Australian Journal of Teacher Education

Volume 47 | Issue 10 Article 5

2022

Using Formative Assessment to Build Coherence Between Educational Policy and Classroom Practice: A Case Study Using Inquiry in Science

Connie Cirkony
University of Tasmania

John Daniel Kenny
University of Tasmania

Recommended Citation
Cirkony, C., & Kenny, J. D. (2022). Using Formative Assessment to Build Coherence Between Educational Policy and Classroom Practice: A Case Study Using Inquiry in Science. *Australian Journal of Teacher Education, 47*(10).
http://dx.doi.org/10.14221/ajte.2022v47n10.5

Using Formative Assessment to Build Coherence Between Educational Policy and Classroom Practice: A Case Study Using Inquiry in Science

Connie Cirkony
John Daniel Kenny
University of Tasmania

Abstract: In this paper we argue that the complexity of education systems can lead to a lack of coherence in the implementation of policy. More effective educational change requires policy-makers and researchers to pay more attention to supporting teachers in classrooms. As an example, we consider decades of research attempts in STEM education to implement learning through inquiry and note there has been little change in teaching practices in classrooms. Using formative assessment in science education as a case study, we developed a rubric for teachers that embeds key aspects of the desired pedagogy. We argue this builds teachers' confidence to implement the change in their classrooms and we claim this same principle may apply to the development of rubrics in other disciplines, which can easily be incorporated by researchers into professional learning programs.

Keywords: teacher change, assessment, representational competence, teacher support, education policy.

Introduction

Globally, the last decade has seen increased development in evidence informed policy and practice in a range of fields, including education (Boaz et al., 2019). Yet, the K-12 school education sector has been described as one with comparatively low levels of investment in research, and poor links between policy, research, and innovation (Burns & Schuller, 2007).

Cain (2019) and Kania et al. (2018) suggest that an education system can be described as a *complex adaptive system*, with an interconnected set of subsystems, each with their own purpose and practices that can impact on the system as a whole. Such systems involve numerous stakeholders, with different perspectives, uneven power dynamics, which makes the outcomes of any given change process uncertain (Checkland, 2012). Change processes require two-way communication and information flows. Checkland (2012) argued that to effect change in such systems, there is a need to focus on action learning as the basis of the change process. Education researchers will need play an important role in enabling this process.

In education, decisions are often made at a policy level, where the parameters driving change can be far removed from what teachers do in classrooms. Levin (2010) and others have argued that this situation is perpetuated by top-down policy-making decisions which preferences large scale quantitative studies, which do not necessarily capture the lived experiences of other key stakeholders, especially the teachers (Hargreaves & Fullan, 2012). This contrasts to an approach which seeks to build *coherence* between policy and practice (Kenny & Cirkony, 2022; Hargreaves & Fullan, 2012) by providing opportunities for the teachers, who are expected to implement the policy in their classrooms, to provide feedback based on their experience and influence its on-going development. Coherence requires alignment of policies, standards, curricula, and assessments – with the active participation of those who implement policy (Heredia, 2020; Penuel et al., 2008). Further, Dolin et al. (2018b) maintained education researchers have contributed to this situation because "it is not the norm for educational research projects to explicitly address policy issues" (p.251), yet they clearly have an important role to play in educational change.

As an example of a policy which has lacked coherence, we will consider Science, Technology, Engineering and Mathematics (STEM) Education. STEM policy has been implemented in many educational systems around the globe and has relevance for K-12 education and tertiary education. However, STEM policy has primarily been driven by national economic development and industry concerns (DeCoito, 2016; Marginson et al., 2013; Sharma & Yarlagadda, 2019; Timms et al., 2018), not research on teacher change. Thus, STEM policy has paid insufficient attention to coherence by assuming that "teachers have, or can develop, expertise in up to four disciplines along with the ability to integrate student-centred learning across these disciplines" (Honey et al. 2014; Kenny & Cirkony, 2022, p.90). This assumption is contrary to research pointing to the inherent teacher change issues involved, and that making connections across disciplines "can be counter-productive" if the ideas in each domain have not been securely learned (Harlen, 2015, p.13). Further, it magnifies issues regarding the many challenges science teachers face with implementing innovations, including inquiry-based teaching and learning (Duit & Treagust, 2012; Fitzgerald et al., 2019), alongside ways to assess these disciplinary-specific approaches.

So, despite decades of advances in research knowledge on developing teachers' expertise (e.g., Brand & Moore, 2011; Cirkony et al., 2022; Garet et al., 2001; Luft et al., 2015) there has been little change in teaching practices in science classrooms, including inquiry-based approaches. Traditional teaching and assessment practices continue to dominate (Duit & Treagust, 2012; Fitzgerald et al., 2019) and student engagement has either continued to fall or remained static in many countries, such as Australia and India (Sharma &Yarlagadda, 2018), as well as Canada (DeCoito, 2016).

This lack of coherence between STEM policy and classroom practice as a likely cause of this failure to translate education research into practice. For effective change in a complex educational system, policy-makers and researchers need to pay more attention to supporting teachers in classrooms. More specifically, Harlen (2015) called for clearer connections between educational policy and the work of science teachers, noting "in many cases a change in policy is required so that innovation is not stifled by existing practices" (p.52).

Honey et al. (2014) added "it is challenging to design assