BACK TO BASICS? IDENTIFYING EDUCATIONAL NEEDS IN STATISTICAL CONSULTING

Sue Finch, Ian Gordon and Graham Hepworth
Statistical Consulting Centre
The University of Melbourne
sfinch@unimelb.edu.au

In statistical consulting the educational needs of clients can vary widely, and a consultant may need to work to develop statistical thinking and/or analytic competencies within a limited timeframe. What can a consultant assume about a client's knowledge and educational needs? What is the relationship between clients' broadly-defined analytic competence and understanding of fundamental statistical concepts and ideas? What are the implications for consulting practice and training in a world with more, 'bigger data'? In our unique study, we carried out surveys of clients and consultants based on actual consultations between experienced applied statisticians and clients. We focus on the consultant data in addressing these questions.

STATISTICAL EDUCATION NEEDS IN A BIG WORLD OF DATA

"A mainstay of the way we have thought about needs for statistical literacy is a dichotomy drawn between the producers and the consumers of statistics; literacy has focused on the needs of the latter. But the value of such a dichotomy is eroding fast. Increasingly people are no longer just passive recipients of data-based reports" (Wild, 2017, p.32). Much has been written about the statistical literacy needs of data users across a broad range of disciplines as there is increasing exposure to data in many forms (see, for example, Wild, Utts & Norton (2018), and the *Statistical Education Research Journal* Special Issue on Statistical Literacy, volume *16*(1), 2017).

The Statistical Consulting Centre (SCC) at the University of Melbourne provides advice to clients from within and outside the University. The clients have a wide range of backgrounds and experience in quantitative research and analysis of data, but all have a practical applied need to collect, use and understand data. These clients minimally need good statistical literacy, and often more extensive statistical knowledge and skills. We carried out a survey of statistical consultations to identify the needs of professionals seeking to work within a quantitative framework across a broad spectrum of applications. Empirical research on this growing group of data users in a world of expanding data is relatively rare.

Framework for examining the statistical content of consultations

A cycle of statistical enquiry that characterizes "the way one acts and what one thinks about during the course of a statistical investigation" has been described by Wild and Pfannkuch (1999, p. 225). The cycle moves through Problem, Plan, Data, Analysis to Conclusion. We use this as a framework for presenting our findings on the statistical content of consultations.

TWO SURVEYS

Two surveys were carried out in relation to a particular project on which each client was receiving advice from the SCC. One survey was completed by the client, the other by the consultant. The data collection was carried out for 12 months commencing June 2015. A client was invited to participate (in person or by email) the first time s/he met with a consultant during the survey period. Any client was only surveyed once during this time, in relation to one project.

The surveys collected information in relation to the interactions the client and consultant had about the substantive content of the project. These interactions may have included telephone conversations, email correspondence, or in-person meetings. Together, these were referred to in the surveys as "the consultation process". Ethics approval was obtained from the University of Melbourne (Ethics approval ID number: 1543825).

Client survey

Eligible clients were University of Melbourne higher degree research students and staff and individuals with quantitative needs from outside academia. The latter included, for example, clients

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

from government, business and industry. Most consultants invited the participation of all eligible clients within the survey timeframe. One consultant who saw a large number of (post-graduate) clients randomly sampled clients to invite; the sampling of clients was done before the consultation. Clients who agreed to participate were sent a link to a survey to be completed via Survey Monkey.

Consultant survey

Consultants completed a survey via Google docs. They identified categories of activity and the statistical content covered in the consultation. The statistical content was described by a list of 89 different statistical items; statistical items referred to concepts, ideas, skills, procedures and principles. An example of a statistical item was "the distinction between samples and populations". The statistical items were grouped into 13 broad statistical themes.

The consultant indicated which of the five stages of the PPDAC cycle were covered. This is described further below. The complexity of the project and the statistical competence of the client was rated. Competence was initially rated in five categories; these were reduced to three categories for analysis – basic, middling or competent. This was a global judgement by the consultant of the client's capacity to work within a quantitative framework; it was based on the discussions and interactions between the client and consultant. As a group, the consultants had a substantial amount of consulting experience; prior to the survey they had routinely recorded information about the level of client statistical competence.

Five consultants participated in the study. The consultant survey was completed for 212 client consultations. A quarter of the consultations were for clients from outside the University.

FINDINGS

This paper focuses on data provided by the consultants; this represents the complex interactions of 5 consultants with over 200 clients. Overall, most clients were satisfied with the consultation they had been receiving; 94% gave an overall rating of the quality of the service as good or very good (highest option offered).

PPDAC cycle

The bottom line of Table 1 shows the percentage of consultations in which some aspect of each of the five stages of the PPDAC cycle was covered. About a quarter of project consultations included discussion relevant to formulation of the research question within a statistical framework – the problem stage – compared with over 80% for the analysis stage. Of course, many consultations covered more than one stage of the cycle.

