analysis in ways that acknowledge and provide insights into potential change.

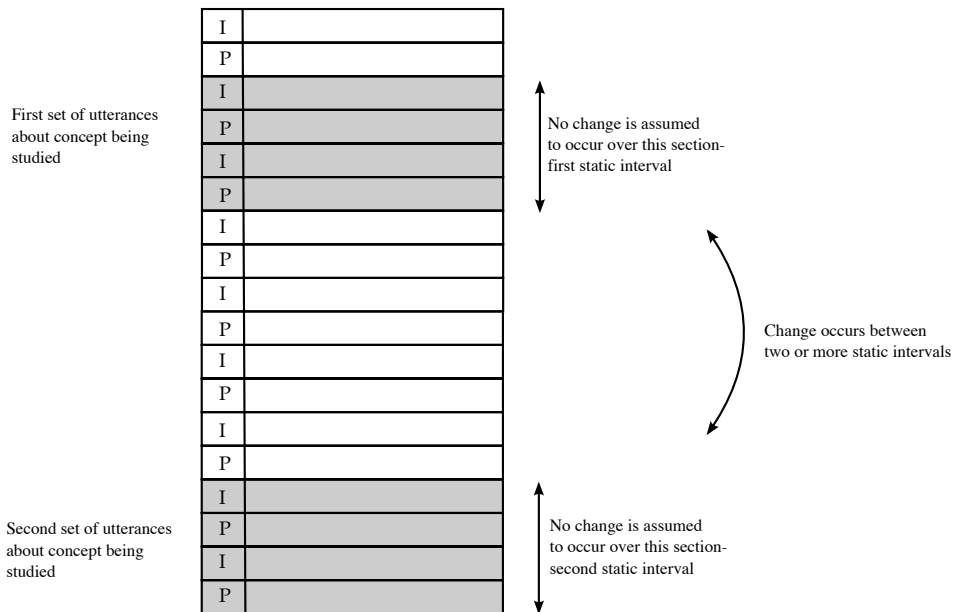


Figure 4: An illustration of the concept of static interval. The image displays a section of transcript. I and P refer to the interviewer’s and participant’s utterances respectively.

In order to report change, researchers will describe a perceived difference between data collected at two or more different times. This approach assumes that change occurs in the interval between these two sections of data but, crucially, the sections of data themselves, the start and end points of the change, are assumed to be essentially static: change occurs *between* the sections rather than *during* the sections. These periods in which no change is assumed to occur are referred to here as static intervals (see figure 4). Siegler (1995, p. 226) draws an analogy between the rapid flow of images in a film and the high density of observations in microgenetic data. A static interval may be likened to an individual frame in a film in which the action appears paused but when played in a sequence, may display evidence of change.

There follows a discussion of two papers, Parnafes and diSessa (2013) and Taber (2008b), to illustrate the concept of the static interval. Parnafes and diSessa (2013) investigated students’ understanding of simple harmonic motion. In the study, described in detail in Parnafes (2010), the students participated in an interview of an hour-and-a-half duration consisting of three sections: a) interaction with physical oscillators (25-40 mins); b) interaction with computer simulations (25-40 mins); and c) concluding discussions (10-15 mins). In their analysis, the authors report that a student reached two different conclusions about the concept of ‘fastness’ in the contexts of an oscillating pendulum and a rod (Parnafes & diSessa, 2013, p. 24). In the case of the pendulum, the student infers that velocity is related to the distance the pendulum travels (that is, when it has a larger amplitude, it is travelling faster); when studying the oscillations of a springy rod, the ‘fastness’ of the motion is linked to the frequency of tapping of the rod on the desk.

Parnafes and diSessa argue:

This is a canonical case of lack of alignment; while Rachel believes she is determining the same kind of information in both situations, 'fast or slow,' she is actually determining two different kinds of information. (Parnafes & diSessa, 2013, p. 24)

By contrast, it is reported that the introduction of friction to a computer simulation '... led to the discovery of an important new relationship' (Parnafes & diSessa, 2013, p. 26). Two interpretations of sequential reports of two different understandings are presented: the first implies two different conceptualisations coexist, the second that a change has occurred. In the first case, the pendulum and the rod are seen as two contexts triggering different interpretations of 'fastness' (as in Mortimer's (1995) conceptual profile model) with an implication that the probe triggers alternative understandings to be brought to mind rather than causing change. In the second case, in the context of the computer simulation, the presentation of one understanding of periodicity followed by another is seen as a 'discovery' rather than a shift between co-existing positions.

The aim of highlighting the difference in conceptualisation is not intended to criticise Parnafes and diSessa's (2013) interpretations of the data, but rather to highlight that a shift in conceptualisation has occurred. In the case of the pendulum and the rod, the two different understandings are treated as though they occupy the same interval in time, that is, they are constructed as existing concurrently. In the case of the 'discovery' of a novel understanding of periodicity, a claim for the occurrence of change is made. The concept of a static interval, defined above as a section of data over which it is assumed change does not take place, may be a useful construct to distinguish these two cases. The two models of 'fastness' are considered to coexist within a single static interval whereas the shift in understanding of periodicity, by implication, implies a transition between two static intervals. A discussion of the difficulty of distinguishing conceptual change from the presence of multiple conceptions is outlined in the next section.

Taber (2008b) presents a different division of time: he interviewed students to examine their application of knowledge of forces and energy to a variety of contexts, introduced sequentially: an apple hanging on a branch, the solar system, and a stretched spring, amongst others. In Taber's (2008b) study, the intention was to explore thinking related to the same underlying scientific ideas in different contexts, and so the study was not intended to be microgenetic. However it is possible to conceptualise the work as a study of potential change: the student is presented with multiple, sequential tasks that, from a formal scientific perspective, access similar conceptual knowledge.

Taber suggested that this kind of interview covering a wide range of topics offered potential to explore the complexity of a learner's cognitive structure (the presence of manifold conceptions, the range of application of a principle from the students perspective, i.e. the extent to which a principle was applied in different contexts). He reported that:

The findings suggest such approaches make it possible in principle to identify situations where a student does not demonstrate ‘target knowledge’ because it is not accessed in a particular context rather than there being a lack in the basic conceptual resources. Alice *did* understand the principle of electrical induction—but did not offer that as a basis for explaining the balloon ‘trick’. In that context, this possibility did not seem to come to mind as a basis for an explanation—she did not make the link. Such ‘failures to connect’ are potentially significant in science teaching and learning. (Taber 2008b, p.1934, italics in original)

The argument was that although the student interviewed, Alice, did not suggest that a balloon charged by friction remained stuck on a wall due to induction, this was a failure of application (i.e. of making a connection, of bringing a principle that could be applied to mind) and not ignorance of the possibility of induction. This was inferred because later in the interview Alice explained van der Waals’ forces in terms of electrical induction. Another example offered in that study concerned the relationship between force and motion. Alice described orbital motion as occurring without net force, as she considered there were two forces (centripetal and centrifugal) which balanced. Alice described this balance of forces at two different points in the interview when asked about the planets orbiting the sun, and later when asked about the moon orbiting the earth. This data could have been interpreted to suggest that Alice did not understand that accelerated force required net force. Yet in the same interview, and indeed between the two points where she talked about orbital motion, Alice discussed the forces acting on a parachutist falling to earth, and described both accelerated motion under net force and terminal velocity where forces balanced. The conclusion drawn was that Alice did understand the relationship between force and acceleration, but did not apply this principle to explaining orbital motion (which she did not consider accelerated motion).

