

around the meanings of key terms during moderation visits.

Assessors sometimes expressed difficulty in separating some of the descriptive qualities within the criteria because, from their perspective, the terminology failed to adequately illustrate differences between the qualities of different performances. This implies that the language used either did not conform to discrete categories or had some overlapping qualities (e.g. 'clear'/'accurate'/'appropriate'/'detailed' or 'basic'/'sound'/'high'), that made it difficult for assessors to fit some performance characteristics to the criteria. Although caution needs to be expressed about making assessment criteria more lengthy (Wiliam, 1998; Wolf, 1995), resolving this issue might involve clarifying the values implicit in the descriptor terminology, perhaps through exemplification, and connecting these meanings through effective communication procedures with assessors' expectations about performance quality. This implies a need to engage assessors in discussions about those aspects of language that they feel hinder their ability to discriminate between performances and to use this as an opportunity to arrive at agreed meanings.

References

- Breland, H. (1983). *The direct assessment of writing skill: a measurement review*. Technical Report No. 83–6. Princeton, NJ: College Entrance Examination Board.
- Bronfenbrenner, U. (1979). *The ecology of human development: experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Crisp, V. & Johnson, M. (2007). The use of annotations in examination marking: opening a window into markers' minds. *British Educational Research Journal*, **33**, 6, 943–962.
- Einhorn, H. J. (2000). Expert judgement: some necessary conditions and an example. In: T. Connelly, H. R. Arkes & K. R. Hammond (Eds.), *Judgement and decision making: an interdisciplinary reader*. 2nd edition, 324–335. Cambridge: Cambridge University Press.
- Elander, J. & Hardman, D. (2002). An application of judgement analysis to examination marking in psychology. *British Journal of Psychology*, **93**, 303–328.
- Engeström, Y. (2001). Expansive learning at work: toward an activity theoretical reconceptualization. *Journal of Education and Work*, **14**, 1, 133–156.
- Huot, B. (1990a). The literature of direct writing assessment: major concerns and prevailing trends. *Review of Educational Research*, **60**, 237–263.
- Huot, B. (1990b). Reliability, validity and holistic scoring: what we know and what we need to know. *College Composition and Communication*, **41**, 210–213.
- Johnson, R. L., Penny, J. & Gordon, B. (2001). Score resolution and interrater reliability of holistic scores in rating essays. *Written Communication*, **18**, 2, 229–249.
- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference and consciousness*. Cambridge, MA: Harvard University Press.
- Kelly, G. A. (1955). *The Psychology of Personal Constructs*. New York: Norton.
- Laming, D. (2004). *Human judgment: the eye of the beholder*. London: Thomson Learning.
- Oates, T. (2004). The role of outcomes-based national qualifications in the development of an effective vocational education and training system: the case in England and Wales. *Policy Futures in Education*, **2**, 1, 53–71.
- Rapport, F., Wainwright, P. & Elwyn, G. (2004). "Of the edgelands": broadening the scope of qualitative methodology. *Journal of Medical Ethics; Medical Humanities*, **31**, 37–42.
- Sanderson, P. (2001). *Language and differentiation in Examining at A Level*. PhD thesis, University of Leeds.
- Vaughan, C. (1991). Holistic assessment: what goes on in the rater's mind? In: L. Hamp-Lyons (Ed.), *Assessing Second Language Writing in Academic Contexts*. Norwood, N.J.: Ablex Publishing Corporation.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge University Press.
- Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, **7**, 2, 225–246.
- Wiliam, D. (1998). *Construct-referenced assessment of authentic tasks: alternatives to norms and criteria*. Paper presented at the 24th Annual Conference of the International Association for Educational Assessment, Barbados.
- Wolf, A. (1995). *Competence-based assessment*. Buckingham: Open University Press.

ASSESSMENT JUDGEMENTS

Annotating to comprehend: a marginalised activity?

Martin Johnson Research Division and **Stuart Shaw** CIE Research

Introduction

One of the important premises underlying this article is that the cognitive processes involved in reading can play a significant role in assessment judgements. Although we acknowledge that not all assessments of performance rely on assessors appraising written texts, many tests use written evidence as an indicator of performance. As a result, it is important to consider the role of assessors' comprehension building when reading candidates' textual responses, particularly where candidates are offered a greater freedom in determining the form and scope of their responses.