Table 1: Percentage of consultations covering each stage of the PPDAC cycle by client competence and overall

Competence	Problem	Plan	Data	Analysis	Conclusion
Basic	26%	23%	35%	79%	44%
Middling	21%	26%	26%	82%	42%
Competent	31%	11%	23%	83%	33%
Overall	26%	20%	27%	82%	40%

Consultants rated clients' statistical competence as basic (n=57), middling (n=89) or competent (n=64). Table 1 shows the percentage of project consultations covering each stage of the PPDAC, according to the rating of client competence. The percentages for the different stages of the cycle are similar for all three levels of client competence. Hence we look in more detail at the broad statistical themes covered.

Broad statistical themes

Consultants indicated which of 89 different statistical items were covered in the project consultation. These were grouped into 13 broad statistical themes; a binary indicator was recorded

if any of the statistical items within each theme had been covered. The groups are thematically related ideas, concepts, procedures or activities; the items within the themes could reflect different complexity of depth of knowledge and understanding.

Figure 1 shows the percentage of project consultations covering each broad statistical theme, according to the rating of client competence; they are organized into relevant stages of the PPDAC cycle. Some themes might apply to more than one stage; four themes, for example, are grouped together under the Data and Analysis stages. Table 2 provides odds ratios from separate logistic regression analyses predicting any coverage of each of the 13 broad themes. It also gives a *p*-value under each broad theme; this is from an overall test of linearity.

Problem stage

Statistical items classified in the Problem stage of the PPDAC cycle included fundamental concepts about working in an empirical framework to collect data to answer a question. This theme of *Fundamentals* was covered in over half the project consultations with basic competence (Figure 1), the odds for those clients being more than two and a half times that for competent clients (Table 2). However *Fundamentals* were still covered for more than 30% of middling and competent clients.

Table 2: Odds ratios comparing levels of competence for the coverage of each of 13 broad statistical themes

Baseline	Level	OR	95% CI	OR	95% CI
		PROBLEM		DATA & ANALYSIS	
		Fundamentals		*Data handling*	
		(p = 0.007)		*(p = 0.005)*	
Competent	Middling	1.06	0.53, 2.12	1.24	0.51, 3.04
Competent	Basic	2.82	1.34, 5.92	3.56	1.47, 8.65
		PLAN			
		Units of observation		*Modelling issues*	
		(p = 0.419)		*(p = 0.003)*	
Competent	Middling	1.62	0.76, 3.47	0.63	0.32, 1.27
Competent	Basic	1.16	0.49, 2.76	0.28	0.13, 0.61
		Measurement & variables		*Modelling methods*	
		(p = 0.066)		*(p = 0.570)*	
Competent	Middling	2.12	1.03, 4.34	0.88	0.35, 2.17
Competent	Basic	2.21	1.01, 4.84	0.61	0.24, 1.59
		Design basics		*Graphs*	
		(p = 0.593)		*(p = 0.685)*	
Competent	Middling	1.24	0.62, 2.45	1.17	0.62, 2.23
Competent	Basic	0.86	0.39, 1.88	1.38	0.67, 2.82
				ANALYSIS & CONCLUSION	
		Sample size (n?)		*Understanding inference*	
		(p = 0.810)		*(p = 0.038)*	
Competent	Middling	1.24	0.58, 2.66	1.31	0.68, 2.50
Competent	Basic	1.28	0.55, 2.94	2.51	1.20, 5.22
		Sampling & randomization		*Applied interpretation*	
		(p = 0.245)		*(p =0.252)*	
Competent	Middling	1.89	0.80, 4.45	1.33	0.69, 2.57
Competent	Basic	1.99	0.79, 5.04	1.85	0.89, 3.82
				Reporting results (p= 0.022)	
Competent	Middling			2.03	0.83, 4.95
Competent	Basic			3.50	1.39, 8.81