It is known that research interviews have something of the structure of Socratic dialogue, and can act as learning contexts for students. It was possible that Alice’s understandings changed through the process of being asked about her ideas, and that her responses at different points in the interview could be seen as reflecting underlying cognitive structure that was evolving during the interview process. In that case, her responses to consecutive probes might be considered to exist in separate static intervals and change might be inferred between them. Indeed some degree of change is likely (just as we hope it will be during a science lesson). However, the assumption made in the study was that any such change was limited enough to allow the whole interview period to be accessing thinking resourced by cognitive structure that was stable enough to be considered to be essentially fixed over the period of data collection (approximately one hour). Taber’s (2008b) conceptualisation of the study can be understood as implying that the data exist within a single static interval and that change to cognitive structure during the interview was negligible.

Taber (2008b) and Parnafes and diSessa (2013) used similar methods: a single interview including multiple, non-identical, probes of a student’s understanding. However, the authors characterise their data differently, in particular they make implicitly different assumptions about the division of reports of cognition over time. The interview is considered either as a single static interval, describing an unchanging

cognitive structure applied in several contexts (as in Taber, 2008b), or the interview is interpreted as containing multiple static intervals, describing an evolving understanding (as in Parnafes and diSessa's (2013) report of a discovery). Both of these interpretations are justifiable, however it is important that authors clearly define their assumptions related to the intervals over which change does and does not occur.

A researcher's choice of static interval will depend on the phenomenon being investigated. Some kinds of change, labelled development, are expected to happen in all normal subjects due to a process of maturation (for example the acquisition of spoken language). Learning is a different kind of change process, which is both highly contingent on opportunities to engage with specific natural phenomena and to experience cultural mediation (such as teaching) of canonical knowledge, and is also dependent upon how such experiences are understood in terms of existing cognitive structure. The investigation of these different phenomena requires different research approaches. Microgenetic studies of development tend to focus on behaviours or the use of strategies. For example, Adolph and colleagues (2008) examined infants' motor development by recording instances of sitting, crawling, walking and other skills. Such phenomena have relatively clear starting and ending points, though there may still be ambiguity, for example, an infant may engage in a behaviour that combines elements of crawling and walking. In these cases, a relatively short duration static interval may be constructed, as the behaviours have a relatively short duration, and are relatively well defined, as the start and end points of change are relatively distinct. However the situation is more complicated if the investigation focuses on learning scientific concepts: as Parnafes and diSessa (2013, p. 15) state '[l]earning a concept simply cannot happen in a single try, as a new strategy usually appears.' It is more challenging to define the start and end points of conceptual learning than for the process of the acquisition of a strategy, hence defining appropriate static intervals for the study of learning is challenging.

The relationship between phenomenon and static interval may explain the focus in developmental psychology on strategy use (e.g. Philips & Tolmie, 2007; Siegler, 1995) and the historical focus in science education research on individual concepts rather than wider explorations of cognitive structure (Taber, 2009, p. 326). Strategies are generally relatively well defined and can be investigated with short static intervals, similarly it is usually expected that discrete conceptions can be elicited within a static interval by suitable probing. However, the extended length of time that would be required to probe a cognitive structure makes it difficult to maintain the assumption that change is not occurring over the observation period. Research interviews intended to explore conceptual understanding at one point in time always involve a compromise between thoroughness (requiring extended exploration across a range of contexts) and resolution (such that ideally conceptualisation does not change, despite a research interview clearly being a learning opportunity). There is an 'in principle' restriction when the researcher's aim is to detail an aspect of a person's cognitive structure at one moment, as both cognition and research are processes and so inevitably occur over time (Taber, 2013).

c) Sampling rate and the rate of change of learning

In order to construct a sufficiently high resolution representation of change, Siegler and Crowley's (1991, p. 606) second criterion for microgenetic studies is that:

‘The density of observations is high relative to the rate of change of the phenomenon’. This requirement might seem paradoxical as it appears that researchers are required to know the rate of change of a phenomenon before it has been investigated. In most cases however, researchers will be able to present evidence from previous research, pilot studies, or studies of related phenomena to make a reasonable estimate of the rate of change of the phenomenon they are investigating (an example discussion, in the context of conceptual change, is presented below). An appropriate choice of sampling rate is of crucial importance to research focused on change, including a number of significant areas in science education research: the nature of conceptual change (Özdemir & Clark, 2007, p. 356), alterations in conceptual profiles (Mortimer, 1995) and the stability of students’ ideas (Taber, 1995). Microgenetic studies may be seen as a subset of longitudinal studies, in which ‘...two or more measures or observations are made at different times of the same individuals or entities’ (White & Arzi, 2005, p. 138). Wertsch (1985, p. 55) has described microgenetic studies as a form of ‘very short-term longitudinal study’ or more precisely, microgenetic studies can be thought of as longitudinal studies in which the rate of sampling is high compared to the rate of change of the phenomenon of interest.

Most longitudinal studies are seen as developing ‘snapshots’ rather than the ‘near continuous flow of information’ (Siegler, 1995, p. 226) seen in microgenetic methods: moment-by-moment analysis is seen as the ‘gold-standard’ of the microgenetic approach (Parnafes & diSessa, 2013, p. 7). When observing change, an overly low density of observations will result in a loss of detail in the data produced (Kuhn, 1995, p. 133). For example, Adolph and colleagues (2008) demonstrated that altering the choice of sampling rate of data on the development of motor skills, led to an apparent stage-like development not seen with higher rates of sampling. If researchers select an overly high sampling rate, the researchers and participants may undertake unnecessary sessions that produce limited evidence of change. To extend Siegler’s (1995, p. 226) analogy between the microgenetic method and cinematic film, if the shutter speed of a camera is too slow, moving objects will appear blurred in images, if the shutter speed is too high, the pictures may come out dark. The choice of sampling rate is a significant research decision as it will influence the representation of change developed in a piece of research.

The Nyquist-Shannon theorem (Lüke, 1999) allows engineers to determine an appropriate rate at which to sample analogue signals (Shannon, 1949, p. 11) but no similar metric exists in the case of learning. In order to fulfil Siegler and Crowley’s (1991, p. 606) condition, researchers need to make and justify claims for the rate or rates of change of the phenomenon they are observing. The studies in table 2 investigated a range of different phenomena, for example, understanding (Haglund et al., 2016; Van Der Steen et al., 2014) or learning (Magnusson, 1996; Nuthall & Alton-Lee, 1993) and a range of different sampling rates were used. In some cases change was observed over multiple measurements within a session (Opfer & Siegler, 2004; Van Der Steen et al., 2014), in others, change was assessed using probes spaced several weeks apart (Soong, 2008). To focus on a particular example, much learning of interest in science education has been conceptualised as conceptual change. Conceptual change has modelled in a number of different ways, some of which include the possibility of change occurring at multiple rates (see table 3). A researcher interested in investigating conceptual change using the microgenetic approach should

make a case that their choice of sampling rate is sufficiently high to adequately represent the change process they are studying.

