



## Directions in study abroad motivational research

### 留学と言語学習の動機における研究とその傾向

Tim Newfields (Toyo University)

本論では、言語学習のモチベーションと留学との動機に関する研究について、モデルとなるような論文と、問題のある論文について比較して述べる。本論の目的は、第一に海外留学の動機研究に関する方法論の良い点と問題点との比較。第二に言語学習へのモチベーションが留学にどのように影響するか、その有無について述べている。最後に留学と言語学習の動機付けの研究の方向性を示唆する。

As the volume of research on study abroad (SA) expands, the task of evaluating the growing number of studies becomes ever more arduous. A wide variety of methods are employed to see how SA might impact foreign language learning motivation. This paper highlights some recent SA research in English on EFL motivation. Three studies that I consider exemplary are contrasted with three that are problematic in some respects. The rationale for this is twofold. One goal is to critically assess some of the methodologies employed in SA research; another is to examine some divergent claims made about SA on language learning motivation.

According to Richards (2009, p. 148) the bulk of current social science research is informed by several competing philosophical frameworks. One framework is positivistic and purports that objective reality can be quantified and measured. Strong claims are often made about precision, universality, and objectivity. More recently, post-positivists have hedged such claims while endorsing the same stringent methodological procedures. Most quantitative research is informed by positivist or post-positivist frameworks.

Another competing research paradigm comes from constructivism. Constructivists contend that knowledge is locally situated and socially co-constructed. This paradigm is suited to a wide range of qualitative research procedures and its proponents tend to renounce assertions of “universal knowledge”. Piaget (1964), von Glasersfeld (1987), and Papert (1994) are but a few of the leading constructivists.

Still another framework informing is critical theory. Broadly speaking, this attempts to examine the social and historical forces shaping the accounts of many phenomena. Critical theorists point out that researchers influence observed outcomes and that subjectivity is inherent in all inquiry. Quantitative research is, from this standpoint, but one of many possible ways of



obtaining information and there are limits on how widely the findings from such research can generalize to other contexts. Leading critical researchers include Schroyer (1973), Chomsky (2000), and Kincheloe (2002, 2005). Critical text analysis and narratology are two widely used procedures in this tradition.

In light of these research frameworks, let us now examine some SA papers published over the last decade to see each how the theoretical framework behind these papers is (or possibly is not) congruent with their research methodologies.

### **Three Exemplary SA Motivational Research Studies**

#### **(1) Chirkov, Vansteeniste, Tao, and Lynch (2007)**

The most sophisticated quantitative SA research in recent years on motivation is probably Chirkov, Vansteeniste, Tao, and Lynch's comparison of two groups of Chinese students in Belgium ( $n=122$ ) and Canada ( $n=98$ ). They used a 21-item questionnaire that highlighted different types of motivation thought postulated by Deci and Ryan (2002) along with four other scales reputedly measuring academic motivation, subjective well being, vitality, and depression. Detailed analyses revealed that intrinsic motivation and autonomy correlated highly with self-reported "success" in study abroad. Underscoring the crucial role of autonomy in shaping motivation, the authors state, "it is not the content of peoples' motivation (intrinsic vs. extrinsic goals) but their level of *autonomy* that actually predict people's [well being]" (p. 199).

Laudibly, Chirkov, Vansteeniste, Tao, and Lynch used a variety of well-known scales that have been applied in previous motivational research. They offer a considerable body of evidence suggesting that the extent that respondents felt in control of their own SA decisions correlates strongly with their "academic motivation" and subjective well being. The least motivated informants were generally studying abroad to satisfy academic requirements or parental pressures. Conversely, the most motivated were doing so from a keen personal interest or belief that SA was "fun".

Also to their credit, Chirkov, Vansteeniste, Tao, and Lynch are transparent about their ideological biases at the onset of their study. They adhere to Deci and Ryan's views on motivation, not attempting to address factors mentioned in competing theories.

Another good feature of this study is the researchers beta-tested their survey, then revised it based on the feedback prior to large-scale sampling. Too many SA surveys are not adequately beta-tested. The fact that the researchers also dropped one of the measurement scales that was found to be statistically unreliable is commendable; it demonstrates a willingness to reject



findings that do not meet a priori standards of reliability or validity.

