Abstract

Three hundred and forty one children (M_{age} = 9,0 years) engaged in a series of science tasks in collaborative, same-sex pairs or did not interact. All children who collaborated on the science tasks advanced in basic level understanding of the relevant task (motion down an incline). However, only boys advanced in their conceptual understanding at a three week post-test. Discussion of concepts and procedural aspects of the task led to conceptual development for boys but not girls. Gender differences in behavioral style did not influence learning. Results are discussed in terms of the links between gender and engagement in conversations, and how gender differences in collaboration may relate to differences in participation in science.

Keywords: collaboration, concept, gender, peer, science learning

Peer collaboration facilitates learning in science (Phelps & Damon, 1989). However, relatively little research has explored whether this learning is influenced by gender. The lack of research into gender and collaborative learning is surprising because gender has a profound influence on children's interactions and conversation dynamics (Leaper & Smith, 2004) and peer collaboration is frequently used in classroom teaching.

The present study examines whether gender influences learning of concepts in a collaborative science task. Questions about when and how children grasp scientific concepts have been a focus of research for developmental psychologists for many years (e.g., Piaget, 1967). By examining collaborative learning we can better understand how conceptual and social factors connect with developmental change. A central objective of the present study was to establish whether any gender differences in the style and content of peer conversations are associated with gender differences in learning, or with different routes to learning.

Gendered communication dynamics in children's interactions reinforce and sustain gender differences and may have longer term consequences for social relationships (Di Donato, Martin, & England, 2014). However, by middle childhood children can anticipate and "work around" the influence of gender on interactions. For instance, a 7-year-old girl can anticipate that a boy will seek to dominate in discussion and as a consequence use a more circumspect approach to persuade him of the merits of her position (Leman, Ahmed, & Ozarow, 2005). By the end of middle childhood, children are quite adept social actors who interact in gendered ways and can respond to a partner's gender. Children at this age can also use peer conversation to make decisions together and to focus on appropriately using varied forms of information in conversation (Gummerum, Leman, & Hollins, 2014). Understanding how gendered conversation styles affect learning is particularly important in the domain of science where

stereotypes of success, ability, and aptitude that are formed early in development can continue to shape men's and women's participation and success in science into adulthood (e.g., Good, Rattran, & Dweck, 2012). A competitive and conflictual atmosphere in discussion in science has also been cited as an element that may discourage women from pursuing careers in the area (see Ceci & Williams, 2007).

Conceptual understanding is an important element in science learning (van Boxtel, van der Linden, & Kanselaar, 2000). Conceptual understanding is distinguished from procedural, basic (or foundational) concepts (Howe, 2009). Basic scientific knowledge relates to single components that can be used to describe a situation or event. For instance, there are many different factors relevant to how far a trolley will travel down an incline, including the gradient of the slope, the weight and position of the trolley, and the characteristics of the surface (e.g., high or low friction). A basic level understanding of the physical processes involved would identify a single element as relevant to determining the distance likely to be travelled by the trolley. In contrast, a conceptual level of understanding requires that children appreciate and successfully coordinate the relations among different factors or variables (see Kuhn & Crowell, 2011). In this respect, conceptual understanding requires the ability to integrate two or more elements to understand all the forces and factors involved (covariation). Studies in science learning contexts suggest that an appreciation of covariation develops systematically from around 3rd grade (9 years) into adulthood (e.g., Kuhn, 1989; Zimmerman, 2000, 2007); although, even in adulthood there can be a failure to grasp all aspects of covariation. Research has examined how children acquire basic and conceptual level knowledge in formal classroom interactions (e.g., Mercer, et al., 2004), but to date no study

has examined whether boys' and girls' collaborations are more or less effective in promoting these different forms of understanding.

