Watching The Detectives: An Initial Report On An Industrial Experiment To Collaborate With The Empirical Software Engineering Research Community[*]

Ken Wallace[†]
BAE Systems, Maritime
Portsmouth, UK

ABSTRACT

Context: Evidence-based practice within healthcare benefits from the Cochrane Collaboration providing trustworthy evidence that can be used to inform clinical decisions. Currently no equivalent resource exists for the practice of software engineering. With the ever increasing dependence of society upon software an analogue of the collaboration for the discipline of software engineering would appear to be of potentially significant societal benefit.

Objectives: This study investigated attitudes of the empirical software engineering community, as exemplified by attendees at an research focused conference, to the hypothesis that an analogue of the Cochrane Collaboration would be beneficial.

Method: An online survey was employed to capture opinions of participants in the EASE'16 conference on the hypothesis. During a presentation at the industrial track of the conference attendees (27) were invited to complete the survey. Subsequently all conference attendees (90) were provided with details of the survey, irrespective of whether they had attended the presentation or not. No closing date for the survey was specified.

Results: In total 9 conference attendees completed the survey: of whom 7 had attended the industrial track presentation. These figures approximate to a response rate of 26% for those attending the presentation and exactly 10% for conference participants overall. This response rate is insufficient to enable statistical analysis of the data obtained through the survey. Trends are, however, evident in the rankings of the responses provided. These trends are consistent with opinions and observations volunteered by a diverse range of individuals active in software engineering and allied fields dating from 2014 to the present.

Conclusions: At this time there is insufficient data to validate the hypothesis. Nevertheless experience, informal observation and opinion continues to accumulate to provide evidence in support of the validity of the hypothesis, albeit principally in anecdotal form.

[*]The identified author is solely responsible for the content herein: attribution to BAE Systems is not permitted.

[†]**ORCID iD** https://orcid.org/0000-0001-7232-0511.

CCS CONCEPTS

• **General and reference** → **Empirical studies**; **Experimentation**; • **Social and professional topics** → **Socio-technical systems**;

KEYWORDS

Software Engineering; Evidence; Cochrane Collaboration; Survey, Principles.

1 INTRODUCTION

In making their original proposal for Evidence-Based Software Engineering (EBSE) in 2004, Kitchenham, Dybå and Jørgensen [11] were explicit in identifying the importance of the Cochrane Collaboration (www.cochrane.org) to the conduct of evidence-based practice in medicine. In doing so they also acknowledged barriers to an equivalent for EBSE, highlighting the international collaboration and funding that would be required to enable the establishment of a comparable network of researchers, professionals, users and other stakeholders with an interest in software engineering. Fast forward fourteen years and the need for a Cochrane analogue for the discipline of software engineering remains, as yet, unfulfilled.

Elsewhere the author has argued [16], as have others [9] that increasing societal dependency on software on a global scale incurs consequential risks should software be unable to achieve a level of dependability commensurate with the trust increasingly being placed in it. Paradoxically such dependence comes at a time when software is disappearing from view [17]. This paradox reflects not only the now well established role of software as the 'building material of choice' [13] for complex systems, but also the extent to which society has reached a utility-based accommodation of software contingent, arguably, upon a belief that it will 'just work'. In a world of faster, better, cheaper few take the time to pause and reflect on whether such a belief is either justified or wise.

With software increasingly assuming the role of 'lightning conductor' for complexity in the rapidly evolving digital world the need for effective independent and rigorous scrutiny of software engineering has never been more acute. This is particularly so

in situations with the potential to cause harm or give rise to significant adverse consequences, regardless of whether economic, environmental or otherwise.

Given the success of the Cochrane Collaboration in healthcare the intent of the present study was to investigate the hypothesis that there are potential benefits of an analogue for the discipline of software engineering.

