Citation: Stump, G., Li, N., Kang, S., Yaghmourian, D., Xu, D., Adams, J., . . . Chi, M. T. H. (2018). Coding dosage of teachers' implementation of activities using ICAP: A video analysis. In E. Manalo, Y. Uesaka, & C. A. Chinn (Eds.), Promoting spontaneous use of learning and reasoning strategies: Theory, research, and practice for effective transfer (pp. 211-225). New York, NY: Routledge.

13

Coding Dosage of Teachers' Implementation of Activities Using ICAP: A video analysis

Glenda S. Stump, Na Li, Seokmin Kang, David Yaghmourian, Dongchen Xu, Joshua Adams, Katherine L. McEldoon, Matthew Lancaster, and Michelene T. H. Chi

Introduction

The ultimate goal of professional development is to facilitate teachers' spontaneous transfer of knowledge gained from education research to their teaching practice in order to improve student outcomes. Researchers in the field of teacher professional development agree that professional education should situate learning opportunities for teachers in the context of their work (e.g., Ball & Cohen, 1999; Kazemi & Hubbard, 2008) and that a professional development "curriculum" should be grounded in the tasks, questions, and problems of practice (Ball & Cohen, 1999, p. 20). Our intervention, instructing in-service teachers about the ICAP (Interactive-Constructive-Active-Passive) framework and its implementation, afforded them the opportunity to apply a theoretical framework in the context of their everyday activities. The instruction was grounded in the tasks that teachers perform on a frequent basis—designing ways to engage their students in learning. Our intervention was an effort to address the ubiquitous question of how student learning can be improved via teachers' instantiation of activities that elicit increased student engagement. In this chapter, we will report our methodology and results from a video analysis conducted to assess teachers' transfer of their learning about the ICAP framework from this professional development activity to their implementation of ICAP activities in the classroom.

The ICAP framework

The ICAP framework (Chi, 2009) links students' overt engagement behaviors to cognitive processes of knowledge change that facilitate learning. It not only explains discrepant findings in education research, but also prescribes characteristics that must be present for optimal learning to occur. The framework specifies four modes of behavioral engagement: passive, active, constructive, and interactive. When students are engaged in passive behaviors, such as listening to a lecture, or silently reading a passage of text, they are paying attention, or receiving information. This allows for "storing" of information without embedding it in a relevant schema, which eventually may detract from its retrieval in dissimilar contexts. Active engagement includes some type of motoric involvement in which students are manipulating the material to be learned in some way, such as highlighting or underlining a passage, matching a word to its definition, or solving a problem using a given formula. This manipulating behavior leads to the activation of prior knowledge and "integration" of the given information for later retrieval. When students engage in constructive behaviors, they generate some type of output containing new information that goes beyond what was given, or engage in the knowledge change process of "inferring." The inferences generated during this level of engagement serve to enrich relevant schema and better prepare students to transfer their understanding to novel situations. The fourth mode of engagement in the framework is the interactive level, in which students engage in the constructive process of inferring with another individual, or the knowledge change process of "co-inferring." During the exchange, they generate new insights together in a give-and-take manner.

The ICAP hypothesis posits that the four modes of engagement—passive, active, constructive, and interactive—are organized in a hierarchical manner and lead to differential effects on student learning. Behaviors enacted during the highest level of engagement (interactive) subsume those enacted in levels below it. For example, a student engaged at an interactive level is also generating information, which is constructive; when generating information, that student is also engaged in a motoric behavior, which is active; and when engaged in a motoric behavior, he or she is also paying attention, which is passive. As each level of engagement from passive to interactive leads to deeper processing of information, the ICAP hypothesis states that student learning during interactive activities will be greater than learning during constructive activities; learning during constructive activities will be greater than learning during active activities; and learning during active activities will be greater than learning during passive activities.

The ICAP hypothesis has been supported when multiple studies were reinterpreted using the ICAP framework (see Chi, 2009; Chi & Wylie, 2014; Fonseca & Chi, 2011 for a complete review). Several studies have shown results as predicted by the hypothesis in lab and naturalistic settings in which activities were implemented and controlled by the researchers (Menekse, Stump, Krause, & Chi, 2013). The current project was an effort to scale implementation of the ICAP framework to a larger sample of teachers than in previous implementations.

Project design

Successful implementation of ICAP in this study meant that teachers were able to transfer their learning about ICAP to their subsequent design and implementation of lessons. This study was part of a larger project designed to implement the ICAP framework in secondary and post-secondary settings. In the final year of the project, information about the framework was delivered as a professional development activity via an online module (the ICAP module) to in-service junior high and high school teachers. Teachers completed the module prior to the start of the school year.

The ICAP module consisted of a pretest, four units of instruction, and a posttest. It began with a general introduction about student engagement followed by an overview of module content. The teachers then completed a 68-question pretest before moving on to the first unit of instruction. The pretest took approximately 40 minutes to complete.

The first unit contained a description of the ICAP hypothesis along with principles of the ICAP framework. Twenty-one comprehension questions were embedded within the online material. Teachers were required to respond to the embedded questions before they could continue with the next activity. For questions that could be easily scored (e.g., multiple choice), teachers received immediate feedback regarding accuracy of their responses and were required to answer the item correctly before moving on to the next activity. However, the teachers' first response was recorded and used for scoring purposes. For questions that required human grading (e.g., short answer), teachers did not receive feedback, and they could move on to the next activity without ensuring their response was correct.

The second unit contained examples of activities at various ICAP levels and focused on the design of ICAP activities by providing practice with modifying or improving given instructional activities to fit criteria of a particular ICAP level. This unit contained 25 embedded comprehension questions. The third unit contained information about and practice with writing assessment items that would measure different levels of student learning from ICAP activities. There were 28 embedded comprehension questions in this unit. The fourth unit contained implementation tips as well as examples of lesson plans that were designed to align with a specific ICAP level. This unit did not contain comprehension questions. After completing the fourth unit, teachers completed a 68-question posttest that took approximately 40 minutes. Completion of the pretest, four units of the module, and posttest required approximately 2 hours.

In the months following completion of the ICAP module, teachers created lesson plans focused on teaching the same course content at different ICAP levels to two classes. Following submission of their lesson plans, teachers received feedback from members of the research team about their planned activities and student assessments. The feedback was provided to correct any misconceptions teachers had about the ICAP framework and to remind them about some specific aspects of ICAP levels related to student behaviors. The teachers revised their lesson plans as recommended by members of the research team and resubmitted them for further review.

Implementation of ICAP lessons

We considered teachers' correct implementation of their ICAP lessons in the classroom as evidence for successful transfer of their learning about the ICAP framework. In this context, transfer cannot be conceived of in the typical two-problem transfer paradigm as described by Lave (1988), in which learners first abstract the deep structure of a problem, recognize that a second problem has the same structure, and then realize that the same procedure applies to both. In our case, teachers were expected to make direct application of an explanatory concept (ICAP principles) to new instances well removed from their initial learning (Perkins & Salomon, 2012). This can be conceptualized as an instance of analogical reasoning in which the teachers mapped exemplars given in the ICAP module to their planned class activities (Gentner, 1983). Nokes (2009) explains that analogical reasoning is most commonly used in this type of situation because it is fast and efficient to the degree that the exemplar surface features and deep structure are a reasonable match to the new situation.

To evidence the degree of transfer by our 13 participating teachers, we videotaped each class in which they implemented their planned lessons. We then analyzed 65 videos for implementation of the lesson at a particular level of engagement from the teachers' perspective. Results from the video analysis helped us to determine if a class planned at a particular ICAP level by teachers (as evidenced in their lesson plans) was delivered at the intended level. We based the judgement of "delivered as intended" on the proportion of time allotted to activities at various levels of engagement (details of the coding will be presented below).

Data

Our data sources for this analysis were 65 class videos from 32 lessons along with written materials that teachers distributed to students during the lesson. Each lesson was delivered at a different ICAP level to two different classes, except for one lesson that was delivered at three different ICAP levels to three different classes. For example, one teacher developed a math lesson on ratios and delivered the lesson at an active level to one class and at an interactive level to another class. The class videos showed the entire class, which was approximately 50 minutes long for junior high students, and 90 minutes for high school students. A typical class involved a pretest; delivery of class content via lecture or readings; question and answer periods in which the teacher would ask questions of the entire class; individual activities such as writing question responses in journals, designing posters, or completing worksheets; activities in which students worked with a partner or group; and a posttest.