In the present study we examined gender differences in the basic and conceptual language used by children in same sex peer collaboration. Nine year old children completed assessments of basic level and conceptual level scientific knowledge relating to understanding motion down an incline. We compared the changes, pre- to post-test, of children who had collaborated on related tasks for three sessions, ten minutes each, over the course of a week with children who did not engage in any interaction or science learning. Additionally, we analysed characteristics of interactions and the content of conversations of children, comparing the interaction dynamics and contents of boys' and girls' conversations. We chose this age group because it is an important phase in the development of conceptual knowledge and related factors (such as covariation, see again Kuhn, 1989).

We proposed three hypotheses. Firstly, we expected that children who engaged in interaction on science tasks would show improvement in performance at post-test. We also sought to establish if gender differences in peer collaborations would lead to differences in the acquisition of basic versus conceptual knowledge. Secondly, we predicted that there would be no gender differences in the content (i.e., what children discussed relating to the science task) of children's conversations: both boys and girls would have access to conceptual and basic forms of thinking about science, and both should be able to reproduce these in conversation. Thirdly, we expected that girl pairs would be more affiliating and boy pairs more assertive in interaction (Leaper & Smith, 2004). Finally, we examined the relations among conversation content, behavioral dynamics, gender, and learning. Previous studies have found that features of conversation (such as content and dynamics) are often not reliable predictors of learning

and that learning gains may be attributable to individual reflection on a topic rather than specific features of the content or style of discussion (Howe, McWilliam, & Cross, 2005; Howe, Taylor Tavares, & Devine, 2014). One possible explanation for this is that the features of interaction could be differentially effective for boys and girls, so we also explored if there are different pathways to learning through collaboration for boys and girls.

Method

Participants

Children (N=341) were recruited from five schools in a metropolitan area of England, United Kingdom across several months in the middle of the school year (November through February 2012-2013). Children were in their fifth or sixth year of formal schooling, mean age 9 years 0 months (184 boys, 157 were girls). The sample was drawn from an area of high ethnic diversity (where 35% of the population is from an white, European ethnic majority group). The participants were of homogeneous socioeconomic status (around 98% of children came from families with incomes in the national lower quartile). Previous research has demonstrated that ethnicity has an influence on interaction dynamics and that ethnicity can also interact with gender in interactions (Leman & Lam, 2008). Therefore we included, in the present analysis, only children from the majority (white European) and principal minority (South Asian) ethnic groups. South Asian children include those of Indian, Pakistani, Sri Lankan and Bangladeshi descent.

Design

All children completed the pre-test of basic and conceptual science knowledge. Children were then placed into one of two groups: collaboration or no collaboration. Allocation to these groups was at random, with no specific selection criteria (e.g., sex, ethnicity). Three

weeks following the interactions phase, children completed a post-test of basic and conceptual science knowledge (see below for further details of the test materials).

Procedure

On the first day, children completed a pencil and paper science test in the classroom (the pre-test), individually at their desks. The following day, a subgroup of children was put into pairs to work on the science tasks ($N=160$). We used a blocking design to assign pairs such that there were roughly equal numbers of all-white, all-Asian, and cross ethnic (Asian-white) pairings with similar gender distributions in each pair type.; European pairs (9 female,7 male), South Asian pairs (21 female, 19 male), cross ethnic pairs (13 female, 11 male). All pairs were same sex.

Children were permitted to explore the apparatus and perform the tasks with the same partner each day for 10-15 minutes. Those who did not participate in the interaction phase formed the control group and engaged in normal classroom activities while those in interaction pairs left the classroom for the period of interaction. Three weeks after the interaction phase, all children completed an individual post-test.

Materials

Science knowledge

Children completed pre-and post-tests of science quiz of 16 questions, four of which tested basic knowledge and 12 tested conceptual knowledge of how the four variables of gradient, weight, starting point, and surface material could predict the extent of motion down an incline.

Interaction tasks

The aim of the science interaction tasks was to increase children's understanding of how four variables of gradient, weight, starting point and surface material explained motion down an incline.