In particular this paper provides an initial report on the results of an experiment to assess attitudes of the empirical software engineering (ESE) research community to this hypothesis. The venue chosen for the experiment was the 20th International Evaluation and Assessment in Software Engineering Conference (EASE'16) held in Limerick, Ireland in 2016. This conference series is *"one of the premier venues for academics and practitioners to present and discuss their empirical software engineering research and their application in practice"* and has a particular focus upon evidence-based approaches [2].

2 THE EXPERIMENT

Attendance at conferences has long constituted an important fixture of academia: for both seasoned campaigners and those at earlier stages of their career looking to establish themselves in their chosen area of study. As well as seeking to advance a given field of knowledge, proponents and organisers of conferences would argue that they should and can provide a means for participants to engage in face to face dialogue sharing views, opinions and knowledge on topics of mutual interest. Such outcomes are beneficial in breaking down some of the barriers to progress that can arise when individuals working on similar topics are not in direct contact with each other on a daily basis.

For practitioners attendance at academic, rather than more practically focused conferences can present difficulties, particularly if the conference is geographically distant. Businesses rightly wish to understand how employee attendance can contribute to the bottom line, especially if the theme of the conference is diffuse, the relevance is not immediately obvious, or the timeliness with which the knowledge acquired can be applied unquantified. Moreover as the number of conferences has grown, concerns amongst practitioners about increasing repetition of familiar themes have been accompanied by rising dissatisfaction regarding the proportion of content which is novel, non-obvious and actionable from the perspective of practitioners. These concerns reflect and are symptomatic of the wider challenges facing academic software engineering research as it seeks to engage with, and demonstrate relevance to industrial practice [3, 4] and answer the familiar 'so what' question.

Recognising that information exchange is a 'two-way street', as a practitioner the author has engaged with the ESE research community through participation in various fora such as the Industrial Track introduced to the EASE conference series in 2014 [1, 15]. A motivation for the experiment reported in this paper was, therefore, to evaluate the benefits of these engagements using responses to the hypothesis provided by EASE'16 attendees as a measure. The method of assessment chosen was a survey, the structure of which is shown in Table 1.

To support the survey several application categories either contributing to and/or potentially benefiting from advancement of

EBSE were employed. The categories, denoted (i) - (v) in Table 2, were defined specifically for the survey: no existing well-recognised construct that could be used for this purpose was identified. These categories were used as a basis for constructing a series of 16 questions addressing aspects of the hypothesis that were compiled into an online survey in advance of the EASE'16 conference.

Communication of the survey was through presentation by the author of experience gained in applying EBSE to industrial practice. Building upon the learning reported in the peer-reviewed Extended Abstract in the conference proceedings the presentation:

(1): considered the role of EBSE in the context of a '5S' model of evidence-based information comprising (empirical) studies, syntheses (systematic reviews), synopses, standards and systems

(2): examined the current state of knowledge in respect of automated and manual static analysis of code and the contribution of EBSE to this particular topic as an example outcome

(3): examined the EBSE User Experience (UX) and compared with other approaches to communicating evidence-based information to decisions makers and practitioners [14]

(4): addressed the question of whether the lessons in respect of industrial adoption, or more specifically the lack thereof, of formal methods over a number of decades could be applied to ensure EBSE avoids the 'Practice Graveyard'.

Further the presentation identified the vision, mission and principles of the Cochrane Collaboration, elaborating the latter and posing the question of whether EBSE would pass a 'Cochrane Test'. For reference the Cochrane Principles are:

(a): **Collaboration** by fostering global co-operation, teamwork, and open and transparent communication and decision making

(b): Building on the **Enthusiasm** of individuals by involving, supporting and training people of different skills and backgrounds.

(c): Avoiding **Duplication** of effort by good management, coordination and effective internal communications to maximize economy of effort.

(d): Minimizing **Bias** through a variety of approaches such as scientific rigour, ensuring broad participation, and avoiding conflicts of interest.

(e): Keeping up-to-date by a commitment to ensure that Cochrane Systematic Reviews are maintained through identification and incorporation of new evidence. (**Currency**)

(f): Striving for **Relevance** by promoting the assessment of health questions using outcomes that matter to people making choices in health and health care.