Finally, the authors include detailed effect size measures (expressed in terms of Cohen's *d*) to facilitate meta-analysis. According to Denis (2003, par. 1), too many quantitative studies neglect such measures.

One significant weakness of this study was that it was not sensitive to parameters outside the theoretical predictions. Positively speaking, this study could be described as "tightly focused". However, a limitation of studies of this sort is that a certain amount of "methodological straight-jacketing" occurs: a combined-method design with some more qualitative and open-ended responses that were coded a posteriori on the basis of the data assembled rather than any specific motivational theory might round out this study well. In fairness, such a procedure would involve a significant amount of extra work.

Also, studies relying entirely on self-reported data are vulnerable to some distortions and biases. To their credit, the authors used a wide range of different research instruments in their study. However, none of the instruments were based on observed behaviors, objective records, or on peer reports. Chirkov, Vansteenkiste, Tao, and Lynch's study has the same weakness as most single-method studies: the lens they used was sharply focused, but we are left to question whether the responses filled out on the paper-and-pencil questionnaire reflect "reality". Critical theorists in particular tend to view such data as one type of narrative discourse rather than as a measure of any "objective truth" (Bleakley, 2004).

Finally, the authors suggest that a parallel exists between *migration motivation* (decisions to shift long term residence to a foreign locale) and *study abroad motivation* (decisions to study overseas temporarily). However, there is no reason to assume that these two phenomena are based on identical motivations. What prompts some students to study at a foreign university for six months may quite differ from what prompts others to live abroad for six years.

## **(2) Kim and Yang (2010)**

Another excellent SA research project was about the learning beliefs and community participation patterns of two Korean university students who spent an academic year in North America. Adopting multiple qualitative research methods, Kim and Yang found that beliefs about L2 learning changed as one individual participated more actively in her L2 community. Another individual in this study remained isolated and did not manifest many positive gains as a consequence of her SA experience. According to Kim and Yang, "the crucial factor for successful SL L2 learning may not be L2-rich contexts per se, but the learners' recognition of



and appreciation for L2 participation” (p. 76). In other words, active involvement in a L2 community – rather than mere physical presence in it – is thought to impact foreign language learning motivation.

A particularly nice feature of their study was the use of triangulated data from semi-structured interviews, language learning autobiographies, and picture-cued stimulated recall tasks. Since readers may not be familiar with the latter two procedures, a brief description is in order.

The *language learning autobiographies* in this study consisted of a series of written answers to questions about the respondents' language learning history, academic backgrounds, and most/least pleasant experiences as L2 language learners. A sample language learner autobiography, likely similar to what Kim and Yang employed, is described by Goetz (2009).

*Picture-cued stimulated recall tasks* were photos from the respondents' blogs about their SA experiences. Basically, the researchers asked the respondents to comment on various photos, transcribing comments for later coding. A more detailed description of this type of activity for younger learners is offered by Fox-Turnbill (2009, 204-217). Collected memorabilia, student drawings, or non-blog photos can be used to elicit memories about various SA experiences, providing valuable springboards for reflective discourse.

Kim and Yang's study has many of the hallmarks of good qualitative research: detailed “thick” narrative descriptions, independent coding by more than one person, and a precisely stated inter-coder reliability coefficient. Importantly, the student participants were given the chance to verify the researcher codings to ascertain whether or not their interpretations reflected their beliefs. This is what Falout and Murphey (2010, pp. 370-380) refer to as “critical participatory looping” and it should be a feature of all qualitative SA research.

The main weakness of Kim and Yang's study concerned their sampling. Their study consisted of only two female respondents. A broader range of respondents might have yielded more diverse response patterns. In particular, the absence of male respondents in their study was perplexing. Although positivistic and post-positivist research generally attempts to be gender-blind, studies that are informed by constructionism and critical theory such as this recognize gender as a key variable in many social interactions. The fact that both researchers were female would probably be ignored in most traditional positivistic studies. In this type of study, however, it might be helpful to have persons of diverse gender and ethnicity code the data. Indeed, gender and ethnicity are variables that may have influenced the participant interactions in subtle ways.



