Exploring the Relationships Between Tutor Background, Tutor Training, and Student Learning: A Problem-based Learning Meta-Analysis

Heather Leary  
*University of Colorado Boulder, heatherleary@gmail.com*

Andrew Walker  
*Utah State University, andy.walker@usu.edu*

Brett E. Shelton  
*Utah State University, brettshelton@boisestate.edu*

M. Harrison Fitt  
*Utah State University, melynda.fitt@gmail.com*

---

**Recommended Citation**

Exploring the Relationships Between Tutor Background, Tutor Training, and Student Learning: A Problem-based Learning Meta-Analysis

Heather Leary, Andrew Walker, Brett E. Shelton, and M. Harrison Fitt

Abstract
Despite years of primary research on problem-based learning and literature reviews, no systematic effort has been made to analyze the relationship between tutor characteristics and student learning outcomes. In an effort to fill that gap the following meta-analysis coded 223 outcomes from 94 studies with small but positive gains for PBL students ($g = 0.24$). Specific sub-group analyses indicate randomly controlled designs may be more sensitive to differences that favor PBL students, even while there is no relationship between tutor content expertise and student learning. Perhaps surprisingly, student learning decreases as tutor experience increases. Limitations and future work are discussed within a context of scholarly and practical significance.

Keywords: meta-analysis, tutors
Introduction

Experimental studies of problem-based learning (PBL) originated almost four decades ago (Neufeld & Barrows, 1974). Since then, several meta-analyses have been published (Albanese & Mitchell, 1993; Dochy, Segers, Van den Bossche, & Gijbels, 2003; Gijbels, Dochy, Van den Bossche, & Segers, 2005; Kalaian, Mullan, & Kasim, 1999; Vernon & Blake, 1993; Walker & Leary, 2009) the most recent of which incorporates findings across several subject areas and educational levels (Savery, 2006; Savery & Duffy, 1995). According to these reviews, effect sizes that favor PBL depend in part on the kind of assessment used (Gijbels et al., 2005; Walker & Leary, 2009), but unexplained variance persists. One potential source for the variance is the facilitator, or tutor. There is primary research examining tutor training, tutor behavior, and the role of the tutor in PBL interventions (Hmelo-Silver & Barrows, 2006; Moust, De Grave, & Gijsselaers, 1990) but there are no meta-analyses that focus directly on PBL tutor background and training.

There is general consensus that training tutors is critical in the PBL approach (Barrows, 1996; Bochner, Badovinac, Howell, & Karimbux, 2002; Dolmans et al., 2002; Eagle, Harasym, & Mandin, 1992; Hmelo-Silver & Barrows, 2006). However, the impact of tutor training on student learning outcomes remains unclear. Training for PBL tutors may include a range of experiences including single workshops where PBL techniques are taught, or intense one-week workshops where tutors learn about PBL and develop tutor skills. Some workshops are sustained, and accompanied by subsequent weekly reminders of what constitutes “good” tutoring techniques (Alleyne et al., 2002; Distlehorst & Robbs, 1998; Mergendoller, Maxwell, & Bellisimo, 2000). In spite of some of those training efforts, there are documented cases of trained PBL tutors turning small group sessions into lectures (Moust et al., 1990).