Crisp and Johnson (2007) note that it is common practice for examiners to annotate scripts when marking. This convention is formalised in the Qualifications and Curriculum Authority (QCA) code of practice (QCA, 2007) which stipulates that a second assessor needs to see any annotations made by a first assessor to gain a full and clear understanding of whether the marking criteria have been applied as intended. Beyond this formalised role, annotation might perform a more general and less formalised function in individual reading comprehension building processes.

Sources (Weiner and Simpson, 2005; Merriam-Webster, 2005) suggest that the definition of the word 'annotation' is to be found in the

15th Century Latin word 'annotare' meaning 'to note or to mark'. The historical importance of the activity is highlighted by Manguel (1997) and Wolfe and Neuwirth (2001) who suggest that it provided a social function, facilitating the sharing of meaning in mediaeval literary cultures. Modern annotation, however, tends to be defined as a discrete activity. Most commonly it is defined as an explanatory note (Weiner and Simpson, 2005), a note added by way of comment or explanation (Merriam-Webster, 2005), a short definition (Nation, 1983), an explanation of word meaning (Pak, 1986), or a critical or explanatory commentary or analysis added to a text (Wiktionary, 2008). Some definitions also allude to the wider impact of annotating on the annotator and any other subsequent reader. Cousins *et al.*, (2000) define annotation as a commentary on an object that the annotator intends to be, and the reader interprets to be, separable from the object itself.

This article considers how annotation might influence reader comprehension building at an informal personal level whilst also fulfilling other more formal functions within assessment processes. It goes on to explore how constraining this informal personalised activity might also influence those comprehension building processes. In order to explore how annotation may impact on text comprehension it is first necessary to ascertain what the literature reveals about the various theories and models of reading comprehension.

Models of reading comprehension

Reading is a complex cognitive activity. Attempts to articulate understandings of the reading comprehension process are neither new nor simple (see Huey in Anderson and Pearson, 1988). Prevailing language processing theories offer insights into the mental processes involved in readers' text comprehension when engaging in different types of real life reading. The intricacy of the cognitive processing activities involved in reading are described in varying degrees of detail by Alderson, 2000; Birch, 2007; Cohen and Upton, 2006; Field, 2004; Grabe and Stoller, 2002; Kintsch and van Dijk, 1978; Perfetti, 1999; Rayner and Pollatsek, 1989; and Urquhart and Weir, 1998. Weir and Khalifa (2008) provide a very helpful overview of a range of contributions to the body of theory concerning reading comprehension. Most of the literature cited draws heavily on first language (L1) research and many of the established theories concerning reading comprehension and language processing resonate with current thinking in the fields of cognitive psychology, psycholinguistics, and language assessment.

Developments in reading research over the last century have highlighted significant shifts in the way that the reading process has been perceived: moving from a bottom-up to a more integrated interactive model via a top-down approach.

Bottom-up processing models of reading comprehension require that the reader utilises a range of orthographic, phonological, lexical, syntactic units in order to progress along the scale of linguistic processing. Beginning with recognition of individual letters, followed by words and then sentences, the reader converges on a sense of textual meaning at both a local and global level. Comprehension on a global level relates to propositional understanding (literal interpretation) beyond the level of the text's microstructure and involves the reader's background knowledge along with their ability to identify arguments; recognise central concepts, key details and textual features, such as gist, coherence, cohesion and rhetorical structure. Local comprehension is related to linguistic knowledge (Cohen and Upton, 2006) and takes place on the micro-

structural, sentence and clause level. Local comprehension is associated with the understanding of micro-propositions such as word meaning and memory, sentential syntax, and textual details, amongst other things.

In top-down models of processing, comprehension is accomplished through the integration of incoming information with the reader's existing knowledge structures. Propositional meaning, or literal interpretation, is built and developed as readers combine what they encounter in the text with the linguistic, content and cultural knowledge they bring to the text. Thus in the act of reading, readers employ existing schemata to both develop 'meaning representation of the text so far' (Weir and Khalifa, 2008, p.6) and to predict subsequent text.

