High Educ (2015) 70:469–486
DOI 10.1007/s10734-014-9849-3

Promoting university students' metacognitive regulation through peer learning: the potential of reciprocal peer tutoring

Liesje De Backer · Hilde Van Keer · Martin Valcke

Published online: 23 December 2014

Abstract Although successful learning in university education can be advanced by students' competence to self-regulate their learning, students often possess insufficient metacognitive regulation skills to regulate their learning adequately. The present study investigates changes in university students' adoption of metacognitive regulation after participating in reciprocal peer tutoring (RPT). A quasi-experimental pretest–posttest design was adopted, involving an experimental ($n = 51$) and two control groups; CG1 ($n = 24$) and CG2 ($n = 22$). Experimental students participated in a RPT intervention during a complete semester. Metacognitive regulation was assessed by means of think-aloud protocol analysis. Results indicate that RPT is promising to promote metacognitive regulation. Experimental students increasingly adopt monitoring, evaluation, and orientation and significantly evolve towards deep-level regulation from pretest to posttest. Except for an increased use of low-level comprehension monitoring, none of the evolutions in experimental students' regulation could be discerned for students in both control groups.

Keywords Metacognitive regulation · Peer tutoring · Think-aloud protocol analysis · Collaborative learning

Introduction

Metacognitive regulation is central to self-regulated learning and contributes to an important extent to students' performance (Meijer et al. 2006; Zimmerman and Schunk 2011). It generally advances academic success, especially in higher education since both organisational structures and academic assignments at this educational level emphasise self-management and independent learning (Nota et al. 2004). Nonetheless, students' metacognitive regulation is often insufficient to self-regulate their learning adequately,

L. De Backer (✉) · H. Van Keer · M. Valcke
Department of Educational Studies, Ghent University, H. Dunantlaan 2, 9000 Ghent, Belgium
e-mail: Liesje.DeBacker@UGent.be

revealing the necessity to design, implement, and evaluate initiatives fostering metacognitive regulation.

The present study examines the impact of reciprocal peer tutoring (RPT) on university students' metacognitive regulation. Unlike most other research, it does not focus exclusively on a particular regulation skill (e.g. monitoring), but takes an integrative perspective when assessing students' metacognitive regulation. It more specifically investigates changes in students' metacognitive orientation, planning, monitoring, and evaluation after participation in RPT. By enhancing our understanding of the differential impact of RPT on students' adoption of regulation skills, the present study offers direct cues to optimally foster students' regulation through RPT.

Theoretical framework

Metacognitive regulation

Metacognition refers to both the awareness and active control students have over their cognitive activities when engaged in learning or academic problem-solving (Brown 1987; Pintrich 2004). The first component, metacognitive knowledge, concerns insights of students about themselves as learners, insight into learning strategies, and insight into the usefulness of these strategies within specific learning conditions (Pintrich 2004; Zimmerman and Schunk 2011). The second component, metacognitive regulation, refers to self-regulatory skills and strategies adopted by students to actively control and coordinate their learning and performance (Greene and Azevedo 2009; Meijer et al. 2006). The focus of the present study is on this second component of metacognition. We distinguish orienting, planning, monitoring, and evaluating as key regulation skills (Brown 1987; Veenman et al. 1997), which unfold over weakly sequenced and recursive problem-solving phases (Greene and Azevedo 2009; Winne and Hadwin 1998). When *orienting*, students engage in task analysis to get acquainted with learning objectives or task demands (Pintrich 2004). This can result in awareness of task perceptions or prior knowledge activation (Meijer et al. 2006; Winne and Hadwin 1998). *Planning* encompasses selecting and sequencing problem-solving strategies, allocating resources, and developing action plans (Greene and Azevedo 2009; Meijer et al. 2006); either at the onset or during task execution (e.g. after a subtask). *Monitoring* involves the online quality control of problem-solving, aimed at identifying inconsistencies and optimising task execution (Meijer et al. 2006; Moos and Azevedo 2009). Comprehension monitoring refers to control activities focusing on the correctness of one's understanding (Pintrich 2004; Veenman et al. 1997), whereas monitoring of progress focuses on the adequateness of problem-solving strategies or the quality of students' progress (Greene and Azevedo 2009; Moos and Azevedo 2009). Finally, *evaluation* involves learners' self-judgement upon completion of problem-solving (Veenman et al. 2005). This can be directed at both learning outcomes and the learning process (Meijer et al. 2006).

Low-level versus deep-level metacognitive regulation

The present study distinguishes low-level and deep-level metacognitive regulation, introducing a more in-depth operationalisation. Low-level *orientation* is solely directed at exploring task demands, whereas deep-level orientation aims at processing task demands and activating prior knowledge (Veenman et al. 2005). Low-level *planning* implies the development of a single problem-solving plan, whereas deep-level planning involves selecting a plan after considering various problem-solving alternatives (Meijer et al. 2006; Veenman et al. 1997). When students check their progress or the comprehensiveness of their understanding, they engage in low-level *monitoring*. Reflective comments on the quality of their perceived progress or elaborative, thought-provoking reflections imply deep-level monitoring (Moos and Azevedo 2009). Correspondingly, low-level *evaluation* involves checking and commenting on either learning outcomes or process factors, whereas deep-level evaluation implies reflective judgements on both (Veenman et al. 2005).

Optimising metacognitive regulation through peer tutoring

Fostering metacognitive regulation requires direct observation of explicitly modelled metacognitive behaviour, along with explanations about regulation skills (Schunk and Zimmerman 2007; Volet et al. 2009). Students should further be challenged to internalise the modelled behaviour at the individual level, by practising and subsequently reflecting upon a variety of self-regulatory strategies. Additionally, learners should be encouraged to discuss and control their learning, refining their metacognitive regulation (Schunk and Zimmerman 2007). Current research centres on modelling by peers during collaborative learning (Volet et al. 2009). During collaborative learning, students discuss meaning and compare their thinking with peers, requiring both regulation of their own cognition and control of how peers learn (Volet et al. 2009). Despite growing consensus on the metacognitive learning opportunities during collaborative learning, empirical evidence on its differential influence on students' adoption of specific regulation skills is limited. The present study aims to enhance our understanding of the metacognitive potential of collaborative learning, by studying the specific impact of reciprocal peer tutoring on university students' use of orientation, planning, monitoring, and evaluation.