Author(s)	Relatively gradual model	Relatively rapid model
Clement, 2008, p. 68	Accretionism	Eurekaism
West and Pines, 1985, p. 5	Conceptual Development	Conceptual Change
Gilbert & Watts, 1983, pp. 88–90	Smooth change	Stepped-change Catastrophic change

Table 3: Models of conceptual change which suggest multiple rates of change occur

Though a growing orthodoxy suggests conceptual change is largely gradual (Nussbaum, 1989, p. 538; Posner et al., 1982, p. 223; Vosniadou, 2008a, p. 12), there is evidence that significant shifts in conceptualisation may occur on shorter timescales, for example, Chi's (1997, p. 230) sudden ontological recategorisations and Clement's (2008, p. 99) model of insight. However, abrupt conceptual change 'does not appear to be the usual road to conceptual change' (Vosniadou, 2008b, p. xvi) and insight is seen as 'the exception rather than the rule' (Fisher & Moody, 2002, p. 67). An insight, a moment of sudden awareness of novel relationships between concepts (Brock, 2015), that appears rapid may be the accumulation of gradual incremental steps in understanding (Nersessian, 1999, p. 14) or the result of a largely tacit incubation period (Smith, 1995, p. 242). It may be that conceptual change, and other cognitive processes, can occur at a variety of rates or at rates which vary over time. Researchers should consider, given a choice of sampling frequency, the types of variation that will be able to be inferred in, and those that will be excluded from, their data.

Three studies in table 2 focused on conceptual change and used relatively high frequencies of observations: Duit, Roth, Komorek & Withers (1998) carried out four observations over two weeks; Wiser and Amin (2001) examined teaching sessions that occurred several mornings a week over five weeks; Opfer and Siegler (2004) analysed changes over multiple probes within a single session. In these cases there appears to be an appropriate fit between the possibly short timescale variation of the phenomenon under investigation, conceptual change, and the frequency of observation. Defining the rate of change of phenomena such as understanding or learning is undoubtedly challenging. However, lacking from many microgenetic studies is a clear statement of the assumed rate or rates of change of the phenomenon being investigated. The relationship between the rate or rates of change of a phenomenon and the sampling frequency is a criterion for microgenetic research and researchers must make explicit their assumptions about these two rates in order to present an argument that their sampling is 'high density' (Siegler & Crowley, 1991, p. 606). For example, in studies which investigate changes in understanding (E.g. Haglund et al., 2016; Van Der Steen et al., 2014), a claim for the use of the microgenetic approach might be supported by an explicit discussion of the rate of change of understanding in comparison to the sampling rate used in the study.

d) Noisiness and stability of data collected with 'high density' sampling

The third issue relating to data collected with the microgenetic method is a difficulty in distinguishing 'signal' from 'noise' (Silver, 2012). In the case of

learning, this might mean discriminating ‘genuine systematic change’ in cognitive structure from transient fluctuations (Lee & Karmiloff-Smith, 2005, p. 257). Taber (2008a, p. 1028) argues researchers should ‘...distinguish between thinking that reflects stable ‘alternative conceptions’ from thinking that constructs a viable but labile response to which the learner has little commitment’ and so be able to discriminate ‘significant progression’ from ‘mental flotsam and jetsam’ (Taber, 1995, p. 5). This is not an easy distinction to make, as learning is increasingly seen as a ‘messy’ process (Taber, 2013) involving the complex interaction of an array of structures (Smith, diSessa, & Roschelle, 1993, p. 148). The distinction between ‘signal’ and ‘noise’ is especially acute in microgenetic research as the high frequency of sampling is reported as uncovering ‘untidy’ change (Flynn et al., 2007, p. 4).

Though stable cognitive elements are often the focus in studies of learning (Petri & Niedderer, 1998, p. 1075; Taber, 1995), transient processes may nonetheless be significant (Siegler, 2006, p. 488). If conceptual change resembles gradual accretion, with few moments of significant advance (Vosniadou, 2008a, p. 12), then it is the accumulation of the myriad ‘false starts’ and ‘relapses to intuitive conceptions’ (Lappi, 2013, p. 2) that constitute learning. For strategy use at least, it is expected that periods of acquisition will be characterised by variability (Adolph et al., 2008, p. 2; Siegler & Chen, 1998, p. 278). As in the case of Penzias and Wilson’s discovery of the cosmic microwave background radiation, the ‘excess noise’ may turn out to be the signal (Partridge, 2007, pp. 46–48). Research is insufficiently advanced to determine whether transient conceptual elements have a significant role in longer-term learning and, therefore, they should not be excluded during analysis. An understanding that appears ad hoc or short-lived within a particular set of data, may, if observations were continued over an extended period, develop into a stable conceptual entity. In the absence of sufficient data researchers should be cautious in discounting the significance of apparently short-lived elements.

An additional difficulty is that a student may possess multiple co-existing understandings that are stable and selectively activated in response to particular contextual cues (Harrison & Treagust, 2000; Mortimer, 1995; Taber, 2000b). When sampling occurs at high densities, it becomes difficult to determine if a report of several different consecutive conceptions is evidence of multiple conceptions reported serially or conceptual change.

Several studies report descriptions of change within a single session (Berland & Crucet, 2016; Opfer & Siegler, 2004; Parnafes & diSessa, 2013; Van Der Steen et al., 2014). Such descriptions of short timescale change are valuable, but it is important that researchers highlight that any claims to stability of change are only valid over the timescale of the probes. Kuhn and Phelps (1982, p. 39) report repeatedly observing students who appeared to ‘solve’ a problem in one session, but when presented with the same problem in the next session, failed to reapply their insight. It seems therefore, that both a high density of observation, to capture short timescale changes in conceptualisations, and an extended period of observation ‘so that shifts in the landscape of cognitive structure may be detected’ (Taber, 2001, p.735) is required to develop a full representation of conceptual change. This approach is seen in Van Der Steen and colleagues (2014) study which reports change within three sessions that occur over a period of six months: both rapid changes in understanding and long-term stabilities in conceptualisation may be examined. This strategy can be conceptualised

as a hybrid approach consisting of a longitudinal series of three periods of microgenetic sampling. Data collection schemes, which allow for variations that occur across a range of timescales may be useful for developing more nuanced models of change.

Tracking change over extended periods of time can be onerous for the researcher and participants; one approach to capturing long-term change is to 'accelerate' or 'miniaturize' (Catán, 1986, p. 260) learning into shorter than normal timescales. Both Catán (1986, p. 260) and Kuhn (1995, p. 137) argue that microgenetic studies should include an element of 'acceleration'. For example, Kuhn, Schuable and Garcia-Mila (1992, p. 286) claim their exposure of ten year-old students to twice weekly 30-45 minute scientific reasoning problems over nine weeks, is an 'acceleration' of the normal acquisition of these skills (for a different approach to acceleration, see Shayer and Adey, 1993). Though this approach has been used in some microgenetic studies, it is not clear if 'accelerated' learning resembles learning on longer timescales (Miller & Coyle, 1999, pp. 212–213). Siegler & Crowley's (1991, p. 606) criteria do not require 'acceleration,' nonetheless any form of intervention is potentially an alteration of the 'natural' learning process. Though intentional 'acceleration' may sometimes be a useful tool for researchers in science education, it is not a necessary feature of microgenetic research.