In the interactive model of reading comprehension, processing takes place in both directions, proceeding simultaneously:

Reading involves the simultaneous application of elements such as context and purpose along with knowledge of grammar, content, vocabulary, discourse conventions, graphemic knowledge, and metacognitive awareness in order to develop an appropriate meaning. (Hudson, 1991, p.83).

Presently, it is widely held that readers construct meaning by processing at different levels concurrently, employing both top-down and bottom-up processing.

Reading comprehension as a metacognitive activity

In their forthcoming volume, *Examining Reading*, Weir and Khalifa present a cognitive processing approach to defining reading comprehension. They identify from the literature within the field of cognitive psychology, certain generic cognitive processes which contribute to the reading process. The cognitive model they use is based on the earlier work of Urquhart and Weir (1998) which expanded Just and Carpenter's (1980, 1987) model and incorporated components from Kintsch and van Dijk (1978, 1983). Central to the model is an understanding of three key constituent features: the *goal setter*, the *processing core*, and the *monitor*. What follows is a very brief description of the role and function of each of these three components. These are considered to be important because annotation may interact with these components and influence key metacognitive functions that facilitate reading comprehension.

The overall goal of reading activity is determined by the *goal setter* which also selects the form of reading which is most likely to realise that goal. Having established a purpose for the reading, the reader is better placed to identify and select the most suitable strategies and determine the type and nature of information which needs to be targeted in the text. Urquhart and Weir (1998) present a helpful matrix in which they identify reading strategies and skills (careful and expeditious) at the local and global levels. 'Strategies' can be thought of as cognisant analytic activities and 'skills' as subliminal, perfunctory abilities (Cohen 1998; Urquhart and Weir, 1998).

The *central processing* element characterises a sequence of reading behaviours. Weir and Khalifa (2008) describe each of these behaviours in detail. Visual recognition, which constitutes the first level of processing, comprises *word recognition* and *lexical decoding*. Word recognition relates to the matching of the form of a word as manifested in written text with a mental representation of the orthographic forms of the language. According to Field (2004), lexical access/decoding, is the 'retrieval of a lexical entry from the lexicon, containing stored

information about a word's form and its meaning'. The next level of processing constitutes an important feature of comprehension. *Syntactic parsing* is concerned with the assembling of words into larger textual units and helps to establish *propositional (core) meaning* at clause and sentence level:

Propositional meaning is a literal interpretation of what is on the page. The reader has to add external knowledge to it to turn it into a message that relates to the context in which it occurred. (Weir and Khalifa, 2008, p.9).

Inferencing, the next higher order level of processing, is a necessary and creative process resulting in the addition of information brought to the text by the reader in an attempt to make the text more meaningful. The reader is now in a position to *build a mental representation* (or model) of the text:

... incoming information has to be related to what has gone before, so as to ensure that it contributes to the developing representation of the text in a way that is consistent, meaningful and relevant. This process entails an ability to identify main ideas, to relate them to previous ideas, distinguish between major and minor propositions and to impose a hierarchical structure on the information in the text. (Field, 2004, p.241).

Creating a text-level structure constitutes the final phase of language processing in which a discourse-level structure is constructed for the entire text.

The *monitor* is a mechanism which provides the reader with feedback regarding the efficacy of the selected reading process. There is a 'symbiotic' relationship between the monitor and goalsetter in that the monitoring process is reliant upon the decisions taken with regard to the type of reading and, therefore, the monitor is triggered in accordance with the goalsetter. Thus each component acts as a metacognitive device that mediates among the reader's range of processing skills and knowledge sources.

Thus the reading process can be thought of as an interaction of the reader's conceptual abilities and process strategies, and their language knowledge and content knowledge.