Peer tutoring is characterised by active academic helping and supporting between peers in small groups or student pairs (Falchikov 2001; Topping 1996). One peer, the tutor, is expected to take a direct pedagogical responsibility by creating learning opportunities through questioning, clarifying, and scaffolding (Chi et al. 2001). The students being cognitively challenged by this peer tutor are called tutees. Reciprocal peer tutoring (RPT) is characterised by the structured exchange of the tutor role among peers in the PT groups/pairs (Topping 1996). Although the available research illustrates that peer tutoring challenges students' metacognitive regulation and particularly promotes their adoption of monitoring (e.g. King 1997; Roscoe and Chi 2008), possible effects on other regulation skills are underexposed.

Aim of the study and research hypotheses

The present study aims to study the changes in university students' adoption of metacognitive regulation skills after participation in RPT. Since literature provides evidence for metacognitive learning opportunities within collaborative learning, we hypothesise that RPT will have a positive impact on students' use of metacognitive regulation skills (hypothesis 1). During tutoring, conceptual discussions can take place at different levels of social interaction: the individual, dyadic, and group level (Volet et al. 2009). Consequently, students experience the need to metacognitively regulate their own, each other's, or the group's cognition, increasing chances for metacognitive engagement. Because more regulatory control often results in the use of profound metacognitive strategies (Chin and Brown 2000; Greene and Azevedo 2009), we additionally hypothesise that RPT will promote students' involvement in deep-level metacognitive regulation (hypothesis 2).

Method

Design

A quasi-experimental pretest–posttest design was adopted, involving an experimental and two control groups. Experimental students participated in the RPT intervention during a complete semester. Both at the start (October) and the end (December) of the semester, the regulation skills of individual participants ($n = 97$) were assessed.

Participants and setting

The study was conducted in a naturalistic university setting. The experimental group (EG) consisted of the complete population of 64 first-year students, Educational Sciences, who already obtained a professional bachelor degree (12.5 % males and 87.5 % females).[1] This gender distribution is representative for the Educational Sciences students population in Flanders (Belgium). Unlike other studies (e.g. Cheng and Ku 2009; Duran and Monereo 2005), the present study did not aim to compare RPT with a different type of student-activating learning (e.g. fixed peer tutoring, self-explaining). Therefore, we preferred not to assign students of the Educational Sciences programme with a professional bachelor degree to either an experimental or control treatment, due to ethical reasons. Such a design would imply students being deprived from the benefits of peer-assisted learning, inherent to RPT. Alternatively, we involved two control groups. The first (CG1) consisted of 24 freshmen in the Educational Sciences programme (12.5 % males and 87.5 % females). Despite differences in participants' age and prior experience in higher education, these students are enrolled in the same curriculum as EG students. This curriculum aims at introducing students into the basic theoretical frameworks of Educational Sciences and consists of general social sciences courses, specialised pedagogical courses, and methodology courses. The second control group (CG2) consisted of 22 first-year students in the Social Welfare Studies programme of the same faculty, who also attained a professional bachelor degree (14.3 % males and 85.7 % females). Although these students are enrolled in a somewhat different curriculum, their background and prior experience in higher education is comparable to EG students. CG2 students' curriculum aims at introducing students into the theoretical frameworks of Social Welfare Studies and consists of general pedagogical courses, specialised courses on Social Welfare Studies, and methodology courses. The didactical approach of these courses is comparable to the curriculum of EG and CG1 students: a dominance of theoretical lectures is alternated with group work, which focuses on writing a theoretical paper with a small group of students. During the research period, students from both control groups attended regular curriculum activities as part of their academic training. They were not involved in tutoring, neither in any comparable

[1] Although all 64 RPT students participated in the pretest, not all of them attended the posttest assessment. Additionally, the tape recordings of some RPT students demonstrated technical problems. Therefore, the data of 13 EG students were excluded from the analysis.

collaborative learning approach, encompassing systematic conceptual interaction and discussions among students in a face-to-face setting.

EG students were randomly assigned to eleven RPT groups. The RPT programme was a formal component of the 5-credit course "Instructional Sciences" (of which the theoretical lectures were part of both EG and CG1 students' curriculum). Informed consent was obtained from all participants in the study.

Intervention

The RPT intervention consisted of eight successive face-to-face sessions (each taking 2 h), in which students tutored each other in small and stable groups of six. The tutor role was changed at each session. As a manipulation check, all RPT groups were observed weekly, to control whether students enacted their roles adequately.

RPT assignments

During each session, students worked on authentic, open-ended group assignments (De Backer et al. 2012). Each assignment comprised of an outline of learning objectives to guide peers' discussions to central course-related topics; a subtask aimed at getting familiar with the theme-specific terminology; and a subtask in which students applied theoretical notions to realistic instructional cases.

Training

All EG students participated in a compulsory tutor training 1 week before the onset of the RPT intervention, focussing on tutor responsibilities and generic tutoring skills (i.e. establishing a safe learning climate, managing peer interactions, questioning, explaining, scaffolding, and providing feedback) (Chi et al. 2001; King 1997). The training used videotaped examples of good and bad practices, which were discussed in depth, role plays in which students experienced multiple tutor responsibilities and received feedback on their tutoring approach, and the in-depth analysis of authentic case studies focusing on specific tutor competences. The outlines of the training were presented in a manual provided to all EG students.