Studies which combine microgenetic and longitudinal approaches (E.g. Van Der Steen et al., 2014) allow for the description of both short timescale variability and the discussion of longer timescale stabilities. An extended period of sampling, across a range of contexts, may assist in distinguishing the presence of multiple conceptions from conceptual change. An alteration in the frequency with which a particular concept is applied in a given context, when observed over a sufficient number of trials, may be taken as evidence of conceptual change. It has been reported that substantive shifts in conceptualisation may take weeks, months or years to occur (Shuell, 1990, p. 531) and so it is difficult to define a duration of observation that would be necessary to support a definitive claim regarding the occurrence of conceptual change. The third of Siegler and Crowley's (1991, p. 606) criteria for microgenetic studies requires that 'the entire period' of change is observed until the phenomenon 'reaches a relatively stable state.' It may be possible to make claims for stability when studying the acquisition of strategies, which may reach a stable state within days or weeks, but similar contentions are much more challenging to achieve with a complex and extended phenomenon such as conceptual learning. The notion of moving towards theoretical saturation (Glaser & Strauss, 1967, p. 61) may be more appropriate than a defined interval of observation for phenomena such as learning or conceptual change. The ideal form of investigation would include microgenetic sampling over an extended period of time to provide evidence of the diversity and stability of conceptual entities.

e) The use of sequences of repeated measures

In defining the conditions of microgenetic studies, Chinn (2006, p. 441) states a participant '...typically encounters similar tasks and measures repeatedly' and implies studies that do not use counterbalanced tasks should not be considered microgenetic (Chinn, 2006, p. 451). Counterbalancing is a technique intended to reduce the effects caused by repeatedly encountering the same task, by changing the

order in which the probes are presented to individual participants (Gaito, 1961, p. 46). This section considers the methodological choices researchers face when choosing the types of probes to use in microgenetic studies. Sequences of measures may be thought of as existing on a continuum of similarity (see figure 5), including sequences of identical probes and sequences of probes which are apparently different but investigate the same conceptual area (the extreme case of sequences of non-identical probes that investigate unrelated conceptual areas is not considered here as it seems an unlikely choice for a researcher). However, defining similarity is not straightforward and perceptions of relatedness may depend on expertise. Chi, Feltovich and Glaser (1981, p. 125) highlighted that experts and novices conceptualised the similarity of problems in different ways and observed that expert physicists' judgements of similarity may be difficult for novices to understand. For example, to an expert, a problem concerning the motion of the mass on a spring and a problem about a block sliding down a slope, may be perceived as similar, because their solution methods share the same 'deep structure' of energy conservation (Chi et al., 1981, pp. 125–127). Measures that have differing surface features may be argued, at least to an expert, to be similar if they share some kind of underlying structure (see figure 5). Of the studies listed in table 2, 15 used sequences of identical repeated measures, 13 sequences of non-identical measures and nine a mix of the two approaches.

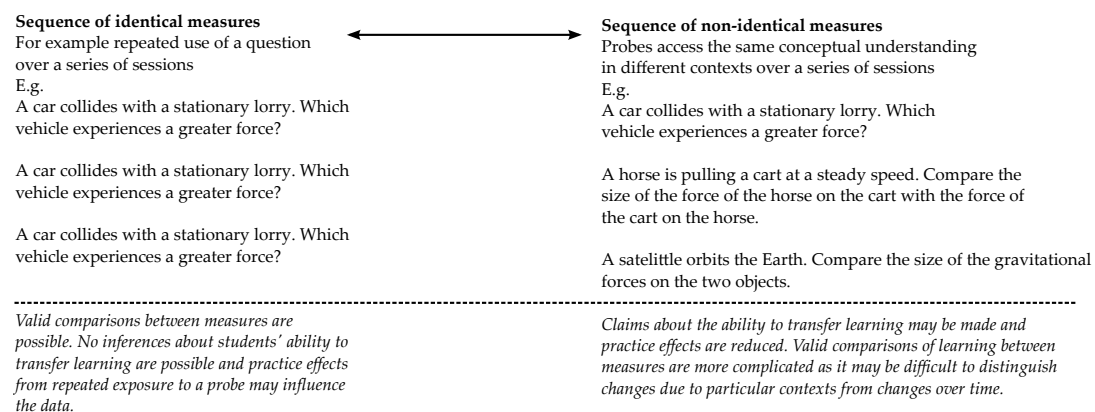


Figure 5: The continuum of similarity of sequences measures

The relatively high frequency application of probes characteristic of the microgenetic approach may produce artefacts in the data (Lee & Karmiloff-Smith, 2005, p. 250), for example, the frequent exposure to novel experiences may differ from typical learning experiences and so influence learning progressions (Miller & Coyle, 1999, p. 212). The nature of assessment affects may vary for sequences of measures found at different places on the continuum above (see figure 5). For example, Siegler and Crowley's (1991) use of a sequence of relatively similar measures in an investigation of addition strategies was criticised by Pressley (1992, p. 1240) who argued that microgenetic studies lead to a confounding of '...time of assessment with experimental effects associated with assessment'. At the other end of the continuum, where identical measures are not used (such as in Parnafes and diSessa's (2013) paper), it is impossible to separate effects due to context from changes over time. If the intention of a study is to produce a narrative of a particular learner's developing understanding across different contexts, this effect is unavoidable and non-identical measures are a necessary choice. The data then, as in

the case-study approach (Flyvbjerg, 2006, p. 221), are inherently bound up with the nature and order of probes and the interaction with the researcher. The onus on a researcher using non-identical measures is not to counterbalance the tasks, but to provide a detailed description of each probe and the participants' responses. The use of non-identical but isomorphic probes is similar to the kind of assessment that a student might routinely encounter in the classroom, and so it might be argued to have stronger analytical generalisability (Kvale, 1996; Taber, 2000a) than a small-scale study using identical measures.

The choice of sequences of measures will be partly informed by the phenomenon being studied. Studies of strategy acquisition might be expected to use highly similar repeated measures, for example, Siegler and Jenkins (1989) repeatedly set participants a task involving the addition of two integers. Contrastingly, studies of the transfer of strategies to novel situations, by definition, require the use of non-identical measures (E.g. Chen & Klahr, 1999). Experts, at least in some domains, appear to have the ability to transfer learning to novel situations (Haskell, 2001, p. 4; Mayer, 2002). It might be argued that transfer, the ability to apply learning to novel domains (Singley & Anderson, 1989), is a key feature of successful learning as '...learning ...almost always generalizes beyond the item on which a new approach was generated' (Siegler, 2006, p. 494). Therefore, it has been argued, a microgenetic study of learning should involve a variety of related, but non-identical tasks that share similar underlying structures (Kuhn, 1995, p. 136).