Annotation as a support for reading comprehension

Anderson and Armbruster (1982) and O'Hara (1996) identify a number of written support activities that are commonly associated with reading. This evidence has led some (O'Hara and Sellen, 1997; Marshall, 2001) to observe that such activities can often operate concurrently with reading activity, frequently being seamlessly integrated with reading activity, and habitually being unselfconsciously generated by the annotator. Wolfe and Neuwirth (2001) also cite a study by Adler *et al.* (1998) which found reader annotation activity occurring in conjunction with reading activity more than 25% of the time, with an additional 22% of annotations being made on separate documents from the reading source document. It appears that the reason for the existence of such practices could relate heavily to the cognitive processes involved with reading comprehension. This observation is supported by research evidence which has found that the complexity of a reading task influences reading performance (Mayes *et al.*, 2001; Just and Carpenter, 1987). Weir (2005) theorises the cognitive complexity of such processes. One particular area of the central

processing core appears to be of specific interest when discussing annotation practices. The process of building a mental model of a text involves reader self-monitoring, which in turn involves the use of working memory. It appears that annotation might perform an important function in mediating reader workload and enhancing comprehension.

There is a body of research which explores how annotating might actually support comprehension building processes. Hsieh *et al.* (2006) highlight evidence from Hartley and Davies (1978) that annotating facilitates textual encoding during the reading process. Textual encoding involves the basic perceptual process of converting a sensory input into subjectively meaningful experience. This encoding process plays a central role in reading comprehension. Weir and Khalifa (2008) outline how the central processing core involves a reader building a mental model of a text through integrating visual textual information with their world knowledge. Annotating might play an important role in this integration. The reason for this might be explained by the way that annotating involves the active integration of a reader's present understanding with new information encountered within the text. Sometimes this might involve the reader paraphrasing or elaborating on textual information in the form of an annotation.

Another important aspect of encoding also involves spatial encoding. Piolat *et al.* (1997) argue that a number of research findings are consistent with the idea that spatial encoding occurs during reading activity and that this is an integral part of a reader building a material representation of the location of textual information. In other words, reading is a spatial activity with the reader's eyes moving from one fixation location to the next to pick up spatially distributed visual information and processing positional information. This interpretation is corroborated by Fischer (1999) who argues that there is both direct and indirect evidence to suggest that memory is used to process information about spatial attributes of texts during reading. This work implies that the act of reading involves the mental spatial tagging of ideas and concepts in a text rather than the tagging of the location of words alone. Such research evidence also reinforces the postulation by Kennedy (1992) of a 'spatial coding hypothesis'. This hypothesis intimates that readers consider texts to behave as physical objects which provide the reader with spatial code in addition to lexical information. A tangible outcome of this hypothesis is demonstrated in studies that highlight how reader information recall correlates positively with increased reader annotation (Hartley and Davies, 1978; Hartley, 1983; Khan, 1994).

Annotating might also perform an important metacognitive function during reading. According to Weir and Khalifa (2008) self-monitoring is a complex metacognitive operation that provides the reader with feedback about the success of their reading processes. McMahon and Dunbar (2003) investigate tools that might support comprehension monitoring and suggest that annotation might support this function. This phenomenon was also observed in a study by Crisp and Johnson (2007). Examiners involved with assessing longer textual answers were observed making annotations whilst reading and they suggested that these annotations provided them with an individual checking function or a means to communicate with themselves about the text being read. This also appears to link to research observations which suggest that annotations might support such a metacognitive function by aiding working memory in a retrospective manner. Marshall (1997) reported that readers' annotations were being used as a visible trace of the reader's attention, especially when the text was in a protracted narrative form. Marshall suggested that these annotations could act as place

markings that subsequently aided the annotator's memory. This suggests that annotation can function as a storage bank of information external to individual working memory.

It might be important to reflect on the idea that annotating practices are potentially highly individualistic in character. Crisp and Johnson (2007) and Shaw (2008) found that some examiners were prone to using annotations idiosyncratically despite the clearly defined expectations of the mark schemes to which they were working. This relates to the view that annotations might be seen to represent the point of convergence between a reader's current knowledge and the propositional aspects of a text that they are encountering at a given time. It is reasonable to assume that the tangible outcome of that encounter would be particular to that situation. This reflexive quality might be important given research which suggests that highly individualistic note-taking can facilitate better information encoding and storage external from working memory (Hartley and Davies, 1978; Hartley, 2002).