Activities were defined as student endeavors that occurred any time during the 50 minute class, such as note-taking during content delivery if the teacher lectured, or completing a worksheet or creating a poster after content delivery. Activities varied considerably in nature, length, and number between lessons. Some teachers designed lessons in which one activity extended over the majority of each class; other teachers initiated two or more shorter activities during a lesson. On average, the lessons contained between one and five activities, or an average of three for each class.

Procedure

Determination of boundaries in videos

Beginning and end boundaries on the video data were set at the beginning of each class (as the teacher began to give instruction or guidance related to the content of the lesson) and at the end of each class. The boundaries excluded time utilized for the pretest and posttest. If more than five minutes were devoted to transition from one portion of the class to another, e.g., teachers pairing up students, collecting worksheets, or students cleaning up, this time was also excluded. The remaining time represented between 80–100% of the total time per class. Boundaries between activities were determined to be the point at which one activity transitioned to another. These boundaries were established during phase 1 coding.

After boundaries were set, the time to be coded was parsed into one-minute segments. This allowed us to calculate the ICAP "dosage" in terms of the total number of minutes allocated for activities at a particular ICAP level. It also allowed us to check inter-rater agreement by making a minute-by-minute comparison of two rater's codes over the class period. The video coding was then conducted in two phases.

Phase 1 coding

During the first phase, we used the teachers' oral instructions for student activities, the content-related questions they asked during class, and any prompts they gave to students as cues to infer teachers' intention to implement activities at a particular ICAP level. We assigned a code to all minutes within the class boundaries using the following nomenclature: passive, active, constructive, interactive, UTC (unable to code – for the time in which an ICAP level could not be determined from teachers' oral instructions or activity descriptions), and NC (no code – for the short periods of time related to non-content instruction or time when students were forming groups, getting equipment, etc.).

Four activities were noted to be the most common—taking notes, completing worksheets, working with a peer, and responding to questions that teachers asked verbally during content delivery. Our coding rules for each of these activities are described below.

Taking notes: The ICAP level for time related to this activity was determined by how teachers orally told students to take the notes. A single activity could elicit different levels of student engagement, depending on the teacher's instructions or scaffolding during the activity. For example, if a teacher used a Power Point presentation to deliver a lecture for twenty minutes and orally told the students to take notes, those twenty minutes were coded as active because students were focusing on particular pieces of content and we assumed students were copying notes from the Power Point. Alternatively, if the teacher orally told students to take notes in their own words, or asked them to answer questions in their notes that required inferences about the content, those twenty minutes were coded as constructive because we assumed students were generating information beyond what was given.

Completing worksheets: The ICAP level for time used for this activity was determined by key words in the teachers' oral instructions. If key words suggesting manipulation of given information (such as "list the important concepts," "select the answer," or "match the items") were used, the related minutes were coded as active. If key words suggesting generation of inferences (such as "predict," "justify," "compare and contrast," "hypothesize," "generate a rule," "self-explain," "propose," "critique," or "create a table/concept map") were used, the related minutes were coded as constructive. If the teachers' verbal instructions did not explicitly reveal an ICAP level—for example, if the teacher stated, "Complete this worksheet," the related minutes were coded as UTC and recoded during phase two.

Working with a peer: The ICAP level for time related to this activity was coded as interactive if the teacher elicited at least a behaviorally interactive activity by instructing students with statements such as "talk to your partner," or "work together."