In contrast to recommendations for training, there is a fair amount of debate about the optimal background for tutors. The PBL tutor is known as the guide or mentor for the student. Tutors prompt students with meta-cognitive questions and provide direction without directly telling the student what to look for and where to go for information. The tutor provides a student-centered learning environment by promoting self-directed learning, the integration of previous knowledge, interaction, and guiding the learning process (Chan, 2008; De Grave, Dolmans, & van der Vleuten, 1999; Hmelo-Silver & Barrows, 2006). Within existing PBL literature, some researchers promote tutors as requiring content expertise (Barrows, 1996; Hmelo-Silver & Barrows, 2006; Dede, 2003), while some argue content experts are not necessary (Barrows, 1986, 1998; Swanson, Stalenhoef-Halling, & van der Vleuten, 1990), and still others claim that content experts and content novices should be used at different stages of PBL instruction (Schmidt, Van Der Arend, Kokx, & Boon, 1994). There may even be an interaction between training and background, with evidence that training is particularly important for tutors with content expertise (De Volder, 1982; Silver & Wilkerson, 1991).
Although there is some primary research on both tutor training and tutor background it is dwarfed by the number of research articles that report PBL interventions in general. Much like previous contributions to the literature on different levels of assessment (Gijbels et al., 2005) or differences between disciplines (Walker & Leary, 2009) there is a need to investigate the claims regarding PBL tutors. The primary purpose of this research is to expand on existing review efforts by investigating the relationships between tutor training, tutor background, and student learning outcomes.

Literature Review

As PBL gained in popularity over recent years, it became associated with several different variations, reflecting institutions modifying PBL to meet their own particular needs (Barrows, 1996). Multiple definitions make characterizing precisely what constitutes PBL challenging and all the more important. For the purposes of this research, PBL is an approach to learning (Barrows, 2002) that includes the following elements:

- Unresolved and ill-structured problems are presented to students who generate multiple thoughts about the cause of the problems, and further, multiple thoughts on the process of how to solve them. While many PBL interventions encourage the finding of solutions, some problem types such as dilemmas (Jonassen, 2000) may not have a resolution.

- A student-centered format must exist in which students determine what they need to learn. Students generate a list of the key issues for a particular problem, identify what they already know, what they need to investigate, and then acquire and apply the missing knowledge.

- Tutors, typically instructors, act as facilitators or guides. Tutors initially ask students meta-cognitive questions about the problem-solving process. Over time, tutors gradually ask students to assume more responsibility for guiding the process through their own questioning.

- Authenticity forms the basis of problem selection, embodied by an alignment to professional or “real world” practice. Such authenticity requires that problems be cross disciplinary and unconstrained, representing the same complexity found by current practitioners.

- Learners typically work together within small groups.

Prior Reviews

A great deal of attention has been given to research on the PBL approach, enough to warrant a synthesis of existing meta-analytic reviews (Barneveld & Strobel, 2009). Both the synthesis and the prior meta-analyses represent large contributions to our understanding
of PBL. The literature is uniform, for instance, on the notion that effect size differences are impacted by the nature of the student assessment. That is, PBL students generally perform better as the focus of assessment moves from knowledge and facts to more complex forms of reasoning (Dochy et al., 2003; Walker & Leary, 2009). PBL also appears to result in better retention over time of what is learned (Barneveld & Strobel, 2009). While these findings are important, there may be methodological flaws in the analyses previously summarized (Barneveld & Strobel, 2009) even within recent publications (Walker & Leary, 2009). That is, many of the existing meta-analyses report sign tests from which misleading conclusions could be formed (Borenstein, Hedges, Higgins, & Rothstein, 2009). Perhaps more importantly, only one meta-analysis examined research design as a factor (Dochy et. al., 2003). It is understandable that a subsequent analysis covering many of the same studies did not examine research design (Gijbels et al., 2005) but after studies from several different subject areas were added (Walker & Leary, 2009), research design was still omitted. Some researchers consider research design and a broader examination of research quality to be an important consideration within all meta-analyses (Wortman, 1994). As part of this review, research design will be considered.

PBL Tutors

There are many definitions of PBL tutoring that reflect different opinions about their role, function and ideal traits (Blumberg, Michael, & Zeitz, 1990). To a certain extent, PBL has evolved over time, which may explain the range of tutor characteristics advocated by various authors (Bochner et al., 2002; Kwizera, Dambisya, & Aguirre, 2001). According to Barrows (1998), the role and function of the PBL tutor is to raise student awareness in higher cognitive thinking and question development. Hmelo-Silver and Barrows (2006) later added that tutors facilitate the collaborative construction of knowledge by students. Their role goes beyond the facilitation of student knowledge construction to progressively turn the role of facilitator over to their students. Tutors model desired behaviors, monitor discussions, and focus student efforts on deep and critical thinking (Hmelo-Silver & Barrows, 2008).