Annotation and assessment

This article has highlighted the potential role of annotating on reading comprehension processes. Moreover, this activity is essentially an informal and potentially highly individualistic activity, influenced by the interaction of a variety of particular factors at a given time.

There have been relatively few studies that have looked at annotating activities in the context of educational assessment, but the limited literature suggests that annotations perform additional functions which are specifically linked to the context of large scale examination processes. Two recent studies at Cambridge Assessment suggest that assessor annotation performs a number of functions beyond supporting reader comprehension.

Crisp and Johnson (2007) report evidence of examiner annotations serving two distinct functions. The first function was to facilitate examiner judgements. The study found that examiners found annotating to be particularly useful for reinforcing their comprehension of protracted texts. The second function was a justificatory one, where annotations communicated the reason for a judgement to other assessors within the system. In this sense annotating supported the confidence of examiners to complete their marking in the knowledge that others would be aware of the reasoning behind their assessment judgements. This confidence factor also parallels the findings of Shaw (2005) who found that examiners used annotations to investigate their own marking consistency. Shaw observed that annotations were used by examiners to provide an efficient means to confirm, deny or reconsider their marking standards both within and across candidates, thereby reassuring the examiners throughout the marking event.

The findings of these studies suggest that annotating activities in large scale assessment systems might be influenced by competing demands beyond the basic requirement to support individual examiner comprehension building. The reason for this might be explained by the accountability concerns attached to large scale assessment, and the related objective of maximising transparent communication within the assessment system.

The case of large scale examinations in the UK provides a useful context within which to discuss this issue. The accountability agenda that pervades education has led to public examinations being the most widely used performance indicators for educational success. The scale of this exercise requires measures to ensure that the examination system

functions in a fair and robust manner. The Code of Practice (QCA, 2007), produced by the examination regulatory authorities in England, Wales and Northern Ireland, outlines procedures that awarding bodies should follow to ensure that examinations are developed and administered within transparent and accountable structures. There is an unambiguous emphasis on clear communication channels between examiners of different seniority to facilitate effective monitoring. Williamson (2003) comments that this function is all the more important in an expanding examinations system, such as that of the UK. Annotations have an important communicative role in this quality control process.

The importance of justificatory annotating undoubtedly influences examiners' practices. This is visibly demonstrated by the extent to which annotating activity is documented in assessment guidelines. Besides the guidance given in mark schemes, at marker standardisation and at coordination meetings, expectations are set out by the QCA. These formalised arrangements state that internal coursework assessments and associated assessment criteria must indicate how credit has been assigned, and that therefore 'Internal assessors are required to annotate the coursework, clearly showing how the marking criteria have been applied' (QCA, 2005, p.19). The most recent QCA code of practice (QCA, 2007) also requires that principle moderators must 'compile exemplar work, annotated to show how the assessment criteria are to be applied' (p.9), in order to ensure that the standards of the unit or component are maintained and consistent with the unit specification and assessment criteria. The code of practice also requires awarding bodies in the UK to 'continue to mark and annotate all scripts in accordance with good practice recognised by the regulatory authorities' (QCA, 2007, p.49).

Wolfe and Neuwirth (2001) suggest that annotations can fulfil a variety of functions, although it appears that in different contexts some functions may dominate others. Wolfe and Neuwirth suggest that annotations can facilitate reading and later writing tasks; eavesdrop on the insights of other readers; provide feedback to writers or promote communication with collaborators; and call attention to topics in important passages. The emphasis in formalised assessment discourse about annotation practices appears to accentuate the functions of eavesdropping on the insights of readers and promoting collaboration with others rather than reading facilitation. The consequence of this in assessment practice is that certain annotation conventions come to be considered acceptable and become 'expected' practice in order to promote transparency and consistency amongst examiners. As a result, it might be argued that the prioritisation of the accountability function could lead to the demotion of the comprehension building function which might rely more on flexible and individualistic annotation practices.