Responding to teachers' oral questions: The ICAP level for time used for this questioning activity was based on the type of question teachers asked and how many students were permitted to respond if an oral response was requested. If teachers asked questions that required students to recall a fact, such as "What makes up the coat of a virus?" and this fact had been presented during class, the related segments were coded as active because students were just recalling information that had been previously given. Key words from the ‘remember’ level of Bloom’s taxonomy (Krathwohl, 2002) such as “define,” “name,” “list,” “describe,” or “restate” were often indicative of recall questions. On the other hand, if teachers asked questions that required students to make an inference or new connection between facts or concepts, the related minutes were coded as constructive because students were often generating information that had not been presented to them. Similar to evaluating the ICAP level of teachers’ verbal instructions for worksheets, key words used when asking questions orally, such as “compare,” “contrast,” “justify,” “defend,” or “evaluate,” were considered to be indicative of questions that required students to make inferences. These keywords are from the ‘understand,’ ‘apply,’ ‘analyze,’ ‘evaluate,’ and ‘create’ levels of Bloom’s taxonomy, as tasks at those levels can require students to make inferences about given information. The exception to use of the above-mentioned words as key words was when the correct response had already been given and students merely had to recall it. For example, if teachers asked an inference question for which the answer had been previously discussed, we coded the allocated time as active because students were only required to recall the correct answer rather than making any new connections between concepts or generating any new understanding.

As mentioned earlier, the second consideration when coding time allocated to question and answer sessions was the number of students permitted to respond. If teachers asked an inference question and students answered by writing a response in their notes, the allocated time was coded constructive because all students were presumably writing responses. Similarly, if teachers told students to answer in unison to the same question, the allocated time was coded as constructive because again, all students (or at least a majority) were responding. However, if teachers told students to raise their hand to respond, and then called on only one student who raised his/her hand, the allocated time was coded as passive, because the majority of the students were not responding and may not have even contemplated a response to the question.

To summarize, during the first phase of coding, we examined the videos for teachers’ oral instructions, questions, and prompts to determine the ICAP level that should have been elicited from their presentation to the students. We defined keywords, many borrowed from Bloom’s taxonomy (Krathwohl, 2002) to code the ICAP levels of teachers’ oral questions or instructions. In addition to the codes representing ICAP levels, we applied UTC for minutes in which the ICAP level was ambiguous from teachers’ descriptions and NC for time that was not related to instruction or activities.

Phase 2 coding

The coding done in phase 1 was based strictly on the teachers’ instructions, in terms of what the teachers orally asked students to do or to answer. However, it is not accurate to code only based on what teachers asked, without also coding for what teachers presented during the course of instruction, because what is presented in instruction can change the ICAP level of a question or an activity. Thus in this second phase of coding, we examined any written materials that accompanied the lessons and reviewed the information that teachers presented during class. Evaluation of Power Point presentations and readings as well as review of class lectures helped us to identify instances in which the answers to inference questions asked during activities had already been provided to students, thus converting constructive activities to active ones because students could simply recall an answer instead of generating it. Evaluation of the worksheets also helped to resolve any ambiguous segments that we were unable to code during the first phase (the cases in which teachers did not orally indicate the type of activity contained in the worksheet), and also helped us to determine if the activities were at one ICAP level or a mixture of levels. Our process of corroborating the accuracy of ICAP levels assigned during phase 1 is described in more detail below.

When reviewing materials for minutes that were coded as passive during the first phase, we examined for any directives that would increase students' level of engagement, such as a worksheet that provided spaces for note-taking during instruction. If such directives were found, the passive code was then changed to the appropriate ICAP level. When reviewing materials for segments that were previously coded as active, we examined the instructions, accompanying questions, or problems to confirm that they only asked students to manipulate the materials in some way, such as copying, selecting the correct answer from information given in class, using a given formula to solve a problem, or underlining important concepts. When reviewing materials related to the segments that were previously coded as constructive, we examined for directives to generate information in some manner, such as connecting previous knowledge to newly learned knowledge, using prior knowledge in a novel way, inducing or inferring new information from given information, making predictions related to newly learned principles, or deducing rules from observed patterns. Key words from Bloom's taxonomy, as described in the phase 1 coding description, were also used in this phase of coding as a guide to identify questions or activities that required recall versus inference-making. As described earlier, activities designated at this level could easily be converted from constructive to active if teachers presented the information elsewhere in the lesson that students were supposed to generate. Thus the classes were closely scrutinized to ascertain that the activities truly required students to exert generative cognitive effort instead of simply recalling information that they had been given. When reviewing materials for minutes that were previously coded as interactive, we examined the instructions, accompanying questions, or problems to confirm that they were constructive (required for true co-construction), rather than active activities in which students were asked to work in pairs. This coding was more stringent than during phase 1 when, rather than examine the nature of the activity, we considered the activity interactive if students were only asked to be behaviorally interactive (work together). However, in phase 2 we still could not determine whether students were actually co-constructing, which is the most stringent measure of interactive engagement as defined by ICAP.

