

From research to research synthesis in CALL

Alex Boulton¹

Abstract. Any research study can only be fully appreciated once it is situated in relation to existing work. This is no mean feat, however, given the sheer quantity and variety of publications to date. Simply relying on one’s background and experience as an expert in the field, coupled with a few internet searches and following up individual references, is likely to lead to a very partial view. This paper argues the need for greater rigour (via meta-analytic and other types of syntheses) to gain a broader, deeper and more balanced understanding of Computer-Assisted Language Learning (CALL).

Keywords: data-driven learning, corpora, research synthesis, meta-analysis.

1. Conducting research

The ‘scientific method’ has been developed in an attempt to reduce human fallibility in exploring the world around us. However, different researchers clearly go about their work in vastly disparate ways even within a single clearly-defined discipline such as CALL. Many attempts have been made to describe the different practices, one of the most common distinctions being between quantitative and qualitative research. A number of surveys have found the former to be prevalent in international journals in applied linguistics (e.g. Richards, 2009), which may fuel a popular perception that it is more prestigious or even more ‘scientific’ in some way.

However, there is disagreement about exactly what qualitative and quantitative methods are, and debate about whether there is a clear boundary between them. On the face of it, any set of data is open to some sort of quantification – and indeed needs to be, otherwise it is impossible to know what to make of discussions of a single example, blog extract, or interview response. Is it representative of a

1. Atilf, CNRS & Université de Lorraine, France; alex.boulton@univ-lorraine.fr

How to cite this article: Boulton, A. (2015). From research to research synthesis in CALL. In F. Helm, L. Bradley, M. Guarda, & S. Thouëсны (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 84-90). Dublin: Research-publishing.net. <http://dx.doi.org/10.14705/rpnet.2015.000314>

more widespread phenomenon, or just an interesting but isolated case? Any data which can be counted but is not (or which stops at the level of raw numbers or percentages), properly invites scepticism from the reader. In the end, the take-home message of many qualitative papers is that the situation is complex (for which the sceptic may read ‘vague’ and ‘subjective’), but that the researchers have found at least some evidence pointing in the right direction (if they are to be believed). Similarly, any overtly quantitative data also needs interpretation for it to make any meaningful contribution. As with qualitative research, it is surprisingly easy to go through the motions and produce poor quantitative studies by simply grinding numbers through an esoteric statistical test chosen for mysterious reasons, leading to a ‘voilà’ moment of $p < .05$. This is markedly unsatisfactory, and has the opposite defect of qualitative studies in being misleadingly simplistic. While each approach is thus easy to criticise on scientific grounds or personal/cultural preference, it seems likely that the most robust research will derive from truly mixed-methods designs.

2. Reviewing research

The scientific enterprise is incremental and no single study will definitively answer any given issue. The question then is how to gain an accurate overview of research to date where even a small field like CALL sees many hundreds of studies published every year, often with conflicting results. The sheer number of publications means it is always possible to find some evidence that justify almost anything (Hattie, 2009, p. 6); it is therefore essential to find ways to bring greater rigour to research synthesis. Considerable advances have been made in this direction since the publication of the seminal paper by Norris and Ortega in 2000. Today, research synthesis has become almost a field in its own right, with a number of handbooks, recommendations by academic associations and scientific journals, and special issues of prestigious journals or collected volumes. Norris and Ortega’s (2010) TimeLine review in *Language Teaching* gives a glimpse of the wealth of work in the area.

Most research synthesis begins with an extensive and principled trawl of the literature related to a clearly defined question, but then can branch in different directions, each with its advantages and disadvantages (see Plonsky, 2014 for an overview). The *narrative synthesis* represents a qualitative approach: it can incorporate any type of study and allows for interpretation and contextualisation by the synthesist, but thereby remains open to charges of subjectivity; and while the picture is carefully nuanced, the final impression may remain correspondingly vague and fuzzy (Han, 2015). Quantitative approaches, on the other hand, attempt

to be more objective in their interpretation of the results, but by definition only cater for studies that provide appropriate quantitative data; like primary quantitative studies, meta-analyses tend to leave a single numerical value as the take-home message for the casual reader, which is simplistic and misleading, and does not do justice to the sub-analyses of moderator variables (for other types of synthesis in CALL, see [Burston 2013, 2015](#); [Felix 2005, 2008](#)).

3. Meta-analyses in CALL

The principle phases consist in outlining the scope of the topic, collecting and selecting publications, coding and extracting the data for analysis, calculating effect sizes and interpreting them according to various moderator variables – all according to stringent, pre-determined criteria. Though many decisions need to be made, the main constant in most meta-analysis is the calculation of the effect size; most in applied linguistics use Cohen's *d*. This basically compares the difference in means between the control and experimental groups (or pre- and post-tests), while taking into account the variance as given in the pooled standard deviations ([Figure 1](#)). Effect size is in many ways more revealing than the more common significance testing; it is recommended if not required by recent APA standards and journals such as *Language Learning*, while some researchers seem to think that *p*-values are at best uninformative and at worst positively harmful, and should be systematically replaced by effect sizes (e.g. [Plonsky, 2011](#)). One major advantage of using a standard measure of effect size is that it enables direct comparison of different studies, which is not possible with *p*-values or narrative syntheses.