Tutor guide

To prepare themselves for the tutor role, tutoring students received a session-specific "tutor guide" 1 week in advance. This offered additional information about the theoretical learning content to focus upon in the RPT session and assured a difference in tutors' and tutees' domain-specific knowledge (Topping 1996). Although the theoretical content of the guide differed across sessions, its structure and design were identical throughout the RPT intervention. In addition to offering theoretical knowledge, the guide exemplified a stepwise problem-solving approach (De Backer et al. 2012).

Interim support

An interim supervision session and two-weekly feedback sessions were organised, to provide tutor support (Falchikov 2001). The supervision was set up in small groups of twelve students and directed by a university staff member, who encouraged students to reflect on the adequacy of their tutor/tutee behaviour. Additionally, the staff member provided group-specific feedback once every 2 weeks.

Instruments

Assessment of students' metacognitive regulation was based on think-aloud protocol analysis. At the start and the end of the research, all participants individually performed an academic task that was videotaped. Participants were instructed to solve the task and to verbalise their problem-solving, resulting in verbal protocols (Greene et al. 2011).

Task

The individual think-aloud task comprised of a theoretical text and a real-life case relevant to the course "Instructional Sciences" (i.e. topic "evaluation" and "social inequity in education" at pretest and posttest, respectively). Students solved thought-provoking questions about the text. In case students stopped verbalising during task performance, they were prompted by the assessor to continue thinking aloud. Apart from a difference in the central topic, all aspects of the think-aloud task and the assessment procedure were identical at pretest and posttest. Both tasks provided a challenging yet comprehensible text, comprised of multiple parts, and had to be executed during the mild time constraint of 30 min (Greene et al. 2011).

Coding instrument

To analyse students' verbal protocols, we developed a literature-based coding instrument, representing a multi-layered model of metacognitive regulation (De Backer et al. 2012). Orientation, planning, monitoring, and evaluation are adopted as the main coding categories and further specified with sub-coding categories (i.e. task analysis, content orientation, planning in advance, interim planning, comprehension monitoring, monitoring of progress, evaluation of learning outcomes, and evaluation of the learning process). Additionally, a dimension on the low-/deep-level approach is included.

Coding strategy

The verbal protocols of all students were transcribed verbatim and coded by two independent and trained coders. They double coded 25 % of the protocols to determine interrater reliability. Cohen's kappa ($\kappa = 0.78$) indicates high overall interrater reliability and good agreement beyond chance for the four main coding categories ($\kappa_{\text{orientation}} = 0.89$, $\kappa_{\text{planning}} = 0.86$, $\kappa_{\text{monitoring}} = 0.82$, and $\kappa_{\text{evaluation}} = 0.89$). The coding procedure followed subsequent phases and focused exclusively on students' verbalised behaviour. Each verbal protocol was initially segmented according to changes in the verbalisation focus, which formed the boundaries of substantial "episodes" (Chi et al. 2001). An episode is conceptualised as a brief segment of the overall verbal protocol that was centred around one particular action. After segmentation, each episode received a general code, indicating whether it concerned a metacognitive, task-executive, or off-task episode. Second, metacognitive episodes were selected for more detailed "statement coding" (Roscoe and Chi 2008). A statement refers to a single thematically consistent verbalisation of a single metacognitive action. Next, the identified metacognitive statements were coded by means of the developed coding instrument. Every statement received a general code (indicating the regulation skill it addressed) and a more differentiated code (referring to the concretised regulation strategies situated in the subcategories of the coding instrument). Additionally, the approach (low level /deep level) to the identified metacognitive statements was coded. Table 1 exemplifies the coding strategy.

Data analysis

First, the frequency of occurrence of all regulation skills and strategies, and of low-/deep-level approaches, was calculated for each protocol per participant. The frequency of occurrence of regulation skills and approaches in individual students' protocols was aggregated for each research condition. These frequencies per condition were used for analysis purposes. Second, pre-analysis investigations were conducted to check both the assumption of normally distributed data (i.e. by means of Kolmogorov–Smirnov tests) and the assumption of homogeneity of variance (i.e. by means of Levene's tests). Since the results revealed that neither assumption was violated, parametric analyses (i.e. *t* tests) were performed to examine the significance of differences in the adoption of metacognitive regulation between conditions and measurement occasions. Third, independent samples *t* tests were conducted (i.e. comparing EG with CG1, EG with CG2, and CG1 with CG2, respectively) to assure that the metacognitive regulation of students in the three conditions was comparable at pretest. Fourth, the frequency of occurrence of regulation skills and approaches at pretest and posttest was compared by means of paired samples *t* tests, to study the impact of RPT on students' use of metacognitive skills (hypothesis 1) and on their deep-level regulation approach (hypothesis 2). Significant differences in students' metacognitive regulation at pretest and posttest were tested for each condition separately (see Tables 3, 4, 5). Cohen's *d* is reported as a measure of effect size of significant differences which were identified in students' adoption of regulation skills and a deep-level regulation approach. The significance level was 0.05 for all analyses.

Results

Descriptive analysis on students' metacognitive regulation

Table 2 demonstrates a dominant involvement of all students in metacognitive monitoring, both at pretest and posttest, whereas their adoption of orientation, planning, and evaluation is rather limited, especially at pretest. Results moreover reveal a clear pretest-to-posttest change towards more frequent use for the majority of metacognitive strategies by EG students. For some metacognitive behaviour, this trend is markedly smaller in both control groups (e.g. comprehension monitoring, evaluation of learning outcomes), while for other metacognitive strategies this positive evolution cannot be discerned at all (e.g. orientation, monitoring of progress). In contrast, the planning behaviour of students appears to evolve similarly for all research groups. Results further reveal students' limited attention to evaluation of the learning process, deep-level monitoring of progress, deep-level planning, and taking into account task perceptions. This is observed at pretest and posttest, in all conditions.[2]

[2] Metacognitive regulation strategies with very low frequency of occurrence (<1 % in Table 2) were removed from further analyses.