Phase 2 coding also allowed us to assign minutes of an ICAP level within an activity if it was comprised of more than one level. For example, if after coding, half of the questions in a worksheet turned out to be recall questions (an active activity) and the other half were inference questions (a constructive activity), we coded 50% of the total minutes allocated for the activity as active and 50% as constructive.

To summarize, during the second phase of coding, we examined any written materials that accompanied the lessons to verify that they (materials) would contribute to or elicit engagement at the same ICAP level teachers articulated to their students. If the materials did not support student engagement at the ICAP level that teachers articulated, we recoded the associated minutes to the appropriate ICAP level.

Inter-rater agreement

During phase 1 coding, we randomly selected 23% of the data—four math classes, four science classes, four language arts classes, and three foreign language classes—to determine inter-rater agreement for our coding. The classes were assigned to six trained raters, ensuring that two raters independently completed coding for each class. Initial agreement between any two raters ranged from 39% to 100%, with only two classes having agreement lower than 85%. For those classes, the raters reviewed the coding protocol and recoded the classes until agreement reached 85%. The six raters then individually coded the remainder of the classes for phase 1.

During phase 2 coding, the same individuals served as second raters for the same fifteen classes. Initial agreement between any two raters ranged from 33% to 98%; this time, eight out of the 15 classes had inter-rater agreement lower than 85%. The major source of discrepancy was determination of the proportion of questions or problems that belonged to the active or constructive category. The raters who coded the same class resolved this disagreement via discussion, and then recoded the segments, reaching an agreement of 85% for each of the eight classes. Six raters individually coded the remainder of the classes for phase 2.

Calculation of ICAP dosage and correct implementation

The final step of the video coding process was to calculate the proportion of class time teachers actually allocated to each ICAP level, based on our coding. These proportions were interpreted as the "ICAP dosage." We then determined whether teachers implemented classes as they intended by using two criteria related to ICAP dosage: 1) the plurality of time was allocated to the ICAP level stated on their lesson plan, and 2) the proportion of that time was greater than 30% of the total class time. For example, if a teacher designed a class to be constructive and allocated 45% of class time to constructive activities, 30% to active activities, and 25% to passive activities, the class was considered to be implemented correctly because the plurality of time was devoted to constructive activities and the time allocated was more than 30% of the total class time.

Results

Table 13.1 Teachers' intended versus implemented class ICAP levels

Intended ICAP level (total classes intended at this level)	Implemented ICAP level using dosage criteria			
	Passive	Active	Constructive	Interactive
Passive (3)	1.5 (50%)	1.5		
Active (19)	2	17 (89%)		
Constructive (20)	3.5	11	5.5 (28%)	
Interactive (23)	4	13	1	5 (22%)

Numbers in bold denote classes that were implemented as intended; percentages were calculated from classes delivered as intended/total number of classes intended at this level. The classes in which time was allocated equally between activities were not considered to be implemented correctly.

When comparing the teachers' intended ICAP level for their classes (as stated in their lesson plans) with the ICAP dosage as determined by video coding, we found that overall, 37 of the 65 classes (57%) were not delivered at the ICAP level that the teacher intended, or what we will refer to as incorrectly implemented. That is, 57% of the classes did not allocate a plurality of the total class time (and greater than 30%) to the intended activity. Table 13.1 shows the number of classes that were actually implemented at each ICAP level (according to our two dosage criteria), out of the total number of classes intended at a given ICAP level. If the passive classes are ignored due to their low number, Table 13.1 shows that the highest proportion of classes implemented correctly was the active classes at 89%. Both the constructive and interactive classes had much lower rates of correct implementation, at around 25% on average.