Effective tutors are defined as expert learners who can model their own learning strategies by asking meta-cognitive questions and focusing on the process of learning. Tutor training teaches them to ask guiding and clarifying questions, facilitate discussion, thinking, revoice, model, and reframe questions and discussion (Hendry, 2009; Moore & Kain, 2011; Zhang, Lundeberg, & Eberhardt, 2011). Tutors are also instructed to be a supporter and facilitator for students in the process of PBL (Dahlgren et al., 1998), but sometimes perception of this role leads to confusion and movement away from a supporter to a resource for content (Papinczak, 2010). The perceived role of a PBL tutor influences the role chosen and adopted in the classroom (Moore & Kain, 2011).
Tutor Content Expertise

Throughout the PBL literature, tutors are characterized along a continuum that includes content experts, typically faculty, to content novices, typically students. Determining which characteristics of a tutor promote student learning the best has been an extensive debate (Chng, Yew, & Schmidt, 2011; Dahlgren et al., 1998; Moore & Kain, 2011; Woltering, Herrler, Spitzer, & Spreckelsen, 2009). At the inception of PBL, McMaster University promoted the idea of using a content novice tutor to keep faculty members from reverting to old teaching habits, such as lecturing (Barrows, 1996). As PBL spread, the approach was refined. Content expertise became less important than facilitation expertise (Barrows & Tamblyn, 1980; De Volder, 1982; Dolmans et al., 2002; Eagle et al., 1992). A review of recommendations for optimal tutor characteristics within PBL literature is often confusing, with some recommendations being closely overlapped while others offer direct contradictions. For example, some researchers recommend that tutors be content experts or faculty with facilitation training (Barrows, 1996; Bochner et al., 2002; Gilkison, 2003; Schmidt & Moust, 1995; Schmidt, Van Der Arend, Moust, Kokx, & Boon, 1993; Schmidt et al., 1994). An early effort that characterizes the impacts of tutor expertise in relation to faculty direction of students (Albanese & Mitchell, 1993) found that expert tutors provide less engagement for student-directed discussion and learning (Davis, Nairn, Paine, Anderson, & Oh, 1992; Silver & Wilkerson, 1991) and were more likely to intervene in student-directed discussion (De Volder, 1982). Since the goal of PBL is to promote self-directed learning, these seem to provide a rationale against the use of expert tutors. Other researchers such as Wilkerson (1992) found physician tutors were well-positioned to synthesize multiple perspectives. Students tend to value tutors’ knowledge of the associated content (Feletti et al., 1982), which seems to support the use of faculty or instructors with related expertise but may contradict the intended meta-cognitive role for tutors (Hmelo-Silver & Barrows, 2006) if that knowledge is used to correct student efforts or lecture.

Perhaps because of these contradictory findings, debate persists with some scholars claiming that content novices positively impact student outcomes (Silver & Wilkerson, 1991), especially in more affective areas like self-directed learning. Still, other researchers view content novices as equally effective for PBL tutoring to content experts (De Volder, De Grave, & Gijselaers, 1985; Hendry, Phan, Lyon, & Gordon, 2002; Kwizera et al., 2001; Moust, De Volder, & Nuy, 1989; Moust & Schmidt, 1994; Park, Susarla, Cox, Silva, & Howell, 2007; Regehr et al., 1995; Steele, Medder, & Turner, 2000; Swanson et al., 1990), claiming that content expertise is at odds with good facilitation because an expert will constantly inject their content knowledge (Des Marchais, Bureau, Dumais, & Pigeons, 1992; Moust et al., 1990; Silver & Wilkerson, 1991) and suppress the student-directed design. Hmelo-Silver and Barrows (2006) report that tutor training is the most important factor in effective PBL, and suggest that content expertise is helpful but not critical for success. Meanwhile, Groves,
Rego, and O’Rourke (2005) found from student surveys that both content expertise and facilitation skills are necessary.

