

What do our school reports really say?

Kristy L. Turner, Fraser Scott et al

Introduction

“As we read the school reports on our children, we realise with a sense of relief that can rise to delight, that nobody is reporting in this fashion upon us” J.B.Priestley, *Delight*, 1949

School reports are an enduring feature of the education landscape. They form part of our personal history, fondly retained by parents well beyond a child’s school leaving age. The Department for Education requires schools in England to report to parents annually (Department for Education, 2015). There is widespread variation in reporting practice and many schools are doing more than is legally required of them (Power and Clark, 2000). While frequent, data focused reports are commonly used, many schools continue to write comment-based reports as part of their reporting regime. As students move into secondary school, reports of their day to day learning become less forthcoming from the students themselves and reports become one of very few channels of home-school communication.

The language used in reports is important; parents commonly express frustration in popular media about mixed messages, errors and impersonal reports (Weale, 2015). At Bolton School, parents receive a comment-based “full” report once a year in addition to more frequent, short, data derived reports. Comments are written by staff in free text format with a restricted character limit; comment banks are not used. These reports represent one of the public faces of the school and as colleagues, we were interested to find out what information could be gleaned from them about our values as a school.

In this paper we report the initial findings of a research project carried out by teachers at Bolton School Boys' Division to analyse the messages we communicate in our written reports.

Methods

A team of 5 teacher researchers examined a sample of Year 11 full reports using inductive thematic analysis. Each statement within the reports’ written comments was coded independently and iteratively refined until a consensus was met. Most of the emergent themes aligned to the VESPA model of non-cognitive skills for success proposed by Oakes and Griffin (2016); however, two additional themes were identified, B (for behaviour) and I (for intelligence), giving our ‘VESPABI’ model (Figure 1).

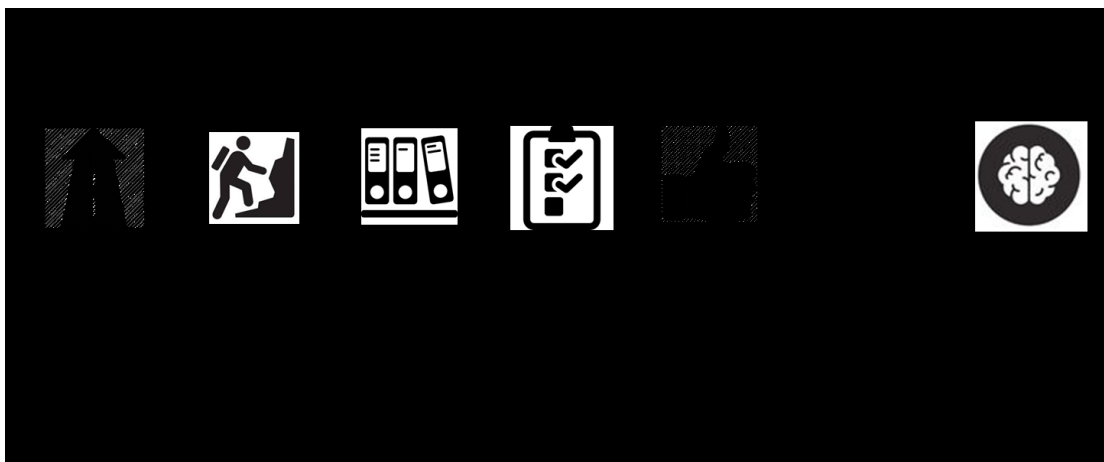


Figure 1: The VESPABI model.

We next sought to generate a quantitative understanding by using the VESPABI model deductively to analyse the written comments within the reports. The lead researcher anonymised a sample of Year 11 full reports (n = 116) consisting of comments from subject teachers and pastoral staff. In order to assess the degree that coders consistently assigned statements to themes of the VESPABI model, an inter-rater reliability estimate was evaluated by a fully-crossed design (Hallgren, 2012). 13 student reports were each coded by 5 teacher researchers and the inter-rater reliability calculated; however, these were not included in the subsequent analysis. The remaining reports from the year group sample (n = 103) were divided between the 5 teacher researchers and were coded independently.