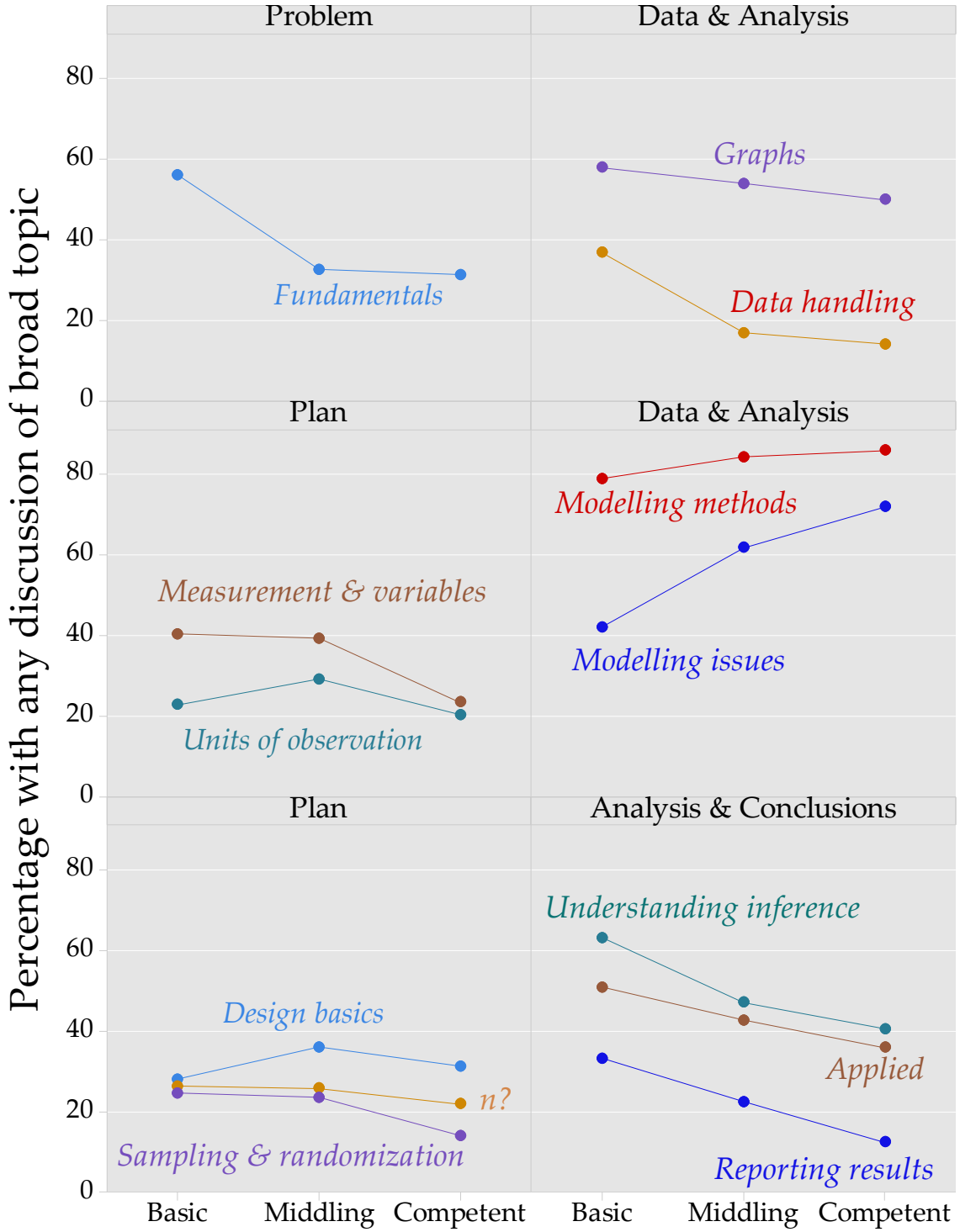

Figure 1: Percentage of project consultations in which each broad statistical theme was discussed by client competence
(*Applied* refers to *Applied interpretation*)

Plan stage

Five broad themes were considered in the Plan stage. For clarity, they are separated in two panels in Figure 1.

Three statistical items were about *Units of observation*, including ideas about the need for replication and to avoid false replication. *Measurement and variables* included, for example, quantification, writing questions, deriving and recoding variables, and ideas of reliability and validity, some of which might be considered to be part of the Data stage. Around a quarter of all project consultations covered *Units of observation*. About 40% covered *Measurement and variables* for basic and middling clients, the odds for these groups being double that for competent clients (Figure 1, Table 2).

Design basics, *Sample size* (labelled *n?* in Figure 1) and *Sampling and randomization* included 5 or 6 statistical items. Differences for *Design basics* and *Sample size* based on competency are less pronounced than for *Fundamentals*, for example. The odds of covering *Sampling and randomization* tended to be lower for competent clients than for the other two groups, although this is not precisely estimated (Table 2).

Data and Analysis stages

Four themes were relevant to the Data and Analysis stages; they appear in two panels in Figure 1. The differences between competence groups vary according to the broad theme (Figure 1). For practical aspects of *Data handling* those with basic competence needed more advice than those rated competent (Table 2). *Graphs* (for data exploration, checking assumptions, representing summary statistics and/or inferences) and/or the principles of good graphs were discussed in over half of all project consultations (Figure 1), but this was not strongly dependent on competence.