**Tutor Training and Experience**

PBL scholars generally agree that tutors should be trained in the process of PBL (Baroffio, Nendaz, Perrier, & Vu, 2007; Daniel, 2004; De Volder, 1982; Hmelo-Silver & Barrows, 2006; Wikerson & Hundert, 1991). To act as an appropriate facilitator, tutors should and often return to process-facilitation skills from training workshops (Barrows, 1998; Dolmans et al., 2002; Hendry, 2009; Hmelo-Silver & Barrows, 2006; Papinczak, 2009). PBL tutors must understand how their role as a tutor changes during the course of a particular problem. In addition, tutors should have a great deal of familiarity with the problem and common approaches to solving it, either as a result of closely collaborating with the case designer or by co-authoring the instructional materials (Chan, 2008; Davis et al., 1992; Johansen, Martenson, & Bircher, 1992). Despite uniform calls for training, some studies have openly used inexperienced tutors with little training (Steinkuehler, Derry, Hmelo-Silver, & Delmarcelle, 2002), but those studies are the exception rather than the norm. Even while PBL scholars agree that tutors need some form of training, very few empirical investigations exist that have examined the relationship between tutor training and student learning (Budé, Imbos, Wiel, Broers, & Berger, 2009; Budé, van de Wiel, Imbos, & Berger, 2011; Chng et al., 2011).

During the course of data collection, an emergent code for tutor experience was developed. This coding was a direct result of seeking to find information about training; several articles instead provided information about experience. In some cases, tutors were identified as having experience but the number of years was not specified. While investigations of tutor content expertise are fairly robust, little has been done to examine tutor experience levels (Park et al., 2007). What has been done is limited by available data to making binary comparisons on presence or absence of tutor experience rather than varying amounts, and only among tutors with expertise. Findings (Park et al., 2007) for the most part indicated no differences, among several student outcomes. The sole significant comparison favored tutors without experience. A more broadly sense of PBL literature indicates students in general appear to benefit from having experienced teachers (Hedges, Laine, & Greenwald, 1994), particularly when controlling for when teachers were first hired (Murnane & Phillips, 1981).

The review of PBL literature indicates that despite a large volume of primary research findings and several meta-analyses (Albanese & Mitchell, 1993; Dochy et al., 2003; Gijbels et al., 2005; Kalaian et al., 1999; Vernon & Blake, 1993; Walker & Leary, 2009), including an early narrative summary of expert and non-expert differences (Albanese & Mitchell, 1993), no meta-analysis exists of the literature across subject areas regarding the impact
of tutor background and tutor training on PBL student outcomes. In addition, the most recent meta-analyses do not address the quality of the research being conducted. The following research questions set out to address this gap in the literature:

1. What is the relationship between research design and student learning?
2. What is the relationship between tutor content expertise and student learning?
3. What is the relationship between tutor training and student learning?
4. What is the relationship between tutor experience and student learning?

Methods and Data Analysis

Based on Barrows (2002), PBL studies included in this analysis used authentic and cross-disciplinary ill-structured problems, with students at the center of learning activities, teachers as facilitators, and small group learning. However, since small group is described as frequent (2002) rather than required and Barrows (1986) has experimented with large group PBL as well, we accepted PBL interventions irrespective of group size.

The literature search began with primary research reported in existing meta-analyses (Albanese & Mitchell, 1993; Dochy et al., 2003; Gijbels et al., 2005; Kalaian et al., 1999; Vernon & Blake, 1993; Walker & Leary, 2009), then used keywords (e.g. problem-based learning, achievement, empirical, high school) obtained from each study to search prominent databases including PsychInfo, ERIC, Education Full Text, and Digital Dissertations. A variety of databases were used to minimize publication bias. Additional referrals were obtained from citations in primary research articles.