Literature suggests that the mode in which a text is presented, either on paper or on screen, also represents another area where systematic pressures come to bear on annotating practices. O'Hara and Sellen (1997) argue that mode can affect reader annotation in a number of ways. One major concern is the degree of physical effort required to annotate in one mode compared with another. They suggest that making paper-based annotations is a relatively effortless procedure and, as a consequence, it factors automatically into the meaning construction process during reading. In contrast, computer-based annotation practices can be impeded by the availability of authentic annotation tools. Keyboards might influence annotating behaviour because they do not accommodate many of the types of mark that readers choose to use when working on paper, therefore making the process less genuine and positively affecting the cognitive demand on the reader.

Again, there is limited empirical literature on modal influence on assessor annotation practices. In one study involving higher education instructors, Price and Petre (1997) present a mixed picture of modal influence. They found that all the instructors in their study had different marking styles on paper and on screen, but that the extent of this modal influence varied between markers. They also found that the annotations used in paper and electronic marking were different, with underlining, circling and highlighting being used less in electronic marking than on paper, despite their availability. Despite these differences, Price and Petre conclude that technology did not impair the assessment practices, and more specifically the annotating processes, for all markers. Other studies suggest emotive and physical dimensions in relation to computer annotating. Greatorex (2004) reports teacher frustration when moderating electronic portfolios. Her study highlights the difficulties that teachers experienced when annotating candidates' work directly, with teachers reporting that there would have been more annotation if portfolios had been presented on paper. Shaw (2008) observes that marker concentration might be adversely affected when assessing on-screen. Not being able to replicate paper and pen practice when applying annotations was a predominant concern amongst trial participants in his study. Participants generally perceived on-screen marking to be physically more demanding than paper marking. Moreover, marking over prolonged periods engendered mental and physical fatigue with the physical process of selecting and applying annotative tools on-line being demanding.

Conclusion

Crisp and Johnson (2007) have suggested that one of the two functions annotations serve is justificatory and that annotating might have an important communicative role in the quality control process in terms of accountability. Annotating has a particular role in assisting with transparent communication between different markers (Williamson, 2003). Accountability is widely recognised to be a multifaceted and complicated concept (Day and Klein, 1987) and 'assumes the requirement to answer to the broader social community' (Kogan, 1986). In an educational context, examination boards offering high-stakes assessments are required to account for or justify certain assessment actions and behaviour for a range of potential community stakeholders. Thus, the notion of accountability is closely related to responsibility, as those who have been given responsibilities – the assessment practitioners – are asked to account for their assessments. If the conventional accountability processes are influenced by the introduction of a new, computer-assisted assessment medium then both the reliability of test scores and the validity of the assessments are potentially compromised.

By bringing together literature about linguistics and annotation practices, both empirical and theoretical, this article suggests that a critical link exists between annotating and reading activities. Moreover, an important aspect of this relationship is associated with reader comprehension building. It is perhaps significant that empirical study into annotating in assessment contexts is very limited and this helps to explain why the extent to which annotating candidate responses influence or affect assessor comprehension is neither known nor fully understood. This is an important observation since the arguments advanced in this article suggest that such an influence is tangible.

Through making the different functions of annotation explicit the intention of this article is to primarily amplify the importance of the impact of annotating on assessor comprehension. It is also intended that

this function be clearly understood in relation to the other accountability and transparency purposes that currently influence how annotation is used in large scale assessment systems. It is also worth noting that this issue has potentially important consequences for ongoing debates about on-line assessment both within Cambridge Assessment, through the current series of on-screen marking trials, and beyond. It is hoped that this article can make a positive contribution to this area of discussion.