The fact that 4 of the 6 potential participants opted out of this study also raises questions about what research incentives should be provided for informants. It may be unrealistic to expect large numbers of persons to participate in time-consuming interviews unless adequate incentives are provided.

### **(3) Goldstein & Kim (2006)**

Goldstein and Kim's study about pre-trip SA motivations had an interesting experimental design that is known as *case controlled study* (Mann, 2003, pp. 54-60). In this type of study, those with a specific outcome (in this case, university SA experience) are retrospectively compared with those lacking that outcome. The goal is to pinpoint factors that may account for the outcome difference. This type of study differs from a *paired propensity cohort study* (Stuart, 2010) in that the grouping is done post hoc rather than a priori. Goldstein and Kim administered a survey packet to 282 first year university students to assess their SA expectations, ethnocentrism, prejudice, intercultural communication apprehensions, foreign language interest and self-reported competence, as well as their tolerance of ambiguity and basic demographic variables. Four years later they checked the university records to ascertain which students had participated in SA programs at least a semester in length. The school records indicated 58% of the respondents ( $N=61$ ) who hadn't opted out of the original survey or left school prematurely had participated in semester or year-length SA programs. The characteristics of this group of students were then compared with the 105 students with no record of any extended SA experience during their university years. Case controlled studies permit a convenient post hoc comparison of two groups.

With the exception of one survey scale in this study, all six of the scales employed obtained Cronbach alpha levels exceeding the .80 minimum suggested by Nunnally (1978, cited in Goldstein and Kim, p. 514). One scale with a Cronbach alpha level of only .76 should have probably been deleted from this study. It may be worth mentioning once again that a hallmark of good quantitative research is that instruments with insufficient reliability or validity are discarded.

Goldstein and Kim then employed one-way ANOVAs and regression analyses to ascertain how the experimental and control groups differed. What they found was SA participation was ". . . associated with favorable expectations of study abroad, lower levels of ethnocentrism, intercultural communication, apprehension, prejudice, and ambiguity intolerance, as well as higher levels of foreign language interest and competence" (p. 517). The authors also found that



students with high levels of ethnocentrism or prejudice were less likely to participate in SA. Female students also were about twice as likely to participate in longer-term SA programs than males. Camilli and Shepard (1987, pp. 87-99) have raised serious questions about the appropriacy of ANOVAs in detecting bias. Ideally, Goldstein and Kim's classical analytical procedures should be reinterpreted within a one-parameter Rasch framework.

One good thing about this study is that the researchers acknowledged some of the sub-scales used were wobbly and in need of revision. They also wisely pointed out that the results might not be generalizable to other populations.

An inherent limitation in this study is that it did not consider short-term SA experiences of nine weeks or less. It also did not consider any not-for-credit SA experiences. This points out how easily data can slip through the cracks of many experimental designs. To some degree their research was also constrained by the theoretical orientations inherent in the six sub-scales. Important data that did not relate directly with those orientations could have easily been disregarded. It is for such reasons I am developing a belief that quantitative research is often best used in tandem with some qualitative methodologies.

Now let us turn our attention to three problematic SA studies.

### **Three Problematic SA Motivational Research Studies**

#### **(1) Hadis (2005)**

Hadis offers a convenient example of some mistakes to avoid when conducting SA research. He sought to gauge the impact of SA on American university students by use of one single online retrospective questionnaire.

Let us begin by considering some of this study's sampling problems, since sampling is a common problem in much SA research. Hadis sent email notices to 772 university students asking them to reply to a lengthy retrospective questionnaire about their SA experiences. Only 95 students with SA experience completed the survey, representing a 12% response rate. The fact that over 87% of the potential informants failed to complete the survey suggests that adequate incentives were not provided. Another possibility is that the survey task load was simply too high. Practical tips on how to design and implement online surveys are offered by Sue and Ritter (2007) as well as Bhaskaran and LeClaire (2010).