We examined other data for factors that may have predicted the teachers' ability to effectively implement ICAP in their classes—teachers' responses to 16 posttest items related to application of ICAP principles from the online ICAP module. These questions asked teachers to explain their understanding of the ICAP framework and apply that understanding to given scenarios. There were four 'Improve' questions that asked teachers to apply their knowledge of ICAP to improve an activity to a higher level, and 12 'Explain' questions that asked them to define ICAP levels and describe activities that could be implemented at each level, or describe why an activity was classified at a particular ICAP level.

To see whether teachers could implement activities with greater fidelity to the ICAP framework if they understood it well, we examined scores from the three highest and three lowest scoring teachers on the 16 posttest items related to the four levels of ICAP. Although there was not a large enough sample to detect significant differences, the descriptive statistics shown in Table 13.2 reveal the differences in total score on the posttest ICAP items between low-scoring teachers (1–3) and high-scoring teachers (4–6). Table 13.2 also shows differences between the higher and lower-scoring teachers' performance when the 16 items were further broken down into the 'Improve' and 'Explain' categories. In addition, the teachers who scored highest on these posttest items also scored significantly higher on more difficult and discriminating posttest items from the entire test than teacherswith lower scores, $t(4) = -3.674$, $p = .021$, demonstrating that, in general, they had a better understanding of ICAP.

Table 13.2 Teachers' scores for items on ICAP module post-test and subsequent implementation of lessons at correct ICAP level

Teacher	*ICAP application question total score (N = 16 items)*	*ICAP application-'improve' question score (N = 4 items)*	*ICAP application-'explain' question score (N = 12 items)*	*Performance on difficult and high discrimination questions (N = 7 items)*	*Proportion of classes taught at intended ICAP level*
1	62.5%	75%	45%	43%	.20
2	62.5%	75%	45%	57%	.67
3	62.5%	75%	45%	29%	.50
4	87.5%	100%	72%	71%	1.00
5	87.5%	100%	72%	86%	.50
6	93.7%	100%	81%	100%	.75

N = 12 teachers who took the posttest; data shown above represents teachers with lowest and highest scores; 'Application' questions were comprised of 'Improve' and 'Explain' questions.

The last column of Table 13.2 shows that higher-scoring teachers tended to deliver a higher proportion of classes at the intended ICAP level than teachers with lower scores. We calculated the proportion of classes that teachers implemented correctly by dividing the number of correct implementations by the total number of classes they planned. This suggests that teachers could transfer their understanding to correct implementation if they understood the ICAP module well.

Discussion

Overall, teachers did not appear to satisfactorily transfer their learning about the ICAP framework when implementing the lesson plans in their classrooms. When evaluating implementation of lessons in the classroom, evidence for transfer was considered to be a match between teachers' intended ICAP level for the class and the level they actually implemented based on our detailed analysis of time allocated for ICAP activities, materials prepared, and content presented. Our results showed that only 43% of the classes were implemented at the teachers' intended ICAP level, with a much higher percentage of classes intended as active implemented correctly (89%) than those intended as either constructive (25%) or interactive (22%).

From a cognitive perspective, one hypothesized reason for the teachers' failure to transfer is that teachers did not learn about ICAP deeply enough (Bransford & Schwartz, 1999; Bransford, Brown, and Cocking, 1999; Chi & VanLehn, 2012). This assumption is supported by our data here—that is, the three teachers who learned material about the ICAP framework more deeply were also more likely to transfer their understanding to practice, as evidenced by our comparison of correct implementation proportions from teachers who scored highest and lowest on 16 posttest application items.

It is interesting to consider how teachers' transfer of ICAP knowledge to practice is the same or different from traditional approaches to assessing transfer. As mentioned earlier, in the traditional transfer literature, transfer is typically conceived of in the context of the two-problem transfer paradigm (Chi & VanLehn, 2012; Lave, 1988), in which learners are expected to first abstract or induce the deep structure or principles underlying the first problem and then apply it to the second problem. In our case, teachers were given the deep structure of the ICAP framework—the cognitive processes associated with knowledge change at four levels of engagement—and exemplar activities for each level. They were then expected to make direct application of this information to their planned class activities, which seems to be an easier transfer than requiring them to first abstract the deep structure. Yet even when the deep structure was presented to them in the ICAP module, in terms of the principles and operational definitions of the ICAP framework, they did not apply ICAP correctly in multiple lesson contexts.