On the first day children worked together on two computer based tasks, each examining the impact of two of the aforementioned variables on the motion of a truck. On day two children were presented with a ramp and a truck which they could experiment with. On the final day children watched a short video which showed a scientist who had invented new skis and then answered questions testing whether they could transfer the knowledge of the truck tasks to a different context. Pairs were given five minutes to experiment with each task and to complete a worksheet which asked about the impact that each variable had individually and also how the two variables interacted to influence how far the truck travelled. Children were also asked to give reasons for their answers.

Children in the control group did not interact with a partner, apparatus, or complete tasks, but did complete the science pre- and post-test quizzes at the same times as the other children.

Conversation measures

Conversation content

In order to establish children's use of different categories of explanation (conversation content) two coders each reviewed video recordings of all the science conversations (each coder viewed half of the videos in the first instance) in order to identify how many times each participant made reference to one of four different categories of explanation or justification for their judgments. Table 1 gives descriptions of the measures of conversation content. A count was made each time an instance of a category was made. The categories were developed from

a prior thematic analysis of conversations. When two categories were combined in one utterance, a code was given for each category.

Behavioral style

We established the levels of assertion and affiliation in conversation using Leaper's (1991) Psychosocial Processes Rating Scheme (PPRS). Video recorded interactions were coded by two judges who were blind to the study hypotheses who separately rated each participant's behavior every 30 seconds on seven-point scales where 1 represented the lowest levels of assertion and affiliation, and 7 the highest levels. Assertion includes verbal and nonverbal behavior from unassertive (e.g., sitting passively) to assertive behavior (e.g., aggression). Affiliation ratings ranged from unaffiliative (e.g., ignoring another child) to interdependent (e.g., cooperation).

Reliability

The two coders rated 12 (15%) of the conversations that had previously been coded by the other. Kappas indicated excellent agreement for content categories (from κ=0.81 to κ=1.00 and behavioral style (assertion κ=0.78; affiliation κ=.88).

Results

Science learning

There were no differences between boys and girls in terms of basic and conceptual level knowledge at pre-test. Separate repeated measures ANOVAs were performed comparing pre- to post-test scores by gender (male or female) and task (science collaboration versus control) on first basic and then conceptual knowledge. Table 2 reports pre- and post-test means for the basic and conceptual science knowledge of boys and girls by condition.

Basic knowledge

Children who participated in science tasks improved more than those in the control group on the basic level, $F(1,277)=12.90$, $p<.001$, $\eta_p^2=.044$. Follow-up, related t tests (Bonferroni corrected) examined pre- to post-test changes in each condition. These indicated significant basic level change in science, $t(141)=6.98$, $p<.001$, Cohen's $d=1.18$, but no change in the control condition, $t(138)=0.58$, $p=.28$, $d=.10$.

There were no differences between boys and girls in the science condition in terms of basic science knowledge, $F(1, 140)=1.73$, $p=.190$.

Conceptual knowledge

A repeated measures ANOVA indicated a significant interaction between condition and the repeated measure (i.e., conceptual knowledge at pre- versus post-test), $F(1,266)=10.27$, $p=.002$, $\eta_p^2=.037$. Children who collaborated on the science tasks showed greater improvements over time compared to the children who did not interact. Follow-up, related t tests indicated a significant change between pre- and post-test in the science condition, $t(134)=2.61$, $p=.01$, $d=.45$, but not in the control condition, $t(134)=1.58$, $p=.12$, $d=.27$, see again Table 2. The ANOVA also indicated a marginally significant interaction between gender, task and the repeated measure, $F(1,266)=3.053$, $p=.082$ $\eta_p^2=.011$. We therefore conducted separate 2 (gender) x 2 (pre- to post-test) repeated measures ANOVAs on participants in each condition separately. There was a significant interaction only in the science condition, $F(1, 133)=5.76$, $p=.018$, $\eta_p^2=.042$. Bonferroni corrected related t tests indicated boys showed improvement on conceptual questions, $t(56)=3.24$, $p=.002$, $d=.87$, whereas girls did not, $t(77)=.502$, $p=.617$, $d=.11$.