The codes used corresponded to the VESPABI model, but were also assessed to determine if they were positive (+), negative (-) or neutral (=). A neutral phrase could be where both a positive and negative indicator was used for a particular code, or be a general piece of advice that does not give a clear indication of a student's particular strength or weakness,

Below are 3 example phrases which discuss organisation and so would be coded with the S (systems) code.

(+S) Joe is well organised and is on track with his design project.

(-S) Joe's exercise book is disorganised with missing sheets and blank pages apparent.

(=S) Joe would benefit from organising his file.

Phrase 3 could be interpreted in 2 ways. It could be considered that everyone would benefit from organising their file, and so a neutral S code would be applied. However another interpretation would be that it is implied that Joe's file is not well organised, and so a -S code could be applied. Such comments generated a lot of discussion amongst the teacher researchers.

Findings

Do teachers agree on what the statements in reports mean?

Inter-rater reliability was assessed using Gwet's AC1 Coefficient giving a value of 0.803 (95% Confidence Interval of 0.781 to 0.824, $P < 0.001$) (Gwet, 2008). This was determined to be 'very good' or 'excellent' agreement as defined by Altman's (1991) or Fleiss' (1981) benchmark scales, respectively. However, initial discussion between teachers generated a great deal of discourse about what the phrases used by colleagues really meant. During this discussion it was apparent that teachers were able to decode the meaning behind their colleagues' phrasing due to their knowledge of teaching particular subjects. When coding the responses teacher researchers were encouraged to look for transparent use of particular statements, such that the meaning would be clear to a parent, rather than extrapolate from their own experience.

The big picture

Only 24.1% of statements across all codes were categorised as either positive or negative, with 75.9% categorised as neutral. This surface level analysis is revealing, suggesting that many of the written comments report little in the way of specific observation of a pupil's non-cognitive skills. Although of course they may report other details such as rate of progress in a subject.

Are some types of statements used more frequently and with a particularly positive or negative focus?

The frequency of non-neutral statements was investigated and revealed that statements coded as P and E were most frequently used (figure 2, panel A). V and B coded statements were used rarely, followed by S, A and I coded statements. This may be expected given that the reports were written about pupils nearing their GCSE examinations and teachers may prioritise commenting on perceived effort and practice.