From an alternate viewpoint, researchers who have studied transfer suggest that multiple measures of transfer should be utilized to assess multiple dimensions of transfer (Nokes, 2009). Thus, taking an Actor Oriented Transfer perspective (Lobato, 2012), which removes the researchers' imposed criteria for successful transfer (in our case, "correctly implemented" meant that time allocated for the desired ICAP level was a plurality of the total class time, and greater than 30%), we examined teachers' implementation to determine if there was any transfer at all. From this different perspective, we noted that there was indeed some transfer by a majority of the teachers. As shown earlier in Table 13.1, the greatest number of classes teachers implemented incorrectly was those intended to be either constructive or interactive. However, in 13 of the 15 constructive classes that were incorrectly implemented (that is, they were implemented as either passive or active), teachers allocated an average of 8.8 minutes or 22% of their total class time to constructive activities. Similarly, in 14 of the 18 interactive classes that were incorrectly implemented, teachers allocated an average of 12.3 minutes or 19% of their total class time to interactive activities. For these classes, the percentage of class time allocated to constructive or interactive activities simply did not meet the plurality or 30% criteria we imposed and thus, by our rules, they were considered unsuccessful cases of transfer. Nevertheless, these teachers did allocate more time to constructive and interactive activities than teachers of the two active and two passive classes that were incorrectly implemented. In those four classes, teachers did not allocate any time to constructive or interactive activities at all.

For the intended constructive or interactive ICAP level classes, the lower amount of time allocated to activities at the intended ICAP level suggested the possibility that teachers considered the ICAP manipulation as an addendum to their former routines, implementing ICAP activities in addition to their usual class activity rather than considering the level of student engagement during the entire class period. To investigate this idea further, we reviewed teachers' initial plans for 22 lessons in which one class was planned at the active level and a comparison class was planned at either a constructive or interactive level. In 17 of the 22 lessons, teachers planned the same activities for both the active class and the constructive or interactive comparison class, adding only an additional activity or two at the intended ICAP level to the comparison class. This finding supports the idea that teachers may have delivered their class in the usual way and then added ICAP activities as an enhancement. This practice is certainly not detrimental to student learning if the usual class routine is highly engaging. However, if teachers continue a past practice of implementing a passive activity, such as lecture, for a majority of class time and then attempt to improve their class by adding a constructive or interactive activity for the last 10 minutes, they may not see any changes in students' learning outcomes due to their use of the ICAP framework.

It is important to note here that we did not tell teachers that all activities within a given class must meet the same intended ICAP level, or that a plurality of class time should be spent at the intended level, so it is not reasonable to hold teachers accountable for this information. We did find however, that the three teachers who scored highest on the posttest application items not only implemented a greater proportion of classes correctly (as shown in Table 13.1), but in general they allocated a greater proportion of class time to intended ICAP level activities than the three low-scoring teachers, as shown in Table 13.3; thus again supporting our premise that teachers who understood ICAP more deeply were more likely to successfully transfer theory into practice by implementing activities that elicited a deeper level of student engagement.

Table 13.3 Class time allocated by teachers scoring high and low on post-test application items

Intended level	*Constructive classes*		*Interactive classes*	
	Average # of minutes allocated to *constructive* activities	Average % of class time allocated to *constructive* activities	Average # of minutes allocated to *interactive* activities	Average % of class time allocated to *interactive* activities
3 low-scoring teachers	19.00	25%	23.57	28%
3 high-scoring teachers	26.75	50%	32.75	69%

12 teachers took the posttest; data shown above represents teachers with lowest and highest scores.