Gender differences in conversations

Independent *t* tests were conducted to examine gender differences in conversational content (see Table 3). These analyses revealed only one significant gender difference in conversation content with boys using more conceptual explanations than girls. In terms of behavioral style, boys' pairs were significantly more assertive than girls' pairs, and girls' pairs were marginally more affiliative than boys' pairs.

Conversation predictors of science learning

In order to examine predictors of learning, two hierarchical linear regression analyses were performed, first on basic level science knowledge and then on conceptual level knowledge. In both analyses, the first block pre-test score was entered as a predictor variable. In the second block gender was entered. In the third block conversation measures (procedural, conceptual, basic, applied, and social) and behavioral measures (assertion and affiliation) were entered. And, in the fourth block, interactive predictors with gender by each conversation measure were entered. Summary statistics for the two regression analyses are given in Table 4 (see online supplemental materials for correlation tables).

For *basic level knowledge*, the inclusion of the predictor performance at pre-test significantly improved the model. Block 2, 3, and 4 showed no significant improvement of the model. The final model was significant overall, $F(16, 122) = 1.79$, $p = .040$, with a total of 8.3% (adjusted R^2) of the variance explained in the post-test basic knowledge scores. Unsurprisingly, higher scores at the pre-test predicted higher scores at post-test. More applied talk was predictive of lower post-test scores and this was moderated by gender, suggesting that the relation differed for boys and girls. After controlling for all predictors, the relation between applied conversation and the post-test scores for boys was negative, r (df = 49) = -.38, $p = .006$, while for girls there was no significant direction, r (df = 72) = .05, $p = .649$;

these two relations differed significantly, $z = -2.44$, $p = .015$. However, applied talk was used very infrequently by boys, and even more so by girls, so it would be unwise to draw strong conclusions based on these findings. Further, more frequent social conversation was predictive of the basic knowledge post-test scores, and this too was moderated by gender. After controlling for all predictors, the relation between social conversation and the post-test scores for boys was more positive, r (df = 49) = .20, $p = .158$, than for girls, r (df = 72) = -.18, $p = .137$; these two relations differ significantly, $z = 2.04$, $p = .041$.

For *conceptual level knowledge*, the inclusion of the predictor performance at pre-test also significantly improved the model. Block 2, 3, and 4 showed no significant improvement of the model. The final model was significant overall, $F(16, 114) = 1.91$, $p = .026$, with a total of 10.1% (adjusted R^2) of the variance explained in the post-test conceptual knowledge scores. Again, higher scores at the pre-test predicted higher scores at post-test. More procedural talk was predictive of higher post-test scores. This was moderated by gender, and after controlling for all predictors, the relation between procedural conversation and the post-test scores for boys was positive, r (df = 45) = .32, $p = .030$, while for girls there was no significant direction, r (df = 68) = -.19, $p = .120$; these two relations differ significantly, $z = 2.66$, $p = .008$. Further, more frequent conceptual conversation was predictive of higher conceptual knowledge post-test scores, and this too was moderated by gender. After controlling for all predictors, the relation between conceptual conversation and the post-test scores was more positive for boys, r (df = 45) = .25, $p = .097$, than for girls, r (df = 68) = -.16, $p = .194$; these two relations differed significantly, $z = 2.09$, $p = .037$.

Discussion

We hypothesized that children who engaged in a series of collaborations on a science task (understanding motion down an incline) would show greater advances in scientific knowledge than children who did not engage in any interaction. This first hypothesis was confirmed and fits the vast majority of findings in the area that demonstrate that peer collaboration promotes science learning (e.g., Howe, 2009). The present study provides important new insights into how the effectiveness of collaboration may differ for boys and girls in the classroom.

We also sought to establish if there were gender differences in learning, and particularly in the acquisition of basic level and conceptual knowledge. Our findings indicated that whereas both boys and girls progressed on the basic level, only boys showed significant improvement in conceptual knowledge through collaboration.