Now let us consider briefly the research instrument. Unfortunately the lack of detail makes it difficult to tell precisely what the instrument was, how many items it contained, or even what its response format was. Lerner and Ogren-Balkama (2009, p. 2) point out a widely accepted



adage: well-written research papers provide enough detail to enable readers to replicate the study. In this respect, Hadis has been most negligent.

Another concern not critically addressed by Hadis is whether post-hoc reflective responses can substitute for a pretest/posttest design. According to Hill and Betz (2005) there are serious validity concerns with using such retrospective reflections in lieu of pretests. The reconstructed memory of events that occurred months or even years ago frequently differs from the fresh impressions of such events due to recall bias and/or other types of bias. Retrospective questionnaires serving as "pretests" are probably not a reliable methodology from a positivist or post-positivist perspective. However, they might offer interesting insights about how people situate their experiences from a critical theory perspective. This illustrates how different research tools can be suited to different philosophical frameworks.

## (2) Meyer (2009)

Using a 26-item pretest/posttest design, Meyer sought to ascertain the effects of a 3-week Canadian SA experience on 31 Japanese high school students. To his credit, he employed 4 well-known attitudinal scales along with some open questions to explore possible reasons behind any changes. He found slight, but statistically insignificant ( $p > .05$ ) increases in the "international posture" (Yashima, Nishide, & Shimizu, 2004) and the motivation to learn English among participants. I will raise one concern about this study's design time frame, then a few apprehensions about the survey items themselves.

The pretest was administered seven months *prior* to the SA departure and the posttest shortly after the SA experience was completed. With this timeframe, it becomes difficult to ascertain whether any changes were due to the actual SA experience or to the pre-departure training. Moreover, it is difficult to know how lasting the changes were: it is not uncommon for some SA students in programs to exhibit a short-term surge in motivation that wanes as they re-socialize back into their familiar milieus. To obtain a more accurate picture of what changes (if any) may have occurred, a longitudinal study with at least two post-tests over an extended time frame should be employed.

Now let us turn our attention to the survey items. A common problem in much SA research is *confirmation bias* (Skeptic's Dictionary, 2011) – this is a sort of expectancy effect that occurs when respondents can discern researcher agendas. If the researcher is working with small numbers of students as a teacher in conditions in which complete anonymity does not exist, respondents may feel compelled to give socially desirable answers rather than those that



are entirely candid. Questions such as “I often try to understand English movies and TV” (Item 7) or “I look forward to my English classes” (Item 14) have an agenda that is all too transparent. As Marsden and Wright (2010, p. 281) suggest, to reduce confirmation bias it is good to use *counterbalanced items* that ask essentially the same thing with a completely different nuance. For example Item 7 could be counterbalanced by a statement such as, "Listening to movies and TV programs in English is a hassle." Item 14 could be counterbalanced by stating, "My English classes usually are boring." Counterbalanced items make it easier for students to express a wider range of viewpoints.

Another concern is that some survey questions might not measure the intended construct. For example, question #11 “I am very interested in the world outside Japan” can be interpreted in various ways. It is quite possible that "the world" does not mean the same thing to all people. In other words, it is very hard to design attitudinal questionnaires that measure *only* a single intended construct. That is why beta testing survey items with a dozen or so students using think-aloud protocols (van Someren, Barnard, & Sandberg, 1994) is recommended: these might offer some insight into what is going on in the minds of potential respondents.

Another point to consider concerns the response format for some questions. For example, question #4 “I don't mind initiating conversations with foreigners” is probably not well suited to a multiple-choice response format: this type of complex statement is probably better handled though an open-response format. We should also reflect on whether most respondents conceive of a "foreigner" as single mental construct. It is quite possible some Japanese classify foreigners into a variety of mental categories. This survey item could create an artificial unidimensionality that might not exist. Indeed, the bulk of quantitative research can be said to have a unidimensional bias: it seeks to generalize and universalize experiences that do not necessarily fit neatly into single empirical categories.