Finally, though it might be assumed that greater similarity between measures reduces variability in responses over short timescales; this is not necessarily the case. Lasry and colleagues (2011) tested one hundred students using the force concept inventory (Hestenes, Wells, & Swackhamer, 1992) and retested them within a week, with no additional instruction. Though they found a high-degree of reliability between individual's *total* scores, on examining individual questions, they discovered roughly a third of responses had changed from the first to second administration. Kuhn and Phelps (1982, p. 40) reported significant variability in participants' strategy use even when encountering identical problems and argue this variability should be treated as 'an important subject of substantive investigation, rather than a methodological source of error.'

In microgenetic studies in science education, the choice of different types of sequences of probes should be related to the aims of the research. For example, where researchers sought to understand changes in students' behaviours in a particular context, for example in understanding a winch (Izsak, 2000) or building towers to within stand earthquakes (Azmitia & Crowley, 2001), identical measures are appropriate. In other cases, the researcher may seek to understand a student's ability to apply ideas across a range of different contexts and so use a range of probes that share 'deep structure.' For example, Duit, Roth, Komorek, and Withers (1998) observed students interactions with a range of scenarios related to chaotic systems and diSessa (2014) examined students' thinking about causality over a number of different scenarios related to temperature equilibration. In these cases, the use of non-identical probes allowed the researchers to understand how students' transferred, or failed to transfer, learning to novel contexts.

The continuum of similarity of measure presents a choice for the researcher; identical measures may be useful for investigating well-defined, sharply focused questions (Cook, 2002, p. 179) concerning phenomena that require a short static interval. However, more complex forms of learning are fundamentally applicable across a range of contexts, and controlling for context would be mistaken (Maxwell, 2004, p. 6). An approach, which uses probes with overlapping 'deep structures,' will provide a broader understanding of change, but an onus is placed on the researcher to justify the claim that the probes access conceptual understanding related to the same scientific concept.

Recommendations arising from a consideration of microgenetic studies in science education

The analysis of papers using the microgenetic method in science education has led to the identification of a number of considerations relevant to researchers who examine change in phenomena using the approach. In many cases, our criticisms are not related to the particular methodological choices made by researchers, but rather that research reports do not always make clear the assumptions that guided their decisions. Therefore, a number of these recommendations suggest that authors should aim to be explicit about certain kinds of assumption that underlie the manner in which they have constructed descriptions of change.

- **Assumptions regarding the rate or rates of change of the phenomenon being studied and the sampling frequency might be explicitly discussed.**

Ideally, researchers should present evidence of meeting Siegler and Crowley's (1991, p. 606) core criterion by providing an argument for an assumed rate or rates of change of the phenomenon being studied and show that the pattern of observations made can be expected to reveal such changes.

- **A discussion of the static interval and its relation to the phenomena being studied may be useful in supporting claims of change.**

Researchers can strengthen claims about change by clarifying their assumptions concerning both the sections of data that are assumed to represent unchanging processing, a static interval, and how claims to change are developed by the comparison of static intervals. To clarify their case, researchers might make explicit the nature of the object of study (a particular strategy, a single concept, a conceptual area) and make a case that the chosen static interval is appropriate for that phenomenon. The static interval may be defined as several sessions, a whole session, or a part of a session, however the researcher must convince the reader that the phenomenon being studied would not be expected to vary significantly over that period.

- **Analysis may be quantitative or qualitative but a defining feature of microgenetic research is a sense of moment-by-moment change.**

The microgenetic approach may be used effectively within research designs collecting and analysing either, or both, qualitative and quantitative data providing researchers clarify the assumptions that support their work and make claims that are consistent with those assumptions. Fine-grained representations of change in both the ideographic and nomothetic traditions will contribute to developing a richer model of change (Gilbert & Watts, 1983; Taber, 2009, p. 391). However, the character of

microgenetic studies is to represent moment-by-moment change and researchers should strive to develop high-density representations of change over time rather than descriptions of change between a number of widely spaced data collection episodes which are characteristic of longitudinal research.

- **Researchers should consider their choice of length of observation period to suit the phenomenon they are studying.**

The extent of observations made is significant for distinguishing conceptual change from the presence of multiple conceptions and for assessing the stability of constructions. However, given that some processes of interest in science education, for example conceptual learning, occur over an extended period of time, Siegler and Crowley's (1991, p. 606) requirement for observation to cover the entire period of change may not be practicable. Useful insights may be derived from more limited observations, as long as the observation period can be expected to reveal a sufficient extent of change to be theoretically important. For example, research suggests that during a two-year chemistry course a student may show significant shifts in the pattern of explanations offered about the nature of chemical bonding without completely abandoning their initial alternative conceptions (Taber, 2001). With some phenomena it may not be practicable to present evidence of a phenomenon having reached 'a relatively stable state' (Siegler & Crowley, 1991, p. 606), however a researcher should present an appropriately extended section of data to enable judgements to be made concerning the typicality and stability of change.

- **Researchers should consider the appropriateness of the chosen sequences of assessments for investigating the phenomenon of interest.**

Sequences of probes with higher or lower degrees of similarity are acceptable though each type influences the data in particular ways. Where probes have lower surface similarity, researchers need to offer an explicit case for how the range of probes used can be considered to offer alternative prompts for accessing the same underlying skills or cognitive resources. Chinn's (2006, pp. 443–444) criticisms of small-scale microgenetic studies apply inappropriate expectations of generalisability and validity to small-N studies, so instead we argue for the importance of researchers justifying the appropriateness of different methods for the study of various phenomena. For example, counterbalanced tasks and identical measures may be used to investigate phenomena that can be studied in relatively short static intervals, for example, problem-solving strategies. However, phenomena such as conceptual learning are more challenging to study and are likely to require longer static intervals and more complex and non-identical sequences of measures to probe transfer across contexts.

- **The microgenetic method may be used in small-scale studies providing caveats regarding generalisability, validity and reliability are provided.**

Similar kinds of justifications to those commonly found in small-scale research projects are required in small-scale microgenetic projects. It would be appropriate for a small-scale study to make use of analytical generalisability rather than statistical generalisability (Taber, 2000a). A qualitative conception of validity (Creswell & Miller, 2000) will be appropriate for such work and may be supported by: a) the use of multiple methods (such as concept-maps or concept inventories) within the microgenetic framework (Shenton, 2004, p. 65); b) a clear statement of theoretical position and assumptions (Creswell & Miller, 2000, p. 127); and c) rich description of data (Onwuegbuzie & Leech, 2007, p. 244). Reliability, in case study research, is

conceptualised as reducing ‘errors and biases’ and may be supported by reporting details of the methods and data (Yin, 2009, p. 45). The argument in this paper, which makes a case for the acceptability and fruitfulness of small-scale, qualitative microgenetic studies in science education, is not intended to criticise the value of large-scale, quantitative microgenetic studies: the approaches should be seen as complementary. As Taber (2009, p. 351) has suggested, educational research might be envisaged as an iterative process with a ‘methodological pendulum’ oscillating between small-scale, in-depth studies and more generalisable approaches.