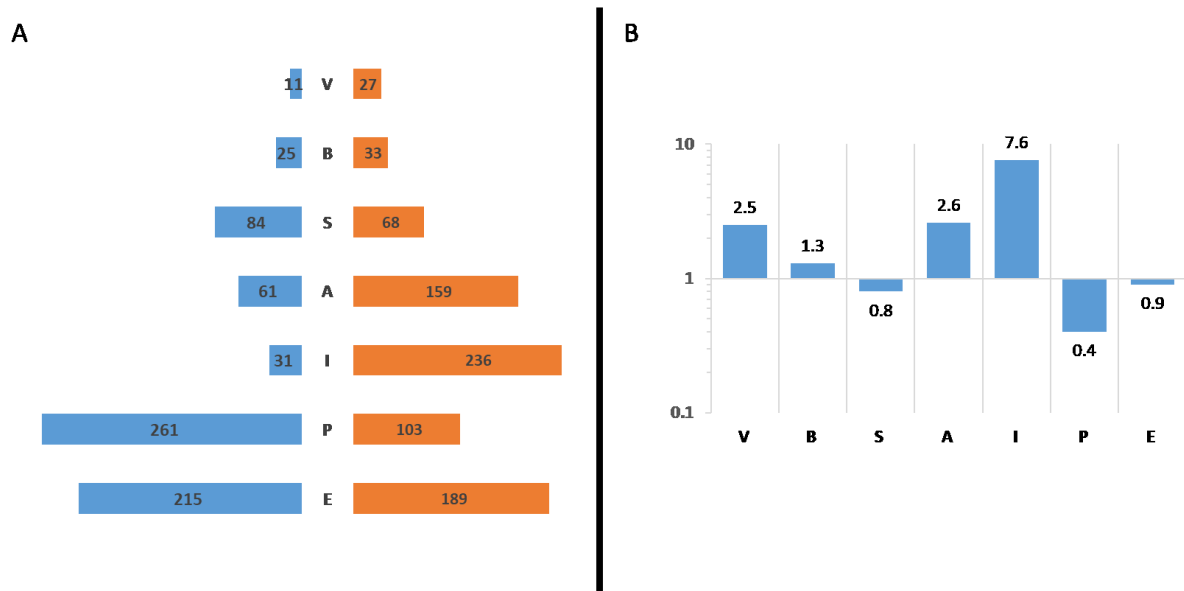


Figure 2: Panel A, frequency of non-neutral statements (positive in orange, negative in blue) of each code ordered from least frequent (top) to most frequent (bottom). Panel B, ratio of positive to negative statements for each code.

The ratio of positive statements to negative statements for each code was also investigated (figure 2, panel B). Statements coded as B, S or E were used positively and negatively with approximately the same frequency. Statements coded as V, A or I were used more positively. Perhaps predictably, teachers were much more likely to report a student was talented or intelligent than to report that they were finding a subject difficult; statements coded as I were 7.6 times more likely to be found as a positive statement rather than a negative statement. Statements coded as P were found to be 2.5 times more likely to be negative statements, again this may be a factor of the report's proximity to external examinations.

Is everything reported potentially actionable by students?

Statements commenting on an individual's intelligence or talent represent the third most coded statement; and it is the most positively used code. There are several reasons why a teacher may comment on intelligence, perhaps to boost confidence or to increase motivation. However, studies suggest that praising ability has negative consequences for students' achievement motivation than praise for effort (Mueller and Dweck, 1998). So given the limited space available, why bother comment on intelligence at all?

Conclusion and recommendations

School reports represent a key communication channel between school and home. It is important that schools are aware of the messages they convey, whether intentionally or unintentionally. The coding method we have described can provide a framework for analysis and generate discourse between colleagues. Our preliminary findings indicate that much of what is written are generic, unactionable statements. Eliminating these statements may lead to clearer guidance, and greater benefits, for students, and also a lighter workload for

teachers. Our follow-up studies include a more thorough investigation of the propensity to comment on 'intelligence' within reports, and also the potential differing interpretations between teachers and parents of comments within reports.

References

Department for Education (2015) *School reports on pupil performance: guide for headteachers*. London: Home Office. Available at: <https://www.gov.uk/guidance/school-reports-on-pupil-performance-guide-for-headteachers> (accessed 2019).

Power S and Clark A (2000) The right to know: parents, school reports and parents' evenings. *Research Papers in Education* 15(1): 25-48.

Wheale S (2015) 'Teachers and parents criticise 'robotic' software-generated school reports', *The Guardian*, 17 July. Available at: <https://www.theguardian.com/education/2015/jul/17/teachers-parents-criticise-robotic-software-generated-school-reports> (accessed 2019).

Oakes S and Griffin M (2016) *The A Level Mindset: 40 activities for transforming student commitment, motivation and productivity*. Carmarthen: Crown House Publishing.

Hallgren KA (2012) Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol.* 8(1): 23–34.

Gwet L (2008). Computing Inter-Rater Reliability and its Variance in the Presence of High Agreement. *British Journal of Mathematical and Statistical Psychology.* (61): 29 – 48.

Altman DG (1991) *Practical statistics for medical research*. London: Chapman and Hall.

Fleiss J (1981) *Statistical Methods for Rates and Proportions*. New Jersey: John Wiley & Sons.

Mueller CM and Dweck CS. Praise for intelligence can undermine children's motivation and performance. *Journal of personality and social psychology.* 75 (1): 33-52.