Having said this, Meyer's work provides some evidence that short-term SA programs do not necessarily produce dramatic results in participant motivational levels. The small sample size and small amount of items for each construct, however, make the results inconclusive. One way to enhance the reliability of studies such as this would be to have at least three items tapping into each key variable so that the study could be used in *factor analysis* (DeCoster, 1998). Another way to enhance this sort of study would be to use *cohort sampling* (Hefferman, n.d.). In Meyer's case, that would probably involve using data from two separate cohorts of SA students over a two-year period rather than data from a single year.





### (3) Newfields (2010)

To make it clear that I am by no means immune from bad writing, I cite one hastily written SA research paper recently published. The fact that it was a nine-month longitudinal study with two different post-return samplings (return day, and 8 months later) is the only redeeming feature of this study. This afforded a chance to ascertain how lasting the impact of a three-week SA experience in England appeared to be. Also, the suggestion that about a third of the respondents made their SA decisions primarily out of touristic motives rather than any desire to improve their English or communicate with foreigners is intriguing, but the sample size ( $n=25$ ) was too small for proper quantitative research and the questionnaire itself should have been beta tested more carefully. This research project could have yielded more reliable and valid results by utilizing either multiple cohort samples over several years and/or by employing qualitative methodologies such as *journaling* (Newbury, 2001) or *in-depth interviews* (Warren & Karner, 2010).

Let me briefly mention three weaknesses of this study simply because many of the small-scale SA research studies have similar limitations.

First of all, the *informed consent procedures* were not congruent with current standards. Until recently it was common practice in many parts of Asia to simply hand out surveys to students and politely ask them to complete them with minimal explanation. As more publications require researchers to adhere to guidelines similar to TESOL's Informed Consent Policy Statement (2007), it has become an accepted procedure to inform students in detail about the study, request their voluntary cooperation, indicate any possible incentives that might exist, and finally remind them that they may opt out of the study. Those doing social science research should probably read the literature review by Wiles, Heath, Crow, and Charles (2005) on informed consent.

A second problem with this article involved the lack of proper *back-translation* procedures of the survey items. As Griffiee (1998) points out, at times it is difficult to capture the same nuance between expressions in two different languages and therefore an independent back-translation procedure is recommended. For research transparency, it is also probably best to have any bilingual research instruments listed in both languages in separate appendices at the end of the paper. Although these were not included in the article due to space constraints, the URLs to both the Japanese and English online surveys were mentioned.

Finally, this study relied entirely on student self-reports. Such information is prone to all sorts of biases and needs to be corroborated with other types of data such as peer reports,



teacher observations, or classroom records. As Bowen (1996) points out, self-reports are not entirely worthless, but they need to be backed up by other sources of information.

## Conclusion

This paper has pointed out some of the research methods being employed to explore the impact of SA experiences on motivation. The more successful research studies mentioned are characterized by an overall congruence between their philosophical frameworks and actual research procedures. The less successful ones exhibit some incongruities in this respect. In particular, the sampling and data collection procedures in many studies fall short of the research goals they espouse.

So what do we actually know about the impact of SA on motivation? I will conclude by venturing two brief hypotheses. First of all, it is probably naïve to assume that SA programs will automatically enhance the motivation to study a foreign language or culture unless participants interact with members of a target group to an extended degree over a period of time. The previously cited study by Meyer (2009) can be interpreted to support the *threshold hypothesis* that posits a requisite level of linguistic input and interaction with members of a target community is needed before significant changes occur.

Second, the studies by Chirkov, Vansteenkiste, Tao, and Lynch (2007) as well as Kim and Yang (2010) made it clear that a sense of "ownership" in the SA decision process may be important in enhancing motivation. Those who are forced to study abroad to fulfill some academic requirement and those who do so voluntarily tend to have differing *initial* levels of language learning motivation. What many motivation studies made clear, however, is that motivation levels can change rapidly in response to critical incidents or significant stress (Fryer, 2012). Shocks such as property theft, racial discrimination, the breakup of an intimate relation with a target language member, or even lower than expected performance on a high-stakes foreign language test can significantly erode foreign language learning motivation. For this reason, Dörnyei (2001, cited in Falout & Maruyama, 2004, p. 3) emphasizes the need to study *demotivational factors* closely in order to understand motivation in a more balanced way. We are now beginning to understand motivation not merely as a psychological construct, but also as a complex array of attitudes and behaviors that are socially mediated to a significant degree.