References

- Adler, A., Gujar, A., Harrison, B. L., O'Hara, K., & Sellen, A. (1998). *A Diary Study of Work-Related Reading: Design Implications for Digital Reading Devices*. Proceedings of CHI '98, 241–248.
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Anderson, R. C. & Pearson, P. D. (1988). A Schema-theoretic View of Basic Processes in Reading Comprehension. In: P. L. Carrell, J. Devine & D. E. Eskey (Eds.), *Interactive Approaches to Second Language Reading*. Cambridge: Cambridge University Press.
- Anderson, T. H. & Armbruster, B. B. (1982). Reader and text-studying strategies. In: W. Otto, & S. White (Eds.), *Reading Expository Material*. London: Academic Press.
- Birch, M. (2007). *English L2 Reading: Getting to the Bottom*. Mahwah: Lawrence Erlbaum Associates.
- Cohen, A. (1998). *Strategies in Learning and Using a Second Language*. London: Longman.
- Cohen, A. D. & Upton, T. A. (2006). *Strategies in Responding to the New TOEFL Reading Tasks* (Monograph No. 33). Princeton, NJ: Educational Testing Services.
- Cousins, S. B., Baldonado, M. & Paepcke, A. (2000). *A Systems View of Annotations*. Technical Report P9910022. Palo Alto, CA: Xerox/Palo Alto Research Center.
- Crisp, V. & Johnson, M. (2007). The Use of Annotations in Examination Marking: Opening a Window into Markers' Minds. *British Educational Research Journal*, **33**, 6, 943–961.
- Day, P. & Klein, R. (1987). *Accountabilities: Five Public Services*. London: Tavistock.
- Field, J. (2004). *Psycholinguistics: the Key Concepts*. London: Routledge.
- Fischer, M. H. (1999). Memory for Word Locations in Reading. *Memory*, **7**, 1, 79–118.
- Grabe, W. & F.L. Stoller (2002). *Teaching and Researching Reading*. London: Longman.
- Greatorex, J. (2004). *Moderated E-portfolio Project Evaluation*. Cambridge: University of Cambridge Local Examinations Syndicate.
- Hartley, J. (1983). Note-taking Research: Resetting the Scoreboard. *Bulletin of the British Psychological Society*, **36**, 13–14.
- Hartley, J. (2002). Notetaking in Non-academic Settings. *Applied Cognitive Psychology*, **16**, 559–574.
- Hartley, J. & Davies, I. K. (1978). Note-taking: A Critical Review. *Innovation in Education and Teaching International*, **15**, 3, 207–224.
- Hsieh, G., Wood, K. R. & Sellen, A. (2006). *Peripheral Display of Digital Handwritten Notes*. Proceedings of the Conference on Human Factors in Computing Systems, Montreal, Quebec, 2006.
- Hudson, T. (1991). A Content Comprehension Approach to Reading English for Science and Technology. *TESOL Quarterly*, **25**, 1, 77–104.
- Just, M. A. & Carpenter, P. A. (1980). A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review*, **87**, 4, 329–354.
- Just, M. A. & Carpenter, P. A. (1987). *The Psychology of Reading and Language Comprehension*. Boston: Allyn and Bacon.
- Kennedy, A. (1992). The Spatial Coding Hypothesis. In: K. Rayner (Ed.), *Eye Movements and Visual Cognition*. New York: Springer-Verlag.