It is important to note that, at this age at least, boys and girls do not differ in performance in similar, independent classroom science tests (e.g., Shepardson & Pizzini, 2010). Nor did our participants differ by gender in terms of their conceptual knowledge at pre-test. It is plainly false to assume that girls are less able to learn scientific concepts. Moreover, there are plenty of other routes to acquiring such scientific knowledge aside from peer collaboration. For instance, conceptual development for girls may occur more often through independent study or teacher-led learning than through peer collaboration (see Tyler-Wood, Ellison, Lim, & Periathiruvadi, 2012).

Given that girls and boys advance differently in terms of conceptual understanding through collaboration, an obvious question is whether aspects of girls' and boys' conversations are associated with this difference. We expected no gender differences in the content of conversations (hypothesis two) because both boys and girls have the same access in conversation to the sorts of ideas and concepts that are relevant to understanding the task.

However this hypothesis was not confirmed: results indicated that boys used more conceptual level language in their interactions than girls.

There were also significant gender differences in behavioral dynamics, confirming our third hypothesis. Boys were more assertive in their interactions with one another, which is consistent with a good deal of previous research that identifies male conversations as more dominance oriented and girls' conversations as more affiliative (Leaper & Smith, 2004).

Previous research (e.g., Howe et al., 2014) has failed to find a reliable causal relation between features of conversation content and dynamics and learning. Our moderation analysis, overall, indicated that the features of interaction may be more active in terms of learning for boys than for girls. For boys, applied talk was negatively associated and social (off topic) talk was positively associated with advances in basic level knowledge. Boys who used more procedural justifications tended to advance more in terms of their conceptual knowledge. This indicates that talking more about the "hands on" aspects of doing the task, something we might naturally associate with successful collaboration in science, were associated with conceptual gains for boys but not for girls. Studies from an educational context point to these active aspects of engagement with science equipment and technology as marking an activity out as being within a male domain (Littleton, Light, Joiner, Messer, & Barnes, 1998). Moreover, our analyses indicated that the use of conceptual justifications was associated with conceptual development only for boys.

Girls' conversational styles are often oriented towards achieving cooperation and consensus, whereas boys' interaction styles may be better suited to learning certain types of information through collaboration because greater conflict and disagreement stimulates a deeper exploration of underlying concepts. Thus, peer collaboration on a science task such as

this, at 9 years of age, appears to be an effective means for boys to learn conceptual information but not for girls. This gender difference may arise from the ways in which children view science and the forms of talk that are triggered by the task in this domain. These gender differences in talk and engagement with the topic in collaboration could entail longer term consequences for academic success and participation in science careers into adulthood.

These findings suggest that interventions that are targeted to encourage argumentation could allow participants to appreciate the constructive role of disagreement and may help to stimulate discussion of conceptual issues in girls' conversations in science. More work is needed to establish if this extends to other age groups and other domains, including domains that are differently gender-marked. Educators need to consider carefully how collaboration is simultaneously a learning and a social activity for children and, consequently, how differing perceptions and orientations towards disagreements can promote effective learning.

References

Ceci, S. J. & Williams, W. M. (Eds) (2007). *Why Aren't There More Women in Science? Top Researchers Debate The Evidence*. Washington: American Psychological Association.

Di Donato, M. D., Martin, C. L., & England, D. (2014). "Gendered Interactions and Their Consequences: A Dynamical Perspective" (pp.20-42), in P. J. Leman & H. Tenenbaum (Eds.), *Gender and Development*. Oxford, UK: Psychology Press.

Good, C., Rattan, A., & Dweck, C.S. (2012). Why Do Women Opt Out? Sense of Belonging and Women's Representation in Mathematics. *Journal of Personality and Social Psychology, 102*, 700-717.

Gummerum, M., Leman, P. J., & Hollins, T. S. (2014). How Do Children Share Information In Groups? *Developmental Psychology, 50,* 2105-2114

Howe, C. (2009). Collaborative group work in middle childhood: Joint construction, unresolved contradiction and the growth of knowledge. *Human Developmen*t, *39*, 71-94.

Howe, C., McWilliam, D., & Cross, G. (2005). Chance favours only the prepared mind: incubation and the delayed effects of peer collaboration. *British Journal of Psychology,* 96, 67-93.