Table 1 Exemplified coding strategy

Verbal protocol excerpt	Episode coding	Statement coding	Approach
It is about assessment and evaluation	Metacognition	Exploring task demands (orientation)	Low level
I first read the instructions to know what I have to do. [student reads silently]	Metacognition	Exploring task demands (orientation)	Low level
Okay, search for different evaluation forms. And the second question, tell for which purposes assessment is used	Metacognition	Processing task demands (orientation)	Deep level
Now I will read the text	Metacognition	Formulating problem-solving plan (planning)	Low level
[student reads] "Regarding the functions of evaluation, we distinguish summative and formative evaluation. Summative occurs at the end of learning, aimed at controlling whether objectives are reached. Formative evaluation intends to provide interim feedback…" [student does not finish sentence]	Task execution		
Wait (…) there is summative and formative… but what is the difference?	Metacognition	Noting lack of comprehension (Monitoring)	Low level
I reread this part. [student rereads silently]	Task execution		
So summative at the end, formative during learning	Metacognition	Demonstrating comprehension by repeating (Monitoring)	Low level
I continue reading (…)	Task execution		

Impact of RPT on students' adoption of metacognitive regulation

Paired samples t tests on EG students' adoption of regulation skills reveal a significant increase in monitoring, evaluation, and (albeit to a lesser extent) orientation from pretest to posttest (see Table 3). Additionally, EG students demonstrate a significantly enhanced involvement in deep-level regulation at posttest. The reported effect sizes for these evolutions are moreover large (see Table 3). No significant changes could be identified in none of the above-mentioned regulation skills, neither in a deep-level regulation approach, for CG1 students (see Table 4) or CG2 students (see Table 5). The planning behaviour of students in all conditions does not differ significantly from pretest to posttest.

The following paragraph outlines the pretest-to-posttest changes in both the frequency of occurrence of metacognitive regulation skills (hypothesis 1) and adoption of deep-level regulation (hypothesis 2) in more detail. Significant changes are addressed for each metacognitive regulation skill separately.

Monitoring

Table 3 reveals a major increase in EG students' adoption of comprehension monitoring from pretest to posttest. Compared to the pretest, EG students demonstrate enhanced involvement in monitoring their understanding at posttest, both by paraphrasing

Table 2 Occurrence of metacognitive regulation at pretest and posttest for the three research conditions (frequencies[a] and percentages[b])

Metacognitive regulation	Pretest						Posttest					
	EG (n = 51)		CG1 (n = 24)		CG2 (n = 22)		EG (n = 51)		CG1 (n = 24)		CG2 (n = 22)	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Orientation	70	10.9	28	10.0	37	14.5	220	10.3	38	8.4	35	8.3
Task analysis	64	9.9	26	9.3	34	13.3	117	8.3	38	8.4	32	7.6
Exploring task demands (LL)	54	8.4	22	7.9	29	11.4	130	6.1	35	7.7	29	6.9
Processing task demands (DL)	10	1.5	4	1.4	5	1.9	47	2.2	3	0.7	3	0.7
Content orientation	5	0.8	2	0.7	3	1.2	35	1.6	0	0.0	1	0.2
Generating hypotheses (DL)	1	0.2	0	0.0	1	0.4	2	0.1	0	0.0	0	0.0
Activating prior knowledge (DL)	4	0.6	2	0.7	2	0.8	33	1.5	0	0.0	1	0.2
Becoming aware of task perceptions	0	0.0	0	0.0	0	0.0	8	0.3	0	0.0	2	0.5
Planning	88	13.7	52	18.6	42	16.5	211	9.9	106	23.4	100	23.7
Planning in advance	28	4.3	14	5.0	17	6.7	76	3.6	29	6.4	27	6.4
Formulating problem-solving plan (LL)	28	4.3	14	5.0	17	6.7	76	3.6	29	6.4	27	6.4
Selecting problem-solving plan (DL)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Interim planning	60	9.5	38	13.6	25	9.8	135	6.3	78	17.0	73	17.3
Formulating problem-solving plan (LL)	60	9.5	38	13.6	25	9.8	125	6.3	78	17.0	73	17.3
Selecting problem-solving plan (DL)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Monitoring	447	69.4	191	68.2	162	63.5	1,472	68.8	286	63.1	263	62.3
Comprehension monitoring	228	35.4	84	30.0	70	27.5	1,019	47.7	189	41.7	202	47.9
Noting lack of comprehension	85	13.2	42	15.0	36	14.4	138	6.5	38	8.4	33	7.8
Summarising main ideas	19	3.0	0	0.0	9	3.5	38	1.8	18	4.0	39	9.2
Demonstrating comprehension by repeating (LL)	95	14.8	28	10.0	19	7.5	445	20.8	105	23.3	89	21.1
Demonstrating comprehension by elaborating (DL)	33	5.1	14	5.0	6	2.4	395	18.5	29	6.4	41	9.8
Monitoring of progress	219	34.0	109	38.9	91	35.7	457	21.4	97	21.4	61	14.4
Checking of progress (LL)	217	33.7	108	38.6	89	34.9	402	18.8	95	21.0	59	14.0

Table 2 continued

Metacognitive regulation	Pretest						Posttest					
	EG ($n = 51$)		CG1 ($n = 24$)		CG2 ($n = 22$)		EG ($n = 51$)		CG1 ($n = 24$)		CG2 ($n = 22$)	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Reflecting on progress (DL)	2	0.3	1	0.3	2	0.8	55	2.6	2	0.4	2	0.4
Evaluation	33	5.1	3	3.2	13	5.1	231	10.7	21	4.6	23	5.5
Evaluating learning outcomes	33	4.9	9	3.2	13	5.1	220	10.3	21	4.6	23	5.5
Checking learning outcomes (LL)	32	0.2	9	3.2	13	5.1	206	9.6	20	4.4	21	5.0
Elaborating on learning outcomes (DL)	1	0.0	0	0.0	0	0.0	14	0.7	1	0.2	2	0.5
Evaluating learning process	0	0.0	0	0.0	0	0.0	11	0.4	0	0.0	0	0.0
Commenting on learning process (LL)	0	0.0	0	0.0	0	0.0	11	0.4	0	0.0	0	0.0
Reflecting on learning process (DL)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