To be included in the analysis, studies selected needed to report quantitative cognitive outcomes focused on student learning or their reasoning processes. The quantitative data needed to report enough information to calculate an effect size. The methods needed to compare a lecture/traditional control group with a problem based learning treatment. Problem based learning had to involve ill-structured problems, student-directed learning, and tutors acting as facilitators (Barrows, 1986). Treatments did not necessarily include closed-loop problem based learning, and employed a full range of group sizes. Manuscripts with only qualitative data or no comparison group were not included.

Each study was independently coded by two researchers for content expertise (expert, novice, mixed), facilitation training (yes, no), years of facilitation experience, study design (random, and non-experimental comparison group; Shadish & Myers, 2004), and finally, effect size. To code for these areas, the manuscripts often reported these parameters directly, providing information such as the number of tutors used in a course. Additional details about individual coding categories are provided alongside results below. Discrepancies, largely due to omission rather than differences of opinion, were resolved until consensus was achieved (Stemler, 2004).
Study outcomes were placed on the common scale of standardized mean difference ($d$). To adjust for differences in variance and sample size and to correct for slight bias (over-estimate of the effect size), each effect size point estimate was converted from Cohen’s $d$ to Hedges’ $g$ with positive effect sizes indicating differences favoring PBL students. An analysis of main effects for content expertise, facilitation expertise and the interaction of content and facilitator expertise was completed using an ANOVA based Q-test. Pairwise differences were explored using a $Z$-test with random effects weights, and the same test was used to determine if any one group of studies significantly differed from zero (Borenstein et al., 2009). All statistical significance testing used an alpha level of 0.05.

**Results**

A total of 223 outcomes with codeable effect sizes from 94 studies were utilized in the meta-analysis. Many of the manuscripts included in the analysis reported more than one outcome. The overall effect size estimate ($g = 0.24$) suggests small differences favoring PBL. Unless otherwise noted, all effect size estimates are statistically greater than zero. For the overall effect size, a test for heterogeneity ($Q = 1307.72, I^2 = 83.0\%, p = 0.01$) suggests that observed outcomes in the sample are markedly different from each other and warrant closer examination.

**Research Design**

The initial research question focuses on the relationship between research design and student learning. As an accessible proxy for methodological rigor of the studies, each was coded as being either a random design (meaning random assignment of participants to the treatment or control condition) or a non-equivalent comparison group (meaning use of intact classrooms, self-selection by participants, or other means of assignment). There was enough data about participant assignment to make a judgment for all outcomes.

**Figure 1.** Learning outcomes by research design. $n_{\text{out}} =$ Outcome N, $n_{\text{treat}} =$ Treatment N, $n_{\text{ctrl}} =$ Control N, $g =$ Effect size (Hedges’ $g$)
As shown in Figure 1, there was a marked difference. Outcomes with random designs \((g = 0.46, z = 4.30, p < 0.01)\) reported stronger learning gains for PBL students \((z = 2.51, p < .01)\) than studies with non-equivalent designs \((g = 0.21, z = 6.14, p < .01)\).

At more than double the value of the overall mean effect, the difference is not only statistically significant but also meaningful, moving from a small effect to a medium difference. Based on this relationship it is likely that the overall effect size is an underestimate of lecture vs problem-based learning differences, to which random designs are more sensitive. While it is justifiable to weight adjust or engage in differential reporting across all studies for the rest of the results, we chose a more conservative approach by keeping the data as is and reporting both random and non-equivalent comparison studies together.

**Tutor Expertise**

Research question two focuses on the level of content expertise held by PBL tutors. Our initial coding was revised, expanding from content expert (meaning tutors had received at least one degree above their learners) and novices (meaning the same degree level as learners) to accommodate a mixture of the two as identified in several \((n = 16)\) study outcomes. Mixed expertise required only one from each category. Figure 2 shows that all variations of tutor expertise were associated with similar student learning gains. None of the pairwise comparisons between level of expertise were significant, and each was significantly larger than zero. Note that in contrast with the other groups content novices \((g = 0.25, z = 4.22, p < 0.01)\) failed a test for heterogeneity \((Q = 19.92, I^2 = 34.7\%, p = .10)\).