It is hoped that these guidelines will be useful to researchers using the microgenetic method. The variability of data produced by the microgenetic method, should not be seen as a hindrance, but rather as a reflection of the inherently changeable nature of cognition (Smith & Thelen, 2003, p. 343). In a model in which learning is seen as ‘messy’ (Taber, 2013), researchers might find it useful to assume that: ‘the signal is the noise’ (Landauer, 1998, p. 658). A potentially fruitful direction for science education research is to engage with the untidy, fine-grained detail of conceptual-structure change. Despite the costs in terms of time and effort (Lee & Karmiloff-Smith, 2005, p. 259), the microgenetic method is an appropriate tool for observing change but researchers who use the method need to be explicit about the assumptions they make in their work. It is hoped this paper will contribute to a conversation regarding the use of the microgenetic method in science education.

References

- Adolph, K. E., Robinson, S. R., Young, J. W., & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review*, *115*(3), 527–543.
- Alberini, C. (2011). The role of reconsolidation and the dynamic process of long-term memory formation and storage. *Frontiers in Behavioral Neuroscience*, *5*(12), 1–10.
- Azmitia, M., & Crowley, C. (2001). The rhythms of scientific thinking: A study of collaboration in an earthquake microworld. In K. Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings* (pp. 51–81). Mahwah, NJ: Erlbaum.
- Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172.
- Berland, L. K., & Cruet, K. (2016). Epistemological Trade-Offs: Accounting for Context When Evaluating Epistemological Sophistication of Student Engagement in Scientific Practices. *Science Education*, *100*(1), 5–29.
- Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765–793.
- Blackmore, S. (2002). There Is No Stream of Consciousness. *Journal of Consciousness Studies*, *9*(5-6), 17–28.
- Bliss, J. (1995). Piaget and after: the case of learning science. *Studies in Science Education*, *25*(1), 139–172.
- Blown, E. J., & Bryce, T. G. K. (2006). Knowledge restructuring in the development of children’s cosmologies. *International Journal of Science Education*, *28*(12), 1411–1462.
- Brock, R. (2015). Intuition and insight: two concepts that illuminate the tacit in science education. *Studies in Science Education*, *51*(2), 127–167.

- Brown, A. L. (2004). Transforming Schools into Communities of Thinking and Learning about Serious Matters. In E. Scanlon, P. Murphy, J. Thomas, & E. Whitelegg (Eds.), *Reconsidering Science Learning* (pp. 74–90). London: Routledge Falmer.
- Caballero, M. D., Greco, E. F., Murray, E. R., Bujak, K. R., Jackson Marr, M., Catrambone, R., ... Schatz, M. F. (2012). Comparing large lecture mechanics curricula using the Force Concept Inventory: A five thousand student study. *American Journal of Physics*, *80*(7), 638–644.
- Catán, L. (1986). The dynamic display of process: Historical development and contemporary uses of the microgenetic method. *Human Development*, *29*(5), 252–263.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098–1120.
- Cheng, M. H. M., & Wan, Z. H. (2015). Unpacking the paradox of Chinese science learners: insights from research into Asian Chinese school students' attitudes towards learning science, science learning strategies, and scientific epistemological views. *Studies in Science Education*, *52*(1), 29–62.
- Chi, M. T. H. (1997). Creativity: Shifting across ontological categories flexibly. In B. T. Ward & S. M. Smith (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 209–234). Washington, DC: American Psychological Association.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.
- Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 3–27). Dordrecht: Springer.
- Chinn, C. A. (1997). *A microgenetic study of learning about the molecular theory of matter and chemical reactions*. Unpublished Ph.D. Thesis, University of Illinois at Urbana-Champaign.
- Chinn, C. A. (2006). The microgenetic method: Current work and extensions to classroom research. In J. Green, G. Camilli, & P. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 439–456). Mahwah, NJ: Erlbaum.
- Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Experimental Education*, *69*(1), 77–97.
- Clement, J. (2008). *Creative model construction in scientists and students*. Dordrecht: Springer.
- Cook, T. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Education Evaluation and Policy Analysis*, *24*(3), 175–199.
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, *39*(3), 124–130.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and Company.
- diSessa, A. (2000). *Changing minds: Computers, learning, and literacy*. Cambridge, MA: MIT Press.
- diSessa, A. (2014). The Construction of Causal Schemes: Learning Mechanisms at the Knowledge Level. *Cognitive Science*, *38*(5), 795–850.
- Driver, R., Leach, J., Scott, P., & Wood-Robinson, C. (1994). Young People's understanding of science concepts: implications of cross-age studies for

- curriculum planning. *Studies in Science Education*, 24(1), 75–100.
- Duit, R., Roth, W., Komorek, M., & Withers, J. (1998). Conceptual change cum discourse analysis to understand cognition in a unit on chaotic systems: towards an integrative perspective on learning in science. *International Journal of Science Education*, 20(9), 1059–1073.
- Duit, R., & Treagust, D. F. (2003). Conceptual change: a powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671–688.
- Eichler, M. L., Del Pino, J. C., & Fagundes, L. da C. (2004). Development Of Cognitive Conducts During A Computer Simulated Environmental Analysis. *Chemistry Education Research and Practice*, 5(2), 157–174.
- Feldon, D. F., & Gilmore, J. (2006). Patterns in children's online behavior, scientific problem-solving: A large-N microgenetic study. In G. Clarebout & J. Elen (Eds.), *Avoiding simplicity, confronting complexity: Advances in studying and designing (computer-based) powerful learning environments* (pp. 117–125). Rotterdam: Sense Publishers.
- Fisher, K., & Moody, D. (2002). Student misconceptions in biology. In J. H. Wandersee & D. Moody (Eds.), *Mapping biology knowledge* (pp. 55–76). New York, NY: Kluwer.
- Flynn, E., Pine, K., & Lewis, C. (2006). The Microgenetic Method Time for change? *Psychologist*, 19(3), 152–155.
- Flynn, E., Pine, K., & Lewis, C. (2007). Using the microgenetic method to investigate cognitive development: an introduction. *Infant and Child Development*, 16(1), 1–6.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245.
- Gaito, J. (1961). Repeated measurements designs and counterbalancing. *Psychological Bulletin*, 58(1), 46–54.
- Garcia-Mila, M., & Andersen, C. (2007). Developmental change in notetaking during scientific inquiry. *International Journal of Science Education*, 29(8), 1035–1058.
- Garcia-Mila, M., Andersen, C., & Rojo, N. E. (2011). Elementary Students' Laboratory Record Keeping During Scientific Inquiry. *International Journal of Science Education*, 33(7), 915–942.
- Gelman, R., Romo, L., & Francis, W. S. (2002). Notebooks as windows on learning: The case of a science-into-ESL program. In N. Garnott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 269–293). Cambridge: Cambridge University Press.
- Gilbert, J. K., & Watts, D. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. *Studies in Science Education*, 10(1), 61–98.
- Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Reserach*. New Brunswick, NJ: Aldine Transaction.
- Goldin-Meadow, S., & Wagner Alibali, M. (2002). Looking at the hands through time: A microgenetic perspective on learning and instruction. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 80–105). Cambridge: Cambridge University Press.
- Ha, J., Lee, K., Lim, S., & Yang, I. (2016). A Microgenetic Study on Changes of Elementary Students' Mental Models Regarding the Particles Constituting Ice, Water, and Water Vapor. *Advanced Science and Technology Letters*, 122, 181–185.