(g): Promoting **Access** by wide dissemination of our outputs, taking advantage of strategic alliances, and by promoting appropriate access models and delivery solutions to meet the needs of users worldwide.

(h): Ensuring **Quality** by applying advances in methodology, developing systems for quality improvement, and being open and responsive to criticism.

(i): **Continuity** by ensuring that responsibility for reviews, editorial processes, and key functions is maintained and renewed.

(j): Enabling wide **Participation** in our work by reducing barriers to contributing and by encouraging diversity.

The audience for the presentation was then invited to provide evidence-based assessments of the Cochrane Test either through completion of the online survey or alternatively via email contact with the author.

Table 1: Survey Structure

Part	Description	Number
1	Respondent experience and knowledge with reference to the application categories: (i) - (v) in Table 2	5
2	Opinion on EBSE and the Cochrane principles (labeled in this report as A - G)	7
3	Follow up questions linked to Parts 1 & 2	4

Table 2: Application Categories and Response Rankings for Questions A - D

Category	Description	A	B	C	D
(i)	Empirical software engineering **research**	1	1	=2	=2
(ii)	Information engineering, science and technology (IT) **research**	2	2	4	4
(iii)	Computer science **research**	3	3	5	5
(iv)	Software engineering **practice** in the commercial, industrial or public sectors	4	5	1	1
(v)	IT **practice**, including information engineering, in the commercial, industrial or public sectors	5	4	=2	=2

In the present report only the responses to Part 2 questions are considered. The 7 questions in this part, identified in this paper as A - G, invited respondents to:

A: **Apply** a 5 point Likert scale to rate the extent to which EBSE has succeeded in improving decision making through the provision of best evidence in each of the application categories

B: **Apply** a 5 point Likert scale to rate the extent to which EBSE has succeeded in improving decision making through the integration of practical experience and human values with evidence in each of the application categories [7]

C: **Rank** by application category the personal importance to the respondent that EBSE achieve impact

D: **Rank** by application category the importance to society that EBSE achieve impact

E: **Apply** a 5 point Likert scale to rate the importance of the Cochrane principles to enable EBSE to achieve impact on a societal scale

F: **Rank** the Cochrane principles in order of importance to increase societal impact of EBSE

G: **Apply** a 5 point Likert scale to rate the extent of barriers that prevent each of the Cochrane principles being achieved for EBSE.

Respondents were advised that the development, delivery and management of educational services, for example teaching, but not associated research, should be accounted for as IT practice in the public sector.

For questions A and B respondents were also asked to provide a textual commentary to explain their chosen ratings.

Approximately a week after the conference ended all attendees (90) were provided with details of the survey, irrespective of whether they had attended the presentation or not. This was achieved through an email circulated by the conference organisers. This email explained the purpose of the survey for those who had not attended the industrial track presentation. No closing date for the survey was specified, although effective February 2017 the survey closed.

3 RESULTS

Intentionally the information provided in this initial report is limited in order to minimise any influence on or bias of future responses that are sought to achieve statistically significant results. Consequently, as noted, only the ranking for the responses to the Part 2 questions are considered. The qualitative nature of questions in Part 3 limits the scope for presentation of an equivalent analysis for this part without incurring a significant risk of influence or biasing future responses.

Table 3: Response Rankings for Questions E-G

Rank	E	F	G
1	Quality	=Quality	Participation
2	Collaboration	=Collaboration	Relevance
3	Bias	Bias	Access
4	Access	Relevance	Bias
5	Participation	Access	Currency
6	=Currency	Currency	Quality
7	=Relevance	Duplication	Collaboration
8	Continuity	Participation	Continuity
9	Duplication	Continuity	Duplication
10	Enthusiasm	Enthusiasm	Enthusiasm

Rankings for Questions A - D are given in Table 2 and those for Questions E - G in Table 3. For both tables 1 is the highest rank and ties in rank are indicated by =.