Howe, C., Taylor Tavares, J. & Devine, A. (2014), Children's conceptions of physical events: Explicit and tacit understanding of horizontal motion. *British Journal of Developmental Psychology, 32,* 141–162.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96,* 674–689.

Kuhn, D. & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22*, 545-552.

Leaper, C. (1991). Influence and Involvement in Children's Discourse: Age, Gender, and Partner Effects. *Child Development, 62*, 797-891.

Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: Talkativeness, affiliative speech, and assertive speech. *Developmental Psychology, 40,* 993–1027.

Leman, P. J., Ahmed, S., & Ozarow, L. (2005). Gender, gender relations and the social dynamics of children's conversations. *Developmental Psychology, 41*, 64–74.

Leman, P. J. & Lam, V. L. (2008). The influence of race and gender on children's conversations and playmate choices. *Child Development, 79,* 1329-1343.

Littleton, K., Light, P. H., Joiner, R., Messer, D., & Barnes, P. (1998) Gender, Task Scenarios and Children's Computer Based Problem Solving. *Educational Psychology, 18*, 327-340.

Mercer, N**.,** Dawes, R., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal*, *30*, 367-385.

Phelps, E. & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81*, 639-646.

Piaget, J. (1967). *Logique et connaissance scientifique.* Paris: Gallimard.

Shepardson, D. P. & Pizzini, E. L. (2010). Gender, achievement, and perception toward science activities. *School Science and Mathematics, 94*, 188-193.

Smith, R. B., Davidson, J., & Ball, P. (2001). Age-related variations and sex differences in gender cleavage during middle childhood. *Personal Relationships,8*, 153-165.

Tyler-Wood, T., Ellison, A., Lim, O., & Periathiruvadi, S. (2012). Bringing Up Girls in

    Science (BUGS): The effectiveness of an afterschool environmental science program for

    increasing female students' interest in science careers. *Journal of Science Education and*

    Technology, 21, 46-55.

Van Boxtel, C., Van der Linden, J. L., & Kanselaar, G. (2000). Collaborative learning tasks

    and the elaboration of conceptual knowledge. *Learning and Instruction, 10*, 311–330

Weisgram, E. S., Bigler, R. S., & Liben, L. S. (2010). Gender, values, and occupational

    interests among children, adolescents, and adults. *Child Development*, *81*, 778-796.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental*

    Review, 20, 99–149.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and

    middle school. *Developmental Review*, *27*, 172-223.

Table 1.

Categories for coding conversation content.

Category	Description
Procedural	Discussions about how to do the task and which variables to try: e.g., "Put the carpet on," "Make it a steep slope," or "Let's do the next question".
Basic	Discussions about the basic properties of the variables, descriptions: e.g., "Carpet is bumpy," "Steep is faster"
Conceptual	Discussions about the concepts of the variables, deeper understanding: e.g. "The bumpy carpet creates more friction" or "The heavy weight gives more force"
Applied	Discussions about the variables in the real world: e.g., "My bike goes faster down a hill"
Social (off-topic)	Discussion unrelated to the topic: e.g., "What did you have for lunch today"

Table 2.

Mean scores and standard deviations for performance on pre- and post-tests by gender and interaction group (collaborative science versus no interaction)

		Pre-Test		Post-test	
		M	*sd*	M	*sd*
Basic level knowledge					
Science group	Boys (N=72)	3.00	*.93*	3.53	*.75*
	Girls (N=84)	2.67	*.96*	3.46	*.72*
No interaction	Boys (N=94)	2.77	*.91*	2.91	*.83*
	Girls (N=74)	2.56	*1.01*	2.64	*.98*
Conceptual level knowledge					
Science group	Boys (N=72)	18.07	*4.23*	20.19	*3.16*
	Girls (N=84)	19.56	*3.07*	18.96	*3.54*
No interaction	Boys (N=94)	18.19	*4.01*	17.83	*3.34*
	Girls (N=74)	18.40	*3.59*	17.76	*3.55*

Note: Basic and conceptual knowledge and pre- and post-tests were assessed using different materials, so although equivalent in their assessment of the underlying knowledge the mean figures are not directly comparable.