## Acknowledgements

*Many thanks to Joe Falout, Andrew Atkins, and Gordon Myskow for their detailed feedback on this article.*



## References

- Bhaskaran, V., & LeClaire, J. (2010). *Online surveys for dummies*. New York: John Wiley & Sons.
- Bleakley, A. (2004). *Education research in the postmodern*. Retrieved from <http://www.edu.plymouth.ac.uk/resined/postmodernism/pmhome.htm>
- Bowen, K. A. (1996). *The sin of omission - punishable by death to internal validity: An argument for integration of qualitative and quantitative research methods to strengthen internal validity*. Retrieved from <http://www.socialresearchmethods.net/Gallery/Bowen/hss691.htm>
- Camilli, G. & Shepard, L. A. (1987). The inadequacy of ANOVA for detecting biased items. *Journal of Educational Statistics*, 12 (1). 87-99.
- Chirkov, V., Vansteene, M., Tao, R. & Lynch, M. (2007). The role of self-determined motivation and goals for study abroad in the adaptation of international students. *International Journal of Intercultural Relations*, 31 (2), 199-222. doi: 10.1016/j.ijintrel.2006.03.002
- Chomsky, N. (2000). *Chomsky on miseducation*. Lanham, MD: Rowman & Littlefield Pub Inc.
- Confirmation Bias. (2011). In R. T. Carroll (Ed.), *The skeptic's dictionary*. Retrieved from <http://skepdic.com/confirmbias.html>
- Deci, E., & Ryan, R. (Eds.), (2002). *Handbook of self-determination research*. University of Rochester Press.
- DeCoster, J. (1998). *Overview of factor analysis*. Retrieved from <http://www.stat-help.com/notes.html>
- Denis, D. J. (2003). Alternatives to null hypothesis significance testing. *Theory & Science*, 4 (1). n.p. Retrieved from [http://theoryandscience.icaap.org/content/vol4.1/02\\_denis.html](http://theoryandscience.icaap.org/content/vol4.1/02_denis.html)
- Falout, J., & Maruyama, M. (2004). A comparative study of proficiency and learner demotivation. *The Language Teacher*, 28(8) 3-9.
- Falout, J., & Murphey, T. (2010). Loop it! Student participatory research. In A. M. Stoke (Ed.), *JALT2009 Conference Proceedings*. (pp. 370-380). Tokyo: JALT. Retrieved from <http://jalt-publications.org/archive/proceedings/2009/E092.pdf>
- Fox-Turnbull, W. (2009). Stimulated recall using autophotography: A method for investigating technology education. In A. Bekker, I. Mottier, & M. J. de Vries (Eds.) *Proceedings of the 22nd Int. Technology & Engineering Educator's Association Conference*. (pp. 204 - 217). Retrieved from <http://www.iteaconnect.org/Conference/PATT/PATT22/FoxTurnbull.pdf>
- Fryer, M. (2012). Understanding L2 motivation in a study abroad context. *Ryūgaku: Explorations in Study Abroad*, 5 (1), 14-26. Retrieved from <http://jalt-sa.org/PDF/5-1Fryer.pdf>
- Goldstein, S. B., & Kim R. I. (2006). Predictors of US college students' participation in study abroad