Response rates for the survey equate to approximately 26% for those attending the presentation and exactly 10% overall for conference participants. These rates are comparable both with those indicated by Kitchenham et al. [10] and experience in industrial practice. All responses were obtained through the online survey. The majority of respondents (66%) had read the accompanying abstract [1] at the time of completion of the survey.

One potential measure of comparison for the foregoing rates is the number of downloads of the extended abstract as registered by the source repository for the conference proceedings (ACM Digital Library). As of December 2017 a total of 56 downloads had been recorded. This figure compares with a mean of 84 for all published conference papers (N=35), placing the noted abstract in the second quartile (47.5 - 69). For reference in the same period there were 47 downloads for the extended abstract of the other industrial presentation included in the published proceedings.

Seven of the nine respondents elected to receive and were provided with preliminary results comparable to those included in this report following closure of the survey.

4 THREATS TO VALIDITY

It has been assumed that attendees at EASE'16 are representative of the ESE research community generally. Noting that the call for papers for the conference invites "*the submission of papers describing original and unpublished results of empirical research in all areas of Software Engineering*" this assumption appears reasonable. As a corollary the author is unaware of any evidence to suggest that

this assumption is either erroneous or inaccurate. This assessment takes into account participation by the author in the Programme Committee for both the EASE'17 and 18 conferences.

5 DISCUSSION

It is evident further responses to the survey will be necessary to enable statistically significant analysis of the results achieved. Increasing **Collaboration** with the ESE research community should enable such an outcome.

Given the initial nature of this report it is not possible to draw any definitive conclusions from the results reported at this time. Nevertheless one observation that can be made is that, based upon experience to date, neither increased **Access** nor **Participation** necessarily translate into increased **Collaboration** to an extent that might be expected or hoped for. Indeed the fidelity with which the response to the survey in the current experiment reproduces that typical of surveys of a general population, rather than a professional one that is incentivised (in theory) to achieve improvement is worthy of note.

For the ESE research community the results of this initial experiment might appear paradoxical given the existential concerns about **Relevance**, the lack of **Participation** from industry and the extent of the divide between research and practice articulated by ESE researchers [3–6, 8, 12]. Hopefully the initial results presented in this report may be of some use to the community when considering these issues.

Learning from experience gained conducting the experiment it is recommended that in respect of **Access**:

1: conference websites be made permanent accessible electronic records of the material **presented**

2: that source material held on the website be made freely available through the use of an appropriate public domain licence.

These recommendations are intended to augment, rather than replace, the conventional (e)print format of conference proceedings and thereby potentially increase the **Relevance** of information and knowledge embodied in the conference.

In terms of future work an expanded survey[1] is now available. There will also be an opportunity for participants in this survey to collaborate on further work regarding the development of a Cochrane analogue. The responses to this survey will hopefully be sufficient to achieve a level of statistical significance that can subsequently be shared with the ESE community. Unsurprisingly therefore **Participation** in this survey is strongly encouraged.

6 CONCLUSIONS

Examining the Cochrane principles it is suggested that the introduction of the Industrial Track to the EASE conference is an example of the Cochrane principles of **Access** and **Participation** in action. Use of the Cochrane principles as a framework for the experiment has also proved beneficial in reducing **Duplication**.

The case for effective independent scrutiny of claims made in respect of the discipline of software engineering that can be tested rigorously and objectively using proven scientific methods is compelling: the ESE research community has an important role to play in this regard. It is to be hoped that this ambition is shared by the

community and that collectively it is willing to participate enthusiastically in a collaborative manner to achieve this outcome. A significant response from attendees at the SER&IP'18 workshop and the wider ICSE'18 community to the follow-on survey would be a very welcome indicator to this effect.