LL low level, *DL* deep level

[a] The columns "frequency" present the absolute frequency of occurrence of a particular metacognitive regulation skill (i.e. orientation, planning, monitoring, or evaluation) or strategy (i.e. at the more concretised level of the coding instrument) within a research condition (i.e. EG, CG1, or CG2) at a particular measurement occasion (i.e. pretest or posttest)

[b] The columns "percentages" present the proportion of a particular metacognitive regulation skill or strategy within a research condition at a particular measurement occasion taking into account all metacognitive regulation utterances that occurred and were assessed in that respective research group at that respective measurement occasion

Table 3 Results of pre- and posttest think-aloud protocol analysis in EG condition: Occurrence of metacognitive regulation skills and low-level versus deep-level regulation

Metacognitive skill	Frequency				$t(df)$	p	Cohen's d
	Pretest		Posttest				
	M^3	SD	M	SD			
Orientation	1.37	1.21	4.21	2.06	−9.10 (50)	<.001	1.73
Task analysis	1.25	1.09	3.33	1.69	−8.93 (50)	<.001	1.53
Exploring task demands (LL)	1.06	0.18	2.49	1.02	−8.37 (50)	<.001	1.55
Processing task demands (DL)	0.18	0.07	0.86	0.11	−5.69 (50)	<.001	1.02
Content orientation	0.10	0.05	0.67	0.08	−5.80 (50)	<.001	1.19
Activating prior knowledge (DL)	0.08	0.04	0.62	0.08	−5.58 (50)	<.001	1.17
Planning	1.72	1.08	2.16	1.06	−1.91 (50)	.062	0.40
Monitoring	8.37	3.33	28.86	8.66	−15.55 (50)	<.001	3.14
Comprehension monitoring	4.10	2.39	19.98	7.14	−14.35 (50)	<.001	3.33
Noting lack of comprehension	1.67	0.99	2.71	1.73	−3.85 (50)	<.001	0.76
Summarising main ideas	0.37	0.10	0.74	0.11	−2.72 (50)	<.001	0.47
Demonstrating comprehension by repeating (LL)	1.47	0.17	8.73	2.06	−16.73 (50)	<.001	3.38
Demonstrating comprehension by elaborating (DL)	0.65	0.14	7.74	4.08	−11.94 (50)	<.001	2.79
Monitoring of progress	4.29	1.75	8.96	3.43	−9.26 (50)	<.001	1.79
Checking progress (LL)	4.29	1.79	7.88	3.02	−8.25 (50)	<.001	1.49
Reflecting on progress (DL)	0.04	0.03	1.06	0.90	−7.38 (50)	<.001	1.72
Evaluation	0.65	0.14	3.53	2.09	−13.42 (50)	<.001	2.58
Evaluating learning outcomes	0.65	0.13	3.31	1.19	−14.06 (50)	<.001	4.03
Checking outcomes (LL)	0.65	0.14	3.04	1.70	−13.98 (50)	<.001	2.59
Elaborating on outcomes (DL)	0.00	0.00	0.27	0.07	−3.68 (50)	.010	1.03

M refers to how often an individual student on average uses a metacognitive regulation skill during think-aloud problem-solving at pretest and at posttest

information and by elaborating on the learning content. Although CG1 students and CG2 students also increase significantly in more comprehension monitoring through paraphrasing at posttest (see Tables 4, 5, respectively), the pretest-to-posttest change is significantly larger for EG students. Regarding students' deep-level regulation approach, the results reveal a major pretest-to-posttest change for deep-level comprehension monitoring for EG students (see Table 3). EG students are, moreover, nearly equally involved in low-level and deep-level comprehension monitoring, whereas control students predominantly engage in low-level comprehension monitoring (see Table 2). CG1 students' and CG2 students' adoption of deep-level comprehension monitoring does not change significantly from pretest to posttest (see Tables 4, 5, respectively).

Further, EG students significantly increase the frequency of monitoring their progress from pretest to posttest (see Table 3). Remarkably, a significant negative pretest-to-posttest change was revealed for both CG1 students (see Table 4) and CG2 students (see Table 5). Although Table 3 additionally demonstrates a significant pretest-to-posttest increase in deep-level monitoring of progress for EG students, their involvement in this respect

Table 4 Results of pre- and posttest think-aloud protocol analysis in CG1 condition: Occurrence of metacognitive regulation skills and low-level versus deep-level regulation