Figure 1. Formula for Cohen's *d*

$$d = \frac{M_2 - M_1}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$$

As a field, CALL is now sufficiently mature to have given rise to several meta-analyses, some of which are given in [Table 1](#); *k* refers to the number of studies covered in the analysis, and *d* is the effect size itself². The value for *d* needs interpreting (just as do *p*-values, which tend to be set arbitrarily at .05 or .01). For applied linguistics, [Oswald and Plonsky \(2010\)](#) find an average effect size of 0.7,

2. Pre/post (within-groups) and control/experimental (between groups) designs are not distinguished in this short paper.

and suggest that 0.4 should be considered small, 1.0 large. The first thing to note from [Table 1](#) is that there are no negative *d*-values in any of the meta-analyses. This is not surprising, since few primary studies set out to discredit an experimental treatment against a control group, and would not expect lower scores in a post-test following treatment. Second, most of the effect sizes are not particularly large, the unweighted mean being just 0.66, with the higher ones mainly derived from smaller samples. Overall, this suggests a medium strength effect of computer-assisted language learning as seen from many different perspectives over many dozens of primary studies involving thousands of learners using a wide variety of tools and techniques. Third, each arrives at a different value; a single meta-analysis does not provide a definitive picture of a field (compare [Plonsky and Brown's \(2015\)](#) discussion of the differing results of 18 meta-analyses of feedback).

Table 1. Meta-analyses in CALL (partly based on [Oswald & Plonsky, 2010](#))

study	year	source	question	<i>k</i>	<i>d</i>
Abraham	2008	<i>Computer Assisted Language Learning</i>	computer-mediated glosses in vocabulary learning	6	1.40
Abraham	2008	<i>Computer Assisted Language Learning</i>	computer-mediated glosses in reading comprehension	11	0.73
Chiu	2013	<i>British Journal of Educational Technology</i>	computer-assisted second language vocabulary instruction	16	0.75
Chiu et al.	2012	<i>British Journal of Educational Technology</i>	digital game-based learning	14	0.53
Cobb & Boulton	2015	Cambridge University Press	data-driven learning	21	1.04
Grgurović et al.	2013	<i>ReCALL</i>	CALL-based language learning	65	0.26
Lin, H.	2014	<i>Language Learning & Technology</i>	CMC and SLA	59	0.44
Lin, H.	2015	<i>ReCALL</i>	CMC in L2 oral proficiency development	25	0.40
Lin, W.C. et al.	2013	<i>Language Learning & Technology</i>	text-based SCMC on SLA	19	0.33
Taylor	2009	<i>CALICO Journal</i>	CALL-based versus paper-based glosses	32	0.49
Yun	2011	<i>Computer Assisted Language Learning</i>	hypertext glosses in vocabulary acquisition	10	0.37
Zhao	2003	<i>CALICO Journal</i>	overall effectiveness of uses of technology in language education	9	1.12

Though as [Grgurović, Chapelle, and Shelley \(2013\)](#) point out, it can be politically useful to be able to quantify the effects of CALL, attempting to account for all the research in a single figure is obviously hugely simplistic. In quantitative research, a major failing in primary studies is that the variation (even if reported in standard deviations) is ironed out in a single overall figure; by definition, a

meta-analysis involves far more variation which is also ignored if we only take away a single figure. Fortunately, meta-analysts do not just provide a single overall figure for effect size, they also discuss and interpret it, and in particular conduct analysis of potential moderator variables precisely to see what factors may explain the variation between studies. While it is not possible to go into details here, the reader is strongly encouraged *not* to take away the simple notion that $d=.66$ for CALL (unless it is politically or strategically expedient to justify budgets or investment at a local level, unethical though that may be), but to consult the various meta-analyses listed to see how each explains the variation it uncovers among the primary studies, and to decide whether the variation between the meta-analyses themselves may be attributed to their specific design or research questions. It is also of course important to go to the relevant primary studies, but approaching them after consulting a meta-analysis may help to keep them in perspective.

4. Conclusions

Meta-analysis can be “an immensely valuable scholarly contribution that brings order to confusion, helps set a future research agenda, and at the same time gives the best evidence-based practical advice” (Cumming, 2012, p. 231), but has its limitations and should never be taken as the ultimate answer to a question. What is needed is always more research: more primary studies of different types (where syntheses can help identify areas in need of work), and more syntheses to help make sense of them – again, both qualitative and quantitative.