Modelling methods referred to statistical techniques, including for example *t*-tests and linear models. As might be expected, this was the most frequently discussed broad theme in over 80% of project consultations; this did not vary strongly by client competence. There were 12 different statistical items included under *Modelling methods*; the most frequently rated items were *Simple inference* (35% overall) including for example *t*-tests, *Linear models* (39%) and *Linear mixed models* (17%). The odds of covering *Simple inference* were 4.68 times for basic clients compared with competent (95% CI: 2.1 to 10.42), and 1.90 times for middling clients compared with competent (95% CI: 0.89 to 4.03). For *Linear models*, the odds were similar for middling and competent clients (OR 1.16, 95% CI: 0.60 to 2.23), and, 0.78 times for basic compared with competent (0.37 to 1.64). For *Linear mixed models*, the odds were 0.56 (95% CI: 0.25 to 1.25) for middling/competent and the odds were around one fifth for basic compared with competent (OR 0.22. 95% CI: 0.07 to 0.72).

In comparison with the results for *Modelling methods*, there were some more substantial differences for *Modelling issues*; the odds were 0.28 for basic relative to competent clients.

Analysis and Conclusion stages

We identified three broad themes relevant to moving from analysis to conclusions. *Understanding inference* included items about interpretation and misinterpretation of *p*-values and confidence intervals, over-reliance on *p*-values, and the quantification of uncertainty. *Applied interpretation* covered practical interpretation of estimates and uncertainty, confounding, attributing causality, and the role of the research design to the validity of conclusions drawn. *Reporting results* was practical advice about presentation and reporting.

There was a decreasing need for covering these themes with increasing client competence (Figure 1). The odds of covering *Understanding inference* were 2.5 times for basic compared with competent (Table 2). Over 40% of project consultations for even competent clients covered *Understanding inference*. Though the differences were a little smaller, in a similar way, 36% of consultations for competent clients covered *Applied interpretation.*

CONCLUSION

For many of the 13 broad themes, there was little variation or decreasing coverage with increasing client competence. In a general sense, decreasing coverage with increasing competence

might be expected – clients with basic competence might need more advice across all aspects of the PPDAC cycle.

Some broad themes at the heart of what it means to think statistically were covered less often for competent compared with basic clients. However they were still covered relatively often for the competent clients. These themes were *Fundamentals* covering basic ideas about samples, population, estimation, inference and variation (31% for competent clients), *Graphs* (purpose, practice and principles of good graphics (50% for competent), *Understanding inference* (41% for competent), and *Applied interpretation* (36% for competent). None of these percentages is small. We conclude that it is unwise to assume that a client judged to be generally competent has a strong understanding of all basic statistical ideas; our data suggest that this may not be the case.

Wild (2017) stated that "The greatest hope for extending the accessibility of the statistical messages in data to a much wider spectrum of the population is visualization. But no matter where you are on the continuum between the webpage viewer and the professional data analyst, you have to be able to make sense of what you are seeing" (p.33). In this context, the need for discussion of *Graphs* in more than half all consultations identifies an important emphasis for statistical training.

The broad theme of *Units of observation* was discussed in 22% of all consultations. Coverage of this theme arose in a majority of cases from discussing the nature of units (rather than replication). This is another fundamental framing statistical idea, and our finding suggests identifying units may be potentially challenging regardless of general competence.

Modelling methods were covered for a high percentage of project consultations for all client groups. The odds of covering *Simple inference* was higher for basic competence than for middling and competent levels, whereas the odds of covering *Linear mixed models* was lower for basic compared with competent clients. There was a tendency for increasing client competence to be related to increasing complexity of the nature of modelling discussed. For *Modelling issues*, however, coverage was around 70% for competent clients – 30% higher than for basic. This broad theme included conceptually important ideas about modelling: how to identify the relevant structure and levels of variation, fitting a model to data, assumptions, transformations, outliers, fixed and random effects, one versus two tailed tests, and avoiding fishing for results. *Modelling issues* includes items that are relevant to most methods of modelling, as well as some that apply to more complex modelling. While a competent client might need and be able to produce more complex analyses, we cannot assume the foundational ideas underlying statistical modelling are strong.

George Cobb (2015) argued that the imperative in developing statistical curricula is to *teach through research*. Our study identifies needs for those trying to *learn through research*. Our characterization shows that the knowledge and skills of our (variously competent) learners cannot be assumed to be built on a scaffold with strong foundations. Building that foundation for future learners, potentially able to learn in many ways other than through a traditional undergraduate curriculum, is the challenge.

REFERENCES

Cobb, G. (2015). Mere renovation is too little too late: We need to rethink our undergraduate curriculum from the ground up. *The American Statistician, 69*(4), 266─282.

Wild, C.J. (2017). Statistical literacy as the earth moves. *Statistical Education Research Journal, 16*(1), 31─37.

Wild, C J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223─248.

Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In *International Handbook of Research in Statistics Education* (pp. 5─36). Springer, Cham.