Metacognitive skill	Frequency				$t(df)$	p	Cohen's d
	Pretest		Posttest				
	M	SD	M	SD			
Orientation	1.17	1.09	1.58	1.17	−1.22 (23)	0.233	0.36
Task analysis	1.08	1.00	1.58	1.17	−1.52 (23)	0.143	0.45
Exploring task demands (LL)	0.92	0.58	1.46	0.83	−2.60 (23)	0.061	0.76
Processing task demands (DL)	0.17	0.13	0.12	0.09	0.25(23)	0.802	0.40
Content orientation	0.08	0.05	0.00	0.00	1.45 (23)	0.162	0.59
Activating prior knowledge (DL)	0.08	0.05	0.00	0.00	1.45 (23)	0.162	0.59
Planning	2.17	1.34	2.50	0.88	−1.09 (23)	0.286	0.30
Monitoring	7.96	2.71	7.50	1.69	1.06 (23)	0.302	0.21
Comprehension monitoring	3.50	1.79	5.12	1.72	−3.43 (23)	0.002	0.87
Noting lack of comprehension	1.75	1.11	1.08	0.87	2.56 (23)	0.071	0.43
Summarising main ideas	0.00	0.00	0.67	0.15	−4.29 (23)	<.001	1.75
Demonstrating comprehension by repeating (LL)	1.17	1.00	2.71	1.62	−4.21 (23)	<.001	1.17
Demonstrating comprehension by elaborating (DL)	0.50	0.22	0.67	0.16	−0.59 (23)	0.558	0.18
Monitoring of progress	4.54	1.69	2.33	1.40	6.20 (23)	<.001	1.45
Checking progress (LL)	4.50	1.69	2.33	1.40	6.11 (23)	<.001	1.39
Reflecting on progress (DL)	0.04	0.03	0.00	0.00	1.00 (23)	0.328	0.40
Evaluation	0.37	0.13	0.58	0.13	−1.00 (23)	0.328	0.32
Evaluating learning outcomes	0.37	0.13	0.58	0.13	1.00 (23)	0.328	0.32
Checking outcomes (LL)	0.37	0.13	0.58	0.13	−1.00 (23)	0.328	0.16
Elaborating on outcomes (DL)	0.00	0.00	0.00	0.00	–	–	–

remains low (2.6 %). CG1 students' and CG2 students' deep-level monitoring of progress does not change significantly from pretest to posttest (see Tables 4, 5, respectively).

Evaluation

Table 3 shows a significant pretest-to-posttest increase in EG students' evaluation of learning outcomes. In contrast, CG1 students' and CG2 students' engagement in evaluating learning outcomes remains limited and does not differ significantly from pretest to posttest (see Tables 4, 5, respectively). Despite significant pretest-to-posttest differences in EG students' deep-level evaluation (see Table 3), the occurrence of the latter remains very limited (0.7 %). No significant pretest-to-posttest changes could be found for CG1 students' and CG2 students' deep-level evaluation (see Tables 4, 5, respectively).

Orientation

Although EG students significantly increase their task analysis and prior knowledge activation from pretest to posttest (see Table 3), their involvement in orientation is less pronounced compared to the changes in their monitoring (i.e. of comprehension and

Table 5 Results of pre- and posttest think-aloud protocol analysis in CG2 condition: occurrence of metacognitive regulation skills and low-level versus deep-level regulation

Metacognitive skill	Frequency				$t(df)$	p	Cohen's d
	Pretest		Posttest				
	M	SD	M	SD			
Orientation	1.68	1.39	1.82	1.53	−0.37 (21)	0.715	0.11
Task analysis	1.54	1.22	1.64	1.29	−0.28 (21)	0.785	0.07
Exploring task demands (LL)	1.32	0.83	1.45	1.10	−0.48 (21)	0.633	0.14
Processing task demands (DL)	0.23	0.11	0.18	0.08	0.44 (21)	0.665	0.10
Content orientation	0.14	0.09	0.09	0.06	0.44 (21)	0.665	0.12
Activating prior knowledge (DL)	0.09	0.05	0.09	0.06	0.00 (21)	0.999	0.00
Planning	1.91	1.44	1.86	0.99	0.13 (21)	0.894	0.04
Monitoring	7.36	2.82	9.04	1.56	−2.18 (21)	0.061	0.76
Comprehension monitoring	3.18	2.03	6.91	1.79	−5.93 (21)	<.001	1.94
Noting lack of comprehension	1.64	0.72	1.00	0.81	2.85 (21)	0.010	0.82
Summarising main ideas	0.41	0.13	1.36	1.00	−3.47 (21)	0.002	1.19
Demonstrating comprehension by repeating (LL)	0.86	0.27	3.36	1.61	−5.56 (21)	<.001	1.72
Demonstrating comprehension by elaborating (DL)	0.32	0.14	1.18	0.38	−2.09 (21)	0.069	0.49
Monitoring of progress	4.14	2.31	2.09	1.26	3.70 (21)	<.001	1.14
Checking progress (LL)	4.14	2.31	2.04	1.17	3.88 (21)	<.001	1..19
Reflecting on progress (DL)	0.09	0.06	0.00	0.00	1.45 (21)	0.162	0.61
Evaluation	0.59	0.21	0.50	0.12	0.36 (21)	0.724	0.11
Evaluating learning outcomes	0.64	0.22	0.50	0.13	0.51 (21)	0.505	0.22
Checking outcomes (LL)	0.64	0.22	0.50	0.13	0.53 (21)	0.613	0.16
Elaborating on outcomes (DL)	0.14	0.09	0.00	0.00	1.37 (21)	0.186	0.58

progress) and evaluation (i.e. of learning outcomes). Table 3 further reveals a significant increase in EG students' adoption of both low-level and deep-level orientation. However, EG students' engagement in deep-level orientation remains limited (3.5 %). No significant pretest-to-posttest changes could be found for CG1 students' and CG2 students' deep-level orientation (see Tables 4, 5, respectively).

Discussion

Impact of RPT on students' adoption of metacognitive regulation

The results indicated that students are predominantly involved in monitoring their problem-solving when executing an academic task, both at pretest and at posttest, in all research conditions. The findings further revealed an increased adoption of monitoring and to a lesser extent evaluation and orientation by EG students at posttest. Except for low-level comprehension monitoring, the above-mentioned pretest-to-posttest increases were not discerned in the control conditions. These results consequently demonstrated that RPT is more beneficiary for fostering students' adoption of monitoring and (albeit to a lesser extent) evaluation and orientation, compared to traditional teaching approaches. A significant impact on students' planning behaviour could, however, not be distinguished, not for EG students neither for the control conditions. This might be due to the design of the think-aloud task, for the latter partially determines the outcomes of protocol analysis (Greene et al. 2011). Since students were expected to solve three thought-provoking questions on a well-structured academic task, the opportunities to plan were probably scarce. Additionally, students might not have felt the need to sequence problem-solving steps within the available time framework. It should further be noted that the developed coding instrument might not have been sensitive enough to capture students' planning appropriately. Future research should aim to conceptualise metacognitive planning in a more specific way in order to optimise its assessment.