- programs: A longitudinal study. *International Journal of Intercultural Relations*, 30, 507-521.  
doi:10.1016/j.ijintrel.2005.10.001
- Goetz, L. A. (2009). *Foreign language learning autobiography*. Retrieved from  
<http://www.longwood.edu/staff/goetzla/201/>
- Griffee, G. (1998). Can we validly translate questionnaire items from English to Japanese? *Shiken: JALT Testing & Evaluation SIG Newsletter*, 2 (2) 11 - 14. Retrieved from  
<http://www.jalt.org/test/PDF/Griffee1.pdf>
- Hadis, B. F. (2005). Gauging the impact of study abroad: How to overcome the limitations of a single-cell design. *Assessment and Evaluation in Higher Education*, 30 (1), 3-19. doi:  
0.1080/0260293042003243869
- Hefferman, C. (n.d.). *Cohort studies*. Retrieved from <http://www.drcath.net/toolkit/cohort.html>
- Hill, L. G., & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation*, 26 (4). 501-517. doi: 10.1177/1098214005281356
- Kincheloe, J. L. (2002). *Teachers as researchers: Qualitative paths to empowerment*, 2nd edition.  
New York: Routledge Palmer.
- Kincheloe, J. L. (2005). *Critical constructivism primer*. New York: Peter Lang.
- Kim, T. Y., & Yang, J. S. (2010). Learner beliefs in study abroad experience: A qualitative case of two Korean ESL learners. *English Teaching*, 65 (2), 75-99. Retrieved from  
<http://works.bepress.com/taeyoungkim/13/>
- Lerner, N., & Ogren-Balkama, M. (2009). *A guide to scientific writing*. Retrieved from  
[http://openwetware.org/images/6/69/Guide\\_to\\_Scientific\\_Writing.pdf](http://openwetware.org/images/6/69/Guide_to_Scientific_Writing.pdf)
- Mann, C. J. (2003). Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*, 20. 54-60. doi:10.1136/emj.20.1.54
- Marsden, P. V. & Wright, J. D. (2010). *Handbook of survey research*. Bingley, West Yorkshire: Emerald.
- Meyer, D. (2009). A change of heart: Motivation and studying abroad. In A. M. Stoke (Ed.) *JALT2008 Conference Proceedings*. 183-192. Tokyo: JALT. Retrieved from  
<http://jalt-publications.org/recentpdf/proceedings/2008/E095.pdf>
- Newbury, D. (2001). Diaries and fieldnotes in the research process. Retrieved from  
<http://www.biad.bcu.ac.uk/research/rti/riadm/issue1/riadmIssue1.pdf>
- Newfields, T. (2010). A case study of longitudinal attitude changes among a cohort of Japanese university students completing a short-term study abroad. *Toyo University Keizai Ronshu*, 35 (1), 27 - 41.  
Retrieved from <http://www.tnewfields.info/Articles/SACasestudy.htm>
- Papert, S. (1994). *The children's machine: Rethinking school in the age of the computer*. New York:



Basic Books.

- Piaget, J. (1964, 1967). *Six études de psychologie* [Six psychological studies]. trans. by A. Tenzer & D. Elkind. New York: Vintage Books.
- Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching*, 42 (2), 147-180. doi:10.1017/S0261444808005612
- Schroyer, T. (1973). *The critique of domination: The origins and development of critical theory*. New York: George Braziller.
- Stuart, E. H. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25 (1), 1-21. doi: 10.1214/09-STS313
- Sue, V. M., & Ritter, L. A. (Eds.). (2007). *Conducting online surveys*. Thousand Oaks, CA: Sage Publications.
- TESOL Inc. (2007). *Informed Consent Policy Statement*. Retrieved from [http://www.tesol.org/s\\_tesol/sec\\_document.asp?CID=476&DID=1014](http://www.tesol.org/s_tesol/sec_document.asp?CID=476&DID=1014)
- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London: Academic Press.
- Von Glasersfeld, E. (1987). *The construction of knowledge: Contributions to conceptual semantics*. Salinas, CA: Intersystems Publications.
- Warren, C. A. B., & Karner, T. X. (2010). *Discovering qualitative methods: Field research, interviews, and analysis* (2nd Edition). New York: Oxford University Press.
- Wiles, R., Heath, S., Crow, G. & Charles, V. (2005). Informed consent in social research: A literature review. *NCRM Methods Review Papers #001*. Retrieved from <http://eprints.ncrm.ac.uk/85/1/MethodsReviewPaperNCRM-001.pdf>
- Yashima, T., Nishide L. Z., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54 (1), 119-152. doi: 10.1111/j.1467-9922.2004.00250.x