Conclusion and recommendations

Taken together, our findings suggest several areas in which instruction about the ICAP framework can be improved to promote teachers' transfer of knowledge to their implementation in the classroom. Our findings when comparing the teachers who scored highest and lowest on the module posttest (as shown in Table 13.3) point clearly to the fact that teachers did not understand the ICAP framework deeply enough to implement activities that elicited higher levels of student engagement for any length of time during their classes. Therefore, we recommend that additional instructional strategies be utilized in subsequent versions of the module to increase deep learning of the ICAP framework.

In particular, teachers themselves need to engage more deeply in identifying the principles underlying the framework. Prior to being given information about ICAP, teachers could be asked to analyze contrasting cases (Schwartz & Bransford, 1998) of ICAP implementation, using video to show correct and incorrect applications of the framework. Comparing examples and nonexamples of a concept to be learned assists the learner with three important transfer mechanisms: constructing generalizations, making connections between generalizations and examples, and making appropriate connections between different examples as part of analogical reasoning (Engle, Lam, Meyer, & Nix, 2012, p. 216). In this case, teachers' analysis of correct versus incorrect implementations of ICAP activities should prepare them to better understand the deep structure of the ICAP framework when it is presented. Additionally, the use of video as the media of instruction for the cases can provide a more engaging experience than text-based scenarios, and is recommended as a methodology that captures the complexity of classroom practice (Borko, Jacobs, Eiteljorg, & Pittman, 2008; Santagata, 2009).

Finally, our results suggest that an additional important point of emphasis for teachers new to ICAP is that utilizing ICAP is not an "add on," but rather a more holistic change in pedagogical approach. Instead of manipulating only one or two activities within a class to a higher level of engagement, teachers must consider students' level of engagement throughout the entire lesson, and adjust all of the instruction to engage students at the highest level that is reasonable. In our study, we explicitly asked teachers to transfer their knowledge of ICAP to their classroom practice. However, had we reframed the use of ICAP as a more holistic approach, it may have facilitated more successful spontaneous transfer of the ICAP framework from theory into practice.

Acknowledgments

The authors are grateful for support from the Institute of Education Sciences (Awards #R305A110090 and #R305A150432) to the last author for the projects "Developing Guidelines for Optimizing Levels of Students' Overt Engagement Activities," and "Developing and Revising Instructional Activities to Optimize Cognitive Engagement." They would also like to thank Emily Bogusch and Christiana Bruchok for their helpful comments on an earlier draft of this manuscript.

References

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), Teaching as the learning profession (pp. 3–31). San Francisco, CA: Jossey-Bass.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. Teaching and Teacher Education, 24(2), 417–436.

Bransford, J., Brown, A. L., & Cocking, R. (Eds.). (1999). Learning and transfer. In How people learn: Brain, mind, experience, and school (pp. 39–66). Washington, DC: National Academy Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. Review of Research in Education, 24(1), 61–100.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73–105.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. Educational Psychologist, 47(3), 177–188.

Chi, M. T. H., & Wylie, R. (2014). ICAP: A hypothesis of differentiated learning effectiveness for four modes of engagement activities. Educational Psychologist, 49(4), 1–25.

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. Educational Psychologist, 47(3), 215–231.

Fonseca, B., & Chi, M. T. H. (2011). The self-explanation effect: A constructive learning activity. In R. Mayer & P. Alexander (Eds.), The handbook of research on learning and instruction (pp. 270–321). New York, NY: Routledge Press.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7(2), 155–170.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts. Journal of Teacher Education, 59(5), 428–441.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into Practice, 41(4), 212–218.

Lave, J. (1988). Cognition in practice: Mind, mathematics, and culture in everyday life. New York, NY: Cambridge University Press.

Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. Educational Psychologist, 47(3), 232–247.

Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. H. (2013). Implementation of differentiated active-constructive-interactive activities in an engineering classroom. Journal of Engineering Education, 102, 346–347.

Nokes, T. J. (2009). Mechanisms of knowledge transfer. Thinking & Reasoning, 15(1), 1–36.

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. Educational Psychologist, 47(3), 248–258.

Santagata, R. (2009). Designing video-based professional development for mathematics teachers in low-performing schools. Journal of Teacher Education, 60(1), 38–51.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. Cognition and Instruction, 16(4), 475–522.