RPT appeared to have a critical impact on students' application of comprehension monitoring. During RPT, students were challenged to approach the learning content critically and to negotiate its meaning, resulting in self-questioning (Roscoe and Chi 2008). It seems plausible to assume that a semester-long experience in this cognitively challenging RPT environment prompted students to internalise this comprehension monitoring behaviour. It should be noted, however, that control students also checked their understanding more often at posttest. Consequently, students' increased involvement in comprehension monitoring might be partially attributed to their experienced need for self-regulation in higher education (Nota et al. 2004). During their first semester at university, all students were presumably faced with the demands for elaborative thinking and self-control of one's understanding, resulting in the adoption of monitoring strategies. Nevertheless, our findings demonstrated that EG students showed a significant larger adoption of monitoring strategies as compared to students in control conditions, implying an added value of RPT.

Our findings further revealed increased monitoring of progress and more evaluative comments on learning outcomes for EG students. In contrast, both control groups did not increase their metacognitive reflection and evaluation, not during the course of problem-solving (i.e. monitoring their progress), neither upon completing it (i.e. metacognitive evaluation). These findings suggest that RPT is a promising approach when aiming to advance students' evaluative reflections. It can be assumed that the key elements of PT (i.e. asking and answering thought-provoking questions, providing knowledge-building explanations, scaffolding, giving feedback.) directly fostered students' self-reflections and evaluative insights (Chi et al. 2001; Roscoe and Chi 2008). Additionally, the design of the RPT learning materials might have promoted students' evaluative engagements, for the assignments and weekly tutor guides systematically outlined learning objectives, which might have served as an evaluative tool (Zimmerman and Schunk 2011).

Given EG students' increased adoption of certain metacognitive regulation strategies, we recommend higher education instructors to implement RPT when aiming to promote university students' metacognitive regulation. The results more specifically revealed that training students to tutor each other stimulates them to start monitoring their learning more frequently. Additionally, the present study suggests that well-structured and goal-oriented group assignments have potential to elicit particular evaluation strategies (albeit less pronounced compared to evoking monitoring acts). Consequently, organising RPT requires educators to carefully design learning materials which can encourage students in regulating their learning.

Impact of RPT on students' deep-level regulation approach

Notwithstanding students' dominant use of low-level regulation, at pretest and posttest, our findings revealed significant effects of RPT on students' deep-level metacognitive regulation. Given that none of the control groups demonstrated significant pretest-to-posttest changes towards a deep-level regulation approach for none of the key regulation skills, the above-mentioned result implies that RPT is more beneficiary to enhance students' engagement in deep-level regulation, compared to traditional teaching approaches. Our findings more specifically demonstrated that EG students outperformed control students in the adoption of deep-level comprehension monitoring. Since tutors were trained to promote tutees' profound reflective thinking by asking critical questions, providing cognitive scaffolds, and giving knowledge-building explanations, it could be assumed that RPT participants observed and eventually internalised these strategies, modelled by their tutors. EG students additionally revealed an increased use of deep-level task analysis, profound monitoring of progress, and deep-level evaluation of learning outcomes. It should be noted, however, that the frequency of occurrence of these deep-level regulation strategies remained low. Students' rather limited involvement in deep-level regulation might be explained by both the need for explicit metacognitive prompts and more extensive opportunities to practise regulation skills (Schunk and Zimmerman 2007). Since orientation and evaluation can only be employed before and upon completion of task execution, respectively, their frequency of occurrence might have been too limited for students to evolve towards frequent deep-level orientation and evaluation.

Based on our results, we advise higher education instructors to implement long-lasting RPT interventions, allowing students the time they need to evolve towards more frequent practice with different deep-level regulation strategies. This practice could increase the chances of students starting to internalise a deep-level approach when regulating their learning. Additionally, it might be advisable to include explicit scaffolds in the RPT learning materials, which can directly prompt students' deep-level regulation during all phases of problem-solving. Such scaffolds might particularly be relevant for evoking orientation, evaluation, and planning, given that these regulation skills remained rather low-level in the current RPT intervention, which was characterised by open-ended RPT assignments, not directing students' regulation. Prompting students' deep-level regulation approach in future research could enhance their regulative engagement in this respect, allowing students to refine and spontaneously adopt a deep-level approach when applying diverse regulation skills.

Limitations and recommendations for future research

Since the present study was conducted in a naturalistic university setting, it was, due to ethical reasons, preferable not to randomly assign students from the same class group to either the experimental or control condition. Although two control groups were involved, these were not completely comparable due to differences in participants' background or their university curriculum. Consequently, caution is needed when interpreting the significant changes in students' metacognitive regulation. Alternative explanations for EG students' increased regulation can be found in both students' cognitive gains and the emphasis on self-regulation in the course "Instructional Sciences". Empirical research demonstrated that students' metacognitive regulation is correlated to their cognitive actions and performance (Greene and Azevedo 2009; Zimmerman and Schunk 2011). Students with higher levels of general and domain-specific knowledge are expected to demonstrate a larger involvement in regulation, often resulting in better performance. Moreover, high levels of knowledge and academic experience are assumed to positively influence the quality of learners' metacognitive skills (Chin and Brown 2000). This study was set up in relation to the course "Instructional Sciences" introducing students in theories about learning and instruction, including self-regulation. The particular course context might have enhanced EG students' problem-solving awareness, resulting in a theory-driven execution of the think-aloud task. It should be noted, however, that metacognitive gains based on students' knowledge of "Instructional Sciences" could equally be expected from CG1 students, since the course was also a formal part of their curriculum. Nevertheless, EG students' increased regulation at posttest was not reflected in the regulation behaviour of CG1 students.