REFERENCES

[1] Steve Anderson, Paul Sagar, Beth Smith, and Ken Wallace. 2016. What We Have Learnt Adopting Evidence-based Software Engineering for Industrial Practice. In *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering (EASE '16)*. ACM, New York, NY, USA, Article 30, 1 pages. https://doi.org/10.1145/2915970.2915981

[2] Sarah Beecham, David Bowes, and Klaas-Jan Stol. 2017. Introduction to the EASE 2016 special section: Evidence-based software engineering: Past, present, and future. *Information and Software Technology* 89 (2017), 14 – 18. https://doi.org/10.1016/j.infsof.2017.05.002

[3] S. Beecham, P. O'Leary, S. Baker, I Richardson, and J. Noll. 2014. Making Software Engineering Research Relevant. *Computer* 47, 4 (Apr 2014), 80–83.

[4] L. Briand. 2012. Embracing the Engineering Side of Software Engineering. *Software, IEEE* 29, 4 (July 2012), 96. https://doi.org/10.1109/MS.2012.86

[5] Jeffrey C. Carver, Oscar Dieste, Nicholas A. Kraft, David Lo, and Thomas Zimmermann. 2016. How Practitioners Perceive the Relevance of ESEM Research. In *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '16)*. ACM, New York, NY, USA, Article 56, 10 pages. https://doi.org/10.1145/2961111.2962597

[6] J. Cleland-Huang. 2015. Toward Meaningful Industrial–Academic Partnerships. *IEEE Software* 32, 1 (Jan 2015), 18–21. https://doi.org/10.1109/MS.2015.20

[7] T. Dyba, B.A Kitchenham, and M. Jorgensen. 2005. Evidence-based software engineering for practitioners. *Software, IEEE* 22, 1 (Jan 2005), 58–65. https://doi.org/10.1109/MS.2005.6

[8] Vahid Garousi, Kai Petersen, and Baris Ozkan. 2016. Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review. *Information and Software Technology* 79 (2016), 106 – 127. https://doi.org/10.1016/j.infsof.2016.07.006

[9] John Hatcliff, Alan Wassyng, Tim Kelly, Cyrille Comar, and Paul Jones. 2014. Certifiably Safe Software-dependent Systems: Challenges and Directions. In *Proceedings of the Future of Software Engineering (FOSE 2014)*. ACM, New York, NY, USA, 182–200. https://doi.org/10.1145/2593882.2593895

[10] Barbara Ann Kitchenham, David Budgen, and Pearl Brereton. 2016. *Evidence-Based Software Engineering and Systematic Reviews*. CRC Press.

[11] Barbara A. Kitchenham, Tore Dybå, and Magne Jørgensen. 2004. Evidence-Based Software Engineering. In *Proceedings of the 26th International Conference on Software Engineering (ICSE '04)*. IEEE Computer Society, Washington, DC, USA, 273–281. http://dl.acm.org/citation.cfm?id=998675.999432

[12] David Lo, Nachiappan Nagappan, and Thomas Zimmermann. 2015. How practitioners perceive the relevance of software engineering research. In *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*. ACM, 415–425.

[13] Art Pyster, Rick Adcock, Mark Ardis, Rob Cloutier, Devanandham Henry, Linda Laird, Harold Lawson, Michael Pennotti, Kevin Sullivan, and Jon Wade. 2015. Exploring the Relationship between Systems Engineering and Software Engineering. *Procedia Computer Science* 44 (2015), 708 – 717. https://doi.org/10.1016/j.procs.2015.03.016 2015 Conference on Systems Engineering Research.

[14] (UK) Government - guidance. 2015. What Works Network. https://www.gov.uk/guidance/what-works-network. (January 2015).

[15] K. R. Wallace. 2014. I'm A Practitioner Help Me Out Here: Can Evidence Based Software Engineering Really Help Navigate The Iron Triangle Successfully ?. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE '14)*. ACM, New York, NY, USA, 453.

[16] K. R. Wallace. 2016. Safe Software: A Suitable Case For An Evidence-Based Treatment?. In *ESREL 2016*. Taylor & Francis, 2403 – 2410.

[17] G. P. Zachary. 2016. The invisible technology. *IEEE Spectrum* 53, 11 (November 2016), 10. https://doi.org/10.1109/MSPEC.2016.7607011

---

[1]https://www.surveymonkey.com/r/CochraneTest.