Table 3.

Gender differences in mean use (utterances) of each measure of conversation content and behavior

	Boys	Girls	t(155)	Cohen's *d*
	(N=72)	(N=84)		
Conversation content				
Procedural	7.92 (4.42)	9.00 (5.56)	1.27	.20
Basic	3.39 (2.66)	4.15 (3.10)	1.64	.26
Conceptual	1.06 (1.54)	.31 (.71)	4.00*	.64
Applied	.07 (.26)	.01 (.11)	1.89†	.30
Social (off-topic)	1.33 (2.69)	1.09 (2.14)	.62	.10
Behavioral measures				
Assertion	2.84 (.86)	2.57 (.92)	1.99*	.32
Affiliation	5.01 (.90)	5.26 (1.09)	1.80†	.29

† *p*<.10, * *p*<.05

Table 4.
Regression analyses summary for predictors of basic and conceptual knowledge at post-test.

	Basic knowledge					Conceptual knowledge				
	Predictor statistics			Block change statistics		Predictor statistics			Block change statistics	
	β	t	p	Significance	R^2	β	t	p	Significance	R^2
Block 1				$F(1,137)=5.75$, $p=.018$	.04				$F(1,129)=9.64$, $p=.002$	.07
Pre-test score	0.158	2.40	.018			0.264	2.57	.011		
Block 2				$F(1,136)=0.03$, $p=.875$	<.01				$F(1,128)=0.81$, $p=.371$	<.01
Pre-test score	0.156	2.32	.022			0.278	3.21	.002		
Gender	-0.020	0.16	.875			-0.536	0.90	.371		
Block 3				$F(7,129)=1.18$, $p=.320$	.06				$F(7,121)=1.14$, $p=.344$	.06
Pre-test score	0.150	2.19	.030			0.323	3.58	.001		
Gender	-0.002	0.12	.991			0.011	0.02	.986		
Procedural	0.009	0.69	.489			0.004	0.06	.951		
Basic	-0.013	0.56	.578			-0.161	1.43	.154		
Conceptual	0.077	1.14	.255			0.381	1.48	.142		
Applied	-0.618	1.77	.079			0.622	0.36	.720		
Social (off-topic) talk	-0.032	0.92	.358			0.122	0.76	.447		
Assertion	<0.001	<0.01	.999			0.561	1.42	.159		
Affiliation	0.025	0.32	.748			0.372	1.05	.295		
Block 4				$F(7,122)=1.97$, $p=.064$	.09				$F(7,114)=1.63$, $p=.134$.	.08
Pre-test score	0.156	2.28	.024			0.305	3.24	.002		
Gender	-1.248	1.37	.175			1.546	0.36	.723		
Procedural	-0.015	0.31	.759			0.445	1.98	.050		

Basic	-0.045	0.57	.573	-0.087	0.24	.815
Conceptual	-0.171	0.79	.433	2.252	2.67	.009
Applied	-3.074	2.70	.008	8.041	1.43	.155
Social (off-topic) talk	0.330	2.02	.046	-0.193	0.25	.803
Assertion	-0.418	1.31	.193	0.514	0.35	.729
Affiliation	-0.221	0.76	.450	0.166	0.12	.903
Gender*Procedural	0.013	0.46	.644	-0.266	2.03	.045
Gender*Basic	0.008	0.17	.866	-0.002	0.01	.993
Gender*Conceptual	0.224	1.50	.137	-1.493	2.37	.019
Gender*Applied	1.779	2.04	.044	-5.472	1.33	.188
Gender*Social	-0.198	2.22	.028	0.212	0.50	.617
Gender*Assertion	0.277	1.49	.138	-0.129	0.15	.881
Gender*Affiliation	0.091	0.53	.598	0.339	0.42	.677