Furthermore, the assessment of students' metacognitive regulation was exclusively based on their verbalised problem-solving. It can be assumed, however, that students do not always explicitly articulate their thinking and regulation (Meijer et al. 2006), implying that the identification of metacognitive utterances in the think-aloud protocols might not have been exhaustive. Another limitation of think-aloud protocol analysis concerns the risk of reactivity (Greene et al. 2011), since asking students to verbalise can increase their attention to their cognitive processing, possibly resulting in more regulation than they would spontaneously demonstrate. Data triangulation through multiple concurrent assessment techniques might be more preferable in future research (Meijer et al. 2006).

Although the results suggested a positive impact of RPT, we only reported short-term effects. Long-lasting interventions are needed to guarantee an enduring impact. Furthermore, the time-consuming nature of think-aloud protocol coding only allowed for data analysis on a relatively small sample (Roscoe and Chi 2008). The study was moreover conducted in a particular setting, with students studying a specific course. Future research preferably involves other student populations, alternative instructional settings, or other tutoring formats to increase the representativeness of the findings.

It should further be noted that although the RPT intervention successfully elicited students' metacognitive monitoring, its effects on the adoption of orientation and evaluation, as well as of deep-level metacognitive regulation (i.e. monitoring and orientation) were less prominent. This might be related to the format of the think-aloud tasks and RPT assignments (Greene et al. 2011; Perry and Winne 2013). Both might have stimulated students to particularly check their comprehension and progress but might have been less appropriate to evoke (and assess) other regulative acts. Future research with alternative task formats is needed to examine whether more changes can be discerned in students' adoption of specific regulation skills and approaches after participation in RPT.

Given the observation that the impact of RPT might be stimulated by specific collaboration patterns within particular tutoring groups (Chi et al. 2001), future process-oriented investigations could also be promising (Roscoe and Chi 2008). Analysis of videotaped RPT sessions could moreover complement the present findings. Directly observing learning and regulation could unravel when, how, and to what extent metacognitive regulation skills are adopted by RPT participants and consequently help to explain the potential metacognitive effectiveness of RPT suggested in the current study.

Conclusion

Since the promotion of metacognitive regulation requires explicit modelling and guided practice, increasing student–staff ratio's challenge university instructors to successfully support students' regulation (Topping 1996). The present study demonstrated that organising RPT could be a valuable alternative. A RPT setting concerns a small-scale learning environment and consequently allows for intensive metacognitive modelling by peers and individualised feedback on internalised regulation skills. The results of the present study moreover demonstrated that RPT has the potential to foster university students' metacognitive regulation, more specifically students' monitoring. Despite students' additional enhanced evaluation and orientation, RPT was considerably less influential towards these regulation skills. The effects regarding students' adoption of deep-level regulation were rather limited as well, given that RPT mainly elicited deep-level comprehension monitoring. Based on our findings, we recommend the implementation of RPT in higher education in order to promote students' metacognitive regulation, more specifically their adoption of (deep-level) monitoring strategies. Additionally, we advise instructors to carefully train students for RPT and to design appropriate learning materials which can elicit regulative acts.

The present study not only has the potential to inspire educators when aiming to foster students' metacognitive regulation, and it also offers empirical insights enhancing our understanding of students' adoption of particular regulation skills. Our findings furthermore raise questions concerning which elements in the RPT setting specifically evoke metacognitive regulation and therefore present new process-oriented research directions to explore regulation in collaborative groups in depth.

References

Brown, A. L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65–116). Hillsdale: Laurence Erlbaum Associates.

Cheng, Y. C., & Ku, H. Y. (2009). An investigation of the effects of reciprocal peer tutoring. *Computers in Human Behavior, 25*, 40–49.

Chi, M., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471–533.

De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional Science, 40*, 559–588.

Chin, C., & Brown, D. (2000). Learning in Science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*, 109–138.

Duran, D., & Monereo, C. (2005). Styles and sequences of collaborative learning in fixed and reciprocal peer tutoring. *Learning and Instruction, 15*, 179–199.

Falchikov, N. (2001). *Learning together. Peer tutoring in higher education*. London: Routledge Falmer.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of sophisticated mental models. *Contemporary Educational Psychology, 34*, 18–29.

Greene, J. A., Robertson, J., & Croker Costa, L. A. (2011). Assessing self-regulated learning using think-aloud methods. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 313–328). New York: Routledge.

King, A. (1997). Ask to think-tell why©: A model to transactive peer tutoring for scaffolding higher level complex learning. *Educational Psychologist, 32*, 221–235.

Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation, 12*, 209–237.

Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning, 4*, 197–216.

Nota, L., Soresi, S., & Zimmerman, B. J. (2004). Self-regulation and academic achievement: A longitudinal study. *International Journal of Educational Research, 41*, 198–215.

Perry, N. E., & Winne, P. H. (2013). Tracing students' regulation of learning in complex collaborative tasks. In S. Volet & M. Vauras (Eds.), *Interpersonal regulation of learning and motivation: Methodological advances* (pp. 45–66). London: Routledge.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407.

Roscoe, R. D., & Chi, M. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science, 36*, 321–350.

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modelling. *Reading and Writing Quarterly, 23*, 7–25.

Topping, K. J. (1996). Effective peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32*, 321–345.

Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs. domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*, 187–209.

Veenman, M. V. J., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science, 33*, 193–211.

Volet, S., Vauras, M., & Salonen, P. (2009). Self- and social regulation in learning contexts: An integrative perspective. *Educational Psychologist, 44*, 215–226.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 279–306). Hilsdale: Erlbaum.

Zimmerman & Schunk. (2011). Self-regulated learning and performance: An introduction and an overview. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). New York: Routledge.