In short, content expertise does not appear to have a relationship with student learning. More than a quarter of the outcomes \((n = 60)\) did not report enough information about tutors to categorize them, while they are omitted from the analysis the overall result is

**Figure 2.** Learning outcomes by tutor content expertise. Of 163 outcomes reporting, the combined effects of expert, novice, and mixed \((g = 0.27, z = 6.75, p < 0.01)\) are similar to the entire set of 223. \(n_{out} = \) Outcome \(n_{treat} = \) Treatment \(n_{cntrl} = \) Control \(g = \) Effect size (Hedges' \(g\))
shown for the entire set \((n = 223)\) to permit comparisons, a feature of all subsequent analyses. Like all subsequent analyses the combined effects of expert, novice, and mixed as a subset of outcomes is statistically identical to the overall effect with every study included.

**Tutor Training**

Research question three shifts toward training for tutors. Training included some mention of professional development, workshops, or other experiences targeted at helping faculty take on the role of PBL tutor. Less than half \((n = 104)\) described whether or not training occurred, a surprise given the general agreement in the literature about the importance of tutor training. Somewhat less surprising is that of those mentioning training only a handful \((n = 3)\) specifically report not training tutors \((g = -0.01, z = 0.03, p = 0.98)\). This result represents the only subgroup estimate that was not statistically greater than zero, suggesting similar learning gains with lecture based interventions. A pairwise comparison yields no differences between trained and untrained tutors. This similarity is likely due to the wide dispersion for untrained tutor interventions.

**Figure 3.** Learning outcomes by tutor training. Of 104 outcomes reporting, the combined effects of training or no training \((g = 0.28, z = 6.03, p < .01)\) are similar to the entire set of 223. \(n_{out}\) = Outcome N, \(n_{treat}\) = Treatment N, \(n_{cntrl}\) = Control N, \(g\) = Effect size (Hedges’ \(g\))

<table>
<thead>
<tr>
<th>Training</th>
<th>(n_{out})</th>
<th>(n_{treat})</th>
<th>(n_{cntrl})</th>
<th>(g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>101</td>
<td>6652</td>
<td>8962</td>
<td>0.29</td>
</tr>
<tr>
<td>no</td>
<td>3</td>
<td>184</td>
<td>178</td>
<td>-0.01</td>
</tr>
<tr>
<td>overall</td>
<td>223</td>
<td>15201</td>
<td>19488</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Outcomes with both trained and untrained tutors were significantly heterogeneous. For those tutors who did receive training, differences in the length of tutor training as well as how sustained training was over time were reported. Since few studies went into detail, coding for these differences is not possible but approaches to training might be responsible for explaining additional variation within groups. Another possible explanation is the experience level of tutors. Several outcomes \((n = 119)\) lack enough information to characterize training reported tutors’ years of experience in facilitating PBL problems.

**Tutor Experience**

Research question four explores the emergent coding for tutor experience. Experience represents the only ratio scale among potential predictors for student learning. Since
experience is expressed in years and with a meaningful zero point, meta-regression is an appropriate analysis. Several of the studies were non-specific about how many years of experience tutors had. For that reason, the data were represented both as a scale, omitting studies that did not report years of experience, and as a forced dichotomy. Data and corresponding analyses are thus presented twice.

A total of 144 outcomes reported enough data or directly reported years of experience. Experience \((n = 144, M = 3.05, sd = 3.29)\) as shown in Figure 4 was skewed positive, with several \((n = 56)\) outcomes associated with 0–1 years of experience. The highest value was 14 for three outcomes.

**Figure 4.** Years of PBL experience.

Meta-regression with a single predictor variable (years of experience) accounted for 9.3% of the variability \((R^2 = 0.093)\) in student learning outcomes. The relationship is inverse \((t = -3.24, p < 0.01)\) with student learning dropping as years of experience increases.