- Haglund, J., Jeppsson, F., & Schönborn, K. J. (2016). Taking on the Heat—a Narrative Account of How Infrared Cameras Invite Instant Inquiry. *Research in Science Education*, 46(5), 685–713.
- Harrison, A. G., & Treagust, D. F. (2000). Learning about atoms, molecules, and chemical bonds: A case study of multiple-model use in grade 11 chemistry. *Science Education*, 84(3), 352–381.
- Haskell, R. E. (2001). *Transfer of Learning: Cognition, Instruction, and Reasoning*. San Diego, CA: Academic Press.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141–158.
- Hogan, K. (1999). Relating students' personal frameworks for science learning to their cognition in collaborative contexts. *Science Education*, 83(1), 1–32.
- Izsak, A. (2000). Inscribing the Winch: Mechanisms by Which Students Develop Knowledge Structures for Representing the Physical World With Algebra. *Journal of the Learning Sciences*, 9(1), 31–74.
- Johnson, K. E., & Mervis, C. B. (1994). Microgenetic analysis of first steps in children's acquisition of expertise on shorebirds. *Developmental Psychology*, 30(3), 418–435.
- Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.
- Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science*, 6(3), 133–139.
- Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, 94(5), 810–824.
- Kuhn, D., & O'Loughlin, M. (1988). The Coordination of Theory and Evidence. In D.Kuhn, E. Amsel, M. O'Loughlin, & M. (Eds.), *The development of scientific thinking skills* (pp. 104–205). New York, NY: Academic Press.
- Kuhn, D., & Phelps, E. (1982). The development of problem solving strategies. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 17, pp. 1–44). New York, NY: Academic Press.
- Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction*, 9(4), 285–327.
- Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage.
- Lachman, S. J. (1997). Learning is a Process: Toward an Improved Definition of Learning. *The Journal of Psychology*, 131(5), 477–480.
- Landauer, R. (1998). Condensed-matter physics: The noise is the signal. *Nature*, 392(6677), 658–659.
- Lappi, O. (2013). Qualitative Quantitative and Experimental Concept Possession, Criteria for Identifying Conceptual Change in Science Education. *Science & Education*, 22(6), 1347–1359.
- Lasry, N., Rosenfield, S., Dedic, H., Dahan, A., & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory. *American Journal of Physics*, 79(9), 909–912.
- Lee, K., & Karmiloff-Smith, A. (2005). Macro- and microdevelopmental research: Assumptions, research strategies, constraints, and utilities. In N. Garnott & J. Parziale (Eds.), *Microdevelopment Transition Processes in Development and Learning* (pp. 243–266). Cambridge: Cambridge University Press.
- Lüke, H. D. (1999). The origins of the sampling theorem. *IEEE Communications Magazine*, 37(4), 106–108.

- Magnusson, S. J. (1996). Complexities of learning with computer-based tools: A case of inquiry about sound and music in elementary school. *Journal of Science Education and Technology*, 5(4), 297–309.
- Marton, F. (1983). Beyond Individual Differences. *Educational Psychology*, 3(3-4), 289–303.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33(2), 3–11.
- Mayer, R. (2002). Rote versus meaningful learning. *Theory into Practice*, 41(4), 226–232.
- Merriam, S. (1995). What Can You Tell From An N of 1?: Issues of Validity and Reliability in Qualitative Research. *PAACE Journal of Lifelong Learning*, 4, 50–60.
- Miller, P. H., & Coyle, T. R. (1999). Developmental change: Lessons from microgenesis. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 209–239). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mortimer, E. (1995). Conceptual change or conceptual profile change? *Science & Education*, 4(3), 267–285.
- Mowrer, R. R., & Klein, S. B. (2001). The Transitive Nature of Contemporary Learning Theory. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of Contemporary Learning Theories* (pp. 1–21). New York, NY: Lawrence Erlbaum Associates, Inc.
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Reply-Reconsolidation: The labile nature of consolidation theory. *Nature Reviews Neuroscience*, 1(3), 216–219.
- Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery* (pp. 5–22). New York, NY: Plenum.
- Niedderer, H., Budde, M., Givry, D., Psillos, D., & Tiberghien, A. (2007). Learning Process Studies. In R. Pintó & D. Couso (Eds.), *Contributions from Science Education Research* (pp. 159–171). Dordrecht: Springer.
- Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28(1), 117–153.
- Nussbaum, J. (1989). Classroom conceptual change: philosophical perspectives. *International Journal of Science Education*, 11(5), 530–540.
- Nuthall, G. (1999). The way students learn: Acquiring knowledge from an integrated science and social studies unit. *The Elementary School Journal*, 99(4), 303–341.
- Nuthall, G., & Alton-Lee, A. (1993). Predicting learning from student experience of teaching: A theory of student knowledge construction in classrooms. *American Educational Research Journal*, 30(4), 799–840.
- O'Brien, G., & Opie, J. (1998). The disunity of consciousness. *Australasian Journal of Philosophy*, 76(3), 378–395.
- Onwuegbuzie, A. J., & Leech, N. L. (2007). Validity and qualitative research: an oxymoron? *Quality & Quantity*, 41(2), 233–249.
- Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' living things concept: A microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology*, 49(4), 301–322.
- Özdemir, G., & Clark, D. B. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science & Technology Education*, 3(4), 336–351.
- Parnafes, O. (2007). What does fast mean? Understanding the physical world through

- representations. *The Journal of the Learning Sciences*, 16(3), 415–450.
- Parnafes, O. (2010). When simple harmonic oscillation is not that simple: The complexity of students' sense making. *The Journal of Science Education and Technology*, 19(6), 565–579.
- Parnafes, O., & diSessa, A. A. (2013). Microgenetic learning analysis: A methodology for studying knowledge in transition. *Human Development*, 56(1), 5–37.
- Partridge, R. B. (2007). *3K: The cosmic microwave background radiation*. Cambridge: Cambridge University Press.
- Pata, K., & Sarapuu, T. (2006). A Comparison of Reasoning Processes in a Collaborative Modelling Environment: Learning about genetics problems using virtual chat. *International Journal of Science Education*, 28(11), 1347–1368.
- Pearsall, N. R., Skipper, J., & Mintzes, J. (1997). Knowledge Restructuring In The Life Sciences: A Longitudinal Study Of Conceptual Change In Biology. *Science Education*, 81(2), 193–215.
- Petri, J., & Niedderer, H. (1998). A learning pathway in high-school level quantum atomic physics. *International Journal of Science Education*, 20(9), 1075–1088.
- Philips, S., & Tolmie, A. (2007). Children's performance on and understanding of the Balance Scale problem: The effects of parental support. *Infant and Child Development*, 16(1), 95–117.
- Piaget, J. (1970). *The Principle of Genetic epistemology*. (W. Mays, Trans.). London: Routledge & Kegan Paul.
- Posner, G., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.
- Pressley, M. (1992). How not to study strategy discovery. *American Psychologist*, 47(10), 1240–1241.
- Pressley, M., Forrest-Pressley, D. L., Elliott-Faust, D., & Miller, G. (1985). Children's Use of Cognitive Strategies, How to Teach Strategies, and What to Do If They Can't Be Taught. In M. Pressley & C. J. Brainerd (Eds.), *Cognitive Learning and Memory in Children* (pp. 1–47). New York, NY: Springer-Verlag.
- Robinson, B. F., & Mervis, C. B. (1998). Disentangling early language development: Modeling lexical and grammatical acquisition using and extension of case-study methodology. *Developmental Psychology*, 34(2), 363–375.
- Rosenthal, V. (2004). Microgenesis, immediate experience and visual processes in reading. In A. Carsetti (Ed.), *Seeing, thinking and knowing: Meaning and self-organisation in visual cognition and thought* (pp. 221–243). Dordrecht: Kluwer Academic.
- Roth, W.-M., & Welzel, M. (2001). From activity to gestures and scientific language. *Journal of Research in Science Teaching*, 38(1), 103–136.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1: Foundations*. Cambridge, MA: MIT Press.
- Rusanen, A. M., & Lappi, O. (2013). What, when and how do the models of conceptual change explain. In *Proceedings of the 35th annual conference of the cognitive science society* (pp. 3331–3336).
- Sander, F. (1930). Structures, totality of experience, and gestalt. In C. Murchinson (Ed.), *Psychologies of 1930: The International University Series in Psychology* (pp. 188–204). Worcester, MA: Clark University Press.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich

- contexts. *Developmental Psychology*, 32(1), 102–119.
- Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1), 10–21.
- Shayer, M., & Adey, P. S. (1993). Accelerating the development of formal thinking in middle and high school students IV: Three years after a two-year intervention. *Journal of Research in Science Teaching*, 30(4), 351–366.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63–75.
- Shuell, T. J. (1990). Phases of Meaningful Learning. *Review of Educational Research*, 60(4), 531–547.
- Siegler, R. S. (1995). How does change occur: a microgenetic study of number conservation. *Cognitive Psychology*, 28(3), 225–73.
- Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology. Volume 2. Cognition, perception, and language*. (pp. 464–510). Hoboken, NJ: Wiley.
- Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology*, 36(3), 273–310.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method: a direct means for studying cognitive development. *American Psychologist*, 46(6), 606–620.
- Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- Silver, N. (2012). *The Signal and the Noise: Why So Many Predictions Fail-but Some Don't*. New York, NY: Penguin.
- Singley, M., & Anderson, J. (1989). The Study of Transfer. In *The Transfer of Cognitive Skill* (pp. 1–41). Cambridge MA: Harvard University Press.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115–163.
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, 7(8), 343–348.
- Smith, S. M. (1995). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 229–251). Cambridge, MA: MIT Press.
- Soong, B. (2008). Learning through computers: Uncovering students' thought processes while solving physics problems. *Australasian Journal of Educational Technology*, 24(5), 592–610.
- Srivastava, A., & Ramadas, J. (2013). Analogy and Gesture for Mental Visualization of DNA Structure. In D. F. Treagust & C. Y. Tsui (Eds.), *Multiple Representations in Biological Education* (pp. 311–329). Dordrecht: Springer.
- Taber, K. S. (1995). Development of student understanding: A case study of stability and lability in cognitive structure. *Research in Science & Technological Education*, 13(1), 89–99.
- Taber, K. S. (2000a). Case studies and generalizability: grounded theory and research in science education. *International Journal of Science Education*, 22(5), 469–487.
- Taber, K. S. (2000b). Multiple frameworks?: Evidence of manifold conceptions in individual cognitive structure. *International Journal of Science Education*, 22(4), 399–417.
- Taber, K. S. (2001). Shifting sands: a case study of conceptual development as competition between alternative conceptions. *International Journal of Science*

- Education*, 23(7), 731–753.
- Taber, K. S. (2008a). Conceptual resources for learning science: Issues of transience and grain-size in cognition and cognitive structure. *International Journal of Science Education*, 30(8), 1027–1053.
- Taber, K. S. (2008b). Exploring conceptual integration in student thinking: Evidence from a case study. *International Journal of Science Education*, 30(14), 1915–1943.
- Taber, K. S. (2009). *Progressing science education: constructing the scientific research programme into the contingent nature of learning science*. Dordrecht: Springer.
- Taber, K. S. (2013). *Modeling Learners and Learning in Science Education*. Dordrecht: Springer.
- Taber, K. S. (2014). The significance of implicit knowledge for learning and teaching chemistry. *Chemistry Education Research and Practice*, 15(4), 447–461.
- Thornton, R. K. (1997). Conceptual dynamics: Following changing student views of force and motion. In *American Institute of Physics Conference Proceedings* (Vol. 399, pp. 241–266).
- Treagust, D. F., & Duit, R. (2009). Multiple Perspectives of Conceptual Change in Science and the Challenges Ahead. *Journal of Science and Mathematics Education in Southeast Asia*, 32(2), 89–104.
- Van Der Steen, S., Steenbeek, H., Van Dijk, M., & Van Geert, P. (2014). A process approach to children's understanding of scientific concepts: A longitudinal case study. *Learning and Individual Differences*, 30, 84–91.
- Veal, W. R. (2004). Beliefs and knowledge in chemistry teacher development. *International Journal of Science Education*, 26(3), 329–351.
- Vosniadou, S. (2008a). Conceptual Change in Learning and Instruction: The Framework Theory Approach. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (1st ed., pp. 11–30). New York, NY: Routledge.
- Vosniadou, S. (2008b). Conceptual change research: an introduction. In S. Vosniadou (Ed.), *International Handbook of research on conceptual change* (1st ed., pp. xiii–xxviii). New York, NY: Routledge.
- Vygotsky, L. (1978). *Mind in Society: the development of higher mental processes* Cambridge, MA: (M. Cole, Trans.). Cambridge, MA: Harvard University Press.
- Werner, H. (1956). Microgenesis and aphasia. *The Journal of Abnormal and Social Psychology*, 52(3), 347–353.
- Wertsch, J. V. (1991). The Problem of Meaning in a Sociocultural Approach to Mind. In A. McKeough & J. L. Lupart (Eds.), *Toward the Practice of Theory-based Instruction* (pp. 31–49). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Wertsch, J. (1985). *Vygotsky and the social formation of mind*. Cambridge MA: Harvard University Press.
- West, L. H. T., & Pines, A. L. (1985). Introduction. In L. H. T. West & A. L. Pines (Eds.), *Cognitive Structure and Conceptual Change* (pp. 1–7). London: Academic Press.
- White, R. T. (1985). Interview protocols and dimensions of cognitive structure. In L. H. T. West & A. L. Pines (Eds.), *Cognitive Structure and Conceptual Change* (pp. 51–59). London: Academic Press.
- White, R. T., & Arzi, H. J. (2005). Longitudinal studies: designs, validity, practicality, and value. *Research in Science Education*, 35(1), 137–149.
- Wiser, M., & Amin, T. (2001). Is heat hot? Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and*

- Instruction*, 11(4), 331–355.
- Yin, R. K. (1981). The case study crisis: Some answers. *Administrative Science Quarterly*, 26(1), 58–65.
- Yin, R. K. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications Incorporated.
- Zohar, A. (1995). Reasoning about Interactions between Variables. *Journal of Research in Science Teaching*, 32(10), 1039–1063.