- Khan, F. (1994). *A Survey of Note-taking Practices*. HP Labs Technical Reports HPL-93-107.
- Kintsch, W. & van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review*, **85**, 363-394.
- Kogan, M. (1986). *Educational Accountability: An Analytic Overview*. London: Hutchinson.
- Manguel, A. (1997). *A History of Reading*. London: Flamingo.
- Marshall, C.C. (1997). *Annotation: From Paper Books to the Digital Library*. Proceedings of the Second ACM International Conference on Digital Libraries; Philadelphia, Pennsylvania.
- Marshall, C.C. (2001). *The Haunting Question of Intelligibility*. Paper presented at the ACM Conference on Hypertext and Hypermedia, Aarhus, August 14-18, 2001.
- Mayes, D. K., Sims, V. K. & Koonce, J. M. (2001). Comprehension and Workload Differences for VDT and Paper-based Reading. *International Journal of Industrial Ergonomics*, **28**, 367-378.
- McMahon, M. & Dunbar, A. (2003). Mark-UP: Facilitating Reading Comprehension through On-Line Collaborative Annotation. In: N. Smythe (Ed.), *Digital Voyages*. Proceedings of the Apple University Consortium Conference, Adelaide, South Australia, September 28 - October 1, 2003.
- Merriam-Webster (2005). *Collegiate® Dictionary, Eleventh Edition*. Springfield, MA: Merriam-Webster.
- Nation, I. S. P. (1983). *Teaching and Learning Vocabulary*. Wellington: English Language Institute, Victoria University.
- O'Hara, K. (1996). *Towards a Typology of Reading Goals*. Rank Xerox Research Centre Affordances of Paper Project Technical Report EPC-1996-107. Cambridge: Rank Xerox Research Centre.
- O'Hara, K. & Sellen, A. (1997). A Comparison of Reading Paper and Online Documents. In: S. Pemberton (Ed.), *Proceedings of the Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery.
- Pak, J. (1986). *The Effect of Vocabulary Glossing on ESL Reading Comprehension*. Unpublished manuscript. University of Hawaii at Manoa.
- Perfetti, C. A. (1999). Comprehending Written Language: A Blueprint for the Reader. In: C.M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language*. Oxford: Oxford University Press.
- Piolat, A., Roussey, J.-Y. & Thunin, O. (1997). Effects of Screen Presentation on Text Reading and Revising. *International Journal of Human-Computer Studies*, **47**, 565-89.
- Price, B. and Petre, M. (1997). *Teaching Programming through Paperless Assignments: An Empirical Evaluation of Instructor Feedback*. Milton Keynes: Centre for Informatics Education Research, Open University.
- QCA (2007). *GCSE, GCE, GNVQ and AEA Code of Practice*. London: QCA.
- QCA (2005). *A Review of GCE and CSE Coursework Arrangements*. London: QCA.
- Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice Hall.
- Shaw, S. (2005). *On-screen Marking: Investigating the Examiners' Experience through Verbal Protocol Analysis*. Internal ESOL Validation and Research Report.
- Shaw, S. (2008). *On-screen Marking of Extended Writing: Towards an Understanding of Examiner On-line Assessment Behaviour*. Internal CIE Research Report.
- Urquhart, A. H. & Weir, C. J. (1998). *Reading in a Second Language: Process, Product and Practice*. London: Longman.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic.
- Weiner, E. S. C. & Simpson, J. A. (Eds.) (2005). *Compact Oxford English Dictionary*. Oxford: Oxford University Press.
- Weir, C. J. (2005). *Language Testing and Validation: An Evidence-Based Approach*. Basingstoke: Palgrave Macmillan.
- Weir, C. J. & Khalifa, H. (2008). A Cognition Processing Approach Towards Defining Reading Comprehension. *Research Notes 31*, February 2008, 2-10.
- Wiktionary (2008). <http://en.wiktionary.org/wiki/annotation>
- Williamson, P. (2003). Setting, Marking and Awarding: The Examination Process. In: K. Tattershall, J. Day, H. James, D. Gillan & A. Spencer (Eds.), *Setting the Standard*. Manchester: AQA.
- Wolfe, J. L. & Neuwirth, C. M. (2001). From the Margins to the Center: The Future of Annotation. *Journal of Business and Technical Communication*, **15**, 3, 333-371.

EXAMINATIONS RESEARCH

Cookery examined – 1937–2007: Evidence from examination questions of the development of a subject over time

Gill Elliott Research Division

Introduction

The teaching of cookery skills in UK schools has become the subject of much debate in recent years. Like its counterpart, needlework, the subject has a history of social change and gender bias. In the early twentieth century, when school examinations began to become widespread, both subjects were highly used in a domestic context. In other words, they were life skills, for at least some part of the population. Initially, undoubtedly, both cookery and needlework were subjects undertaken by girls, in the same way as woodwork and metalwork were 'for' boys. In the 1970s and

early 1980s there was more integration of boys to the subjects. However, as school subjects, they became increasingly a minority option by both sexes, until they almost disappeared altogether in the 1980s.

As we approach the end of the first decade of the twenty first century, needlework remains a minority option at GCSE, mostly taken by girls (across all awarding bodies in 2006, 45,950 girls took the textiles option of Design & Technology GCSE as opposed to 1,515 boys) and is no longer necessary to any individual as a 'life skill' – nobody suggests that the 21st century family should return to making a substantial number of their own clothes, as was commonly the case into the 1950s at least.