Parallel results are found with the forced dichotomy. As shown in Figure 5, tutors without experience \((g = 0.31, z = 5.03, p < 0.01)\) reported stronger learning gains for PBL students than tutors with experience \((g = 0.19, z = 3.99, p < 0.01)\) at a statistically significant \((z = 3.07, p < 0.02)\) level.

To further explore these results, outcomes were split according to the length of intervention. Studies reporting a program wide or multiple class PBL intervention \((n = 101)\) continue to have an inverse relationship \((t = -4.03, p < 0.01)\), which accounts for 18.7% of
the variability in student learning ($R^2 = 0.19$). The inverse relationship does not hold true for single class or partial class implementations ($n = 38$) in which there is no predictive relationship between experience and student learning ($t = 0.03, p = 0.97$).

**Discussion**

Of the prior meta-analytic reviews in PBL, only one (Dochy et al., 2003) examined research design. Their classification scheme relied on literature specific designs such as historical cohort control groups, and comparisons between institutions as well as pure random assignment. They found far fewer ($n = 4$) random study designs and no significant differences when they were used. About 10% of our outcomes were classified as random and by contrast were associated with about twice the effect size gains as non-equivalent comparison group designs. The random studies ranged from 1976 (Barrows & Tamblyn) to 2010 (Lin, Lu, Chung & Yang) and assuming they encapsulate the population of PBL quantitative research, one of the most influential outcomes of this study is the likely underestimate of the effect of PBL on student learning in the existing review base.

The lack of differences between content novices and content experts makes a significant contribution to the PBL literature. Some scholars have advocated for content experts (Hmelo-silver & Barrows, 2006), while others have argued expertise is not important (Swanson et al., 1990). Still others have actually evolved their opinion over time (Barrows, 1986, 1996, 1998). According to the results of this meta-analysis, content expertise is not a significant factor, which seems to coincide with the primary research. Park et al. (2007) had no uniform findings in favor of either expert or non expert-tutors. Chng et al. (2011) found the facilitation process to be more important than content expertise. This result has important ramifications for PBL practitioners. Faculty, coded as content experts, are much
more expensive than graduate students or peers, coded as content novices. If facilitation can be done equally well at a lower cost, it may help address the criticism of PBL as an expensive intervention, particularly as it scales to larger numbers of students (Stemler, 2004).

The uniform calls for tutor training appear to be vindicated in these results. While the overall impact of studies with trained tutors is similar to the combined effect of all studies, untrained tutors are associated with student gains that are similar to a lecture-based approach, the only sub-group of outcomes in the analysis that was not statistically greater than zero. Less clear is the exact form of training taking place in each study. Like the overall effect size, outcomes with trained tutors exhibit significant heterogeneity. Some scholars are calling for training to be more than a one-time professional development or workshop activity (Hendry, 2009; Zhang et al., 2011). They suggest that tutors have follow-up sessions with their trainer, are observed by a trainer, and engage in feedback from their students (surveys and verbal discussions) to grow and develop as a tutor.

Tutor experience is a much more complicated picture. At first glance, there appears to be a modest negative relationship between years of experience as a tutor and student learning. Upon further examination, that relationship is exclusive to implementations of PBL that go beyond a single class. In both cases this finding directly contradicts wider literature about teaching experience (Hedges et al., 1994; Murnane & Phillips, 1981). When examining single or partial class interventions, there is no relationship between student learning and tutor experience, which parallels findings from a primary research PBL study (Park et al., 2007). The “role” of the tutor is implicated in tutor experience, knowledge or years experience (no positive learning outcomes), while specific training is a guiding factor for student learning. Given the nature of the data in the meta-analysis, causes for these findings are purely speculative. It is possible that over time, tutors in a program experience fatigue, or that program level changes are the underlying cause. It is possible that preparing for a new method or new course material means a heightened state of activity on the part of the tutor.