References

- Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: a meta-analysis. *Computer Assisted Language Learning*, 21(3), 199-226. doi:10.1080/09588220802090246
- Burston, J. (2013). Mobile-assisted language learning: a selected annotated bibliography of implementation studies 1994-2012. *Language Learning & Technology*, 17(3), 157-225.
- Burston, J. (2015). Twenty years of MALL project implementation: a meta-analysis of learning outcomes. *ReCALL*, 27(1), 4-20. doi:10.1017/S0958344014000159
- Chiu, Y.-H. (2013). Computer-assisted second language vocabulary instruction: a meta-analysis. *British Journal of Educational Technology*, 44(2), E52-E56. doi:10.1111/j.1467-8535.2012.01342.x
- Chiu, Y.-H., Kao, C.-W., & Reynolds, B. L. (2012). The relative effectiveness of digital game-based learning types in English as a foreign language setting: a meta-analysis. *British Journal of Educational Technology*, 43(4), E104-E107. doi:10.1111/j.1467-8535.2012.01295.x

- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *Cambridge handbook of corpus linguistics* (pp. 478-497). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139764377.027
- Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.
- Felix, U. (2005). What do meta-analyses tell us about CALL effectiveness? *ReCALL*, 17(2), 269-288. doi:10.1017/S0958344005000923
- Felix, U. (2008). The unreasonable effectiveness of CALL: what have we learned in two decades of research? *ReCALL*, 20(2), 141-161. doi:10.1017/S0958344008000323
- Grgurović, M., Chappelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology supported language learning. *ReCALL*, 25(2), 165-198. doi:10.1017/S0958344013000013
- Han, Z. (2015). Striving for complementarity between narrative and meta-analytic reviews. *Applied Linguistics*, 36(3), 409-415. doi:10.1093/applin/amv026
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Lin, H. (2014). Establishing an empirical link between computer-mediated communication (CMC) and SLA: a meta-analysis of the research. *Language Learning & Technology*, 18(3), 120-147.
- Lin, H. (2015). Computer-mediated communication (CMC) in L2 oral proficiency development: a meta-analysis. *ReCALL*, 27(3), 261-287. doi:10.1017/S095834401400041X
- Lin, W. C., Huang, H. T., & Liou, H. C. (2013). The effects of text-based SCMC on SLA: a meta-analysis. *Language Learning & Technology*, 17(2), 123-142.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528. doi:10.1111/0023-8333.00136
- Norris, J. M., & Ortega, L. (2010). Research synthesis [Research timeline article.] *Language Teaching*, 43(4), 461-479. doi:10.1017/S0261444810000200
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: choices and challenges. *Annual Review of Applied Linguistics*, 30, 85-110. doi:10.1017/S0267190510000115
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: a meta-analysis. *Language Learning*, 61(4), 993-1038. doi:10.1111/j.1467-9922.2011.00663.x
- Plonsky, L. (2014). Study quality in quantitative L2 research (1990–2010): a methodological synthesis and call for reform. *Modern Language Journal*, 98(1), 450-470. doi:10.1111/j.1540-4781.2014.12058.x
- Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31(2), 267-278. doi:10.1177/0267658314536436
- Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching*, 42(2), 147-180. doi:10.1017/S0261444808005612

- Taylor, A. M. (2009). CALL-based versus paper-based glosses: is there a difference in reading comprehension? *CALICO Journal*, 27(1), 147-160. doi:10.11139/cj.27.1.147-160
- Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: a meta-analysis. *Computer Assisted Language Learning*, 24(1), 39-58. doi:10.1080/09588221.2010.523285
- Zhao, Y. (2003). Recent developments in technology and language learning: a literature review and meta-analysis. *CALICO Journal*, 21(1), 7-27.

Published by Research-publishing.net, not-for-profit association
Dublin, Ireland; info@research-publishing.net

© 2015 by Research-publishing.net (collective work)
© 2015 by Author (individual work)

Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy
Edited by Francesca Helm, Linda Bradley, Marta Guarda, and Sylvie Thouéšny

Rights: All articles in this collection are published under the Attribution-NonCommercial -NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Under this licence, the contents are freely available online (as PDF files) for anybody to read, download, copy, and redistribute provided that the author(s), editorial team, and publisher are properly cited. Commercial use and derivative works are, however, not permitted.



Disclaimer: Research-publishing.net does not take any responsibility for the content of the pages written by the authors of this book. The authors have recognised that the work described was not published before, or that it is not under consideration for publication elsewhere. While the information in this book are believed to be true and accurate on the date of its going to press, neither the editorial team, nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, expressed or implied, with respect to the material contained herein. While Research-publishing.net is committed to publishing works of integrity, the words are the authors' alone.

Trademark notice: product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Copyrighted material: every effort has been made by the editorial team to trace copyright holders and to obtain their permission for the use of copyrighted material in this book. In the event of errors or omissions, please notify the publisher of any corrections that will need to be incorporated in future editions of this book.

Typeset by Research-publishing.net
Fonts used are licensed under a SIL Open Font License

ISBN13: 978-1-908416-28-5 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-29-2 (Ebook, PDF, colour)
ISBN13: 978-1-908416-30-8 (Ebook, EPUB, colour)

Legal deposit, Ireland: The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.

Legal deposit, United Kingdom: The British Library.
British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2015.