There are limitations in this work. Due to a lack of information in some studies, not all outcomes were included in each analysis. This limitation is tempered by the fact that each subset of outcomes had a combined effect that was similar to the overall estimate for every outcome. Finally, as a meta-analysis, the review covers only quantitative studies and excludes a rich qualitative PBL literature.

Future work remains. It is clear as a result of this analysis that a great deal of variability persists. We join other calls for a clear language to describe what precisely is meant by content expertise (Park et al., 2007) and thus facilitate the succinct reporting of tutor expertise. Given the space constraints for journal articles the field also needs a brief language to characterize the training that tutors undergo, similar to the taxonomy of various PBL implementations (Barrows, 1986) or typology of problem types (Jonassen, 2000) that
are already available. There is an extensive literature regarding development of faculty in their role of teaching for higher education as a whole. Review work (Stes, Min-Leliveld, Gijbels, & Van Petegem, 2010) suggests some best practices that may well inform faculty development specific to PBL. For example, increases in faculty learning, faculty behavior, student learning and institutional change is coincides with training extended over time as compared with single contact formats. This finding matches best practices in technology teacher professional development (Lawless & Pelligrino, 2007). In addition to best practices, this literature might inform attempts to characterize the features of tutor training. Finally, a meta-synthesis that includes qualitative literature on PBL tutors may clarify results in this meta-analysis in addition to contributing unique findings.

In terms of scholarly significance, it appears that the pursuit of ecological validity, while admirable in the context of PBL, may be masking some of the impacts of the intervention. The inverse relationship identified between experience and student learning should be investigated in primary research and, if replicated, scholars should examine potential causes. Practitioners can take heart and save money by pursuing the use of peer level tutors. Effort should be devoted to training tutors, but the field might do well in revisiting and extending scholarship about the most effective means (Moust et al., 1990) of preparation.

References


Barrows, H. S. (2002). Is it truly possible to have such a thing as dPBL? Distance Education, 23(1), 119–122. http://dx.doi.org/10.1080/01587910220124026


Kalaian, H. A., Mullan, P. B., & Kasim, R. M. (1999). What can studies of problem-based learning tell us? Synthesizing and modeling PBL effects on National Board of Medical Examination


**Heather Leary** is a research associate in the Institute of Cognitive Science at the University of Colorado at Boulder. She earned her doctoral degree in the Department of Instructional Technology and Learning Sciences at Utah State University in 2012. Her dissertation was a meta-analysis focused on self-directed learning in problem-based learning. Leary’s research interests include problem-based learning, teacher professional development, teacher assessment, open education, 21st-century skills, meta-analysis, and machine learning algorithms. Correspondence regarding this article should be directed to Heather Leary at 594 UCB, Boulder, Colorado, 80309; email: heatherleary@gmail.com.

**Andrew Walker** is an associate professor in the Department of Instructional Technology and Learning Sciences at Utah State University. His research interests include collaborative information filtering for educational resources, technology profession development for in-service teachers, meta-analysis, and problem-based learning. He has collaborated on several mixed methods research projects but his own methodological expertise is quantitative.

**Brett E. Shelton** is an associate professor in the Department of Instructional Technology and Learning Sciences at Utah State University. He uses a variety of mixed-method research approaches to study vision, perception, cognition, and the design and assessment of innovative technologies for learning. Other interests include immersive and interactive learning environments, data visualizations, open education, instructional simulations and educational gaming. He now directs the IDIAS Institute, which has projects based on the development of hand-held applications for learning, as well as virtual world training applications that use unique design attributes to facilitate after action review and assessments.

**M. Harrison Fitt** is an instructional designer at Western Governors University. She earned her doctoral degree in the Department of Instructional Technology and Learning Sciences at Utah State University in 2011. Her dissertation centers around doctoral writing, with an emphasis on the doctoral dissertation literature review. Her research interests include graduate education, the adult learner, and factors surrounding the successful implementation of problem-based learning in online learning environments.
Appendix

Manuscripts Included in the Analysis


**volume 7, no. 1 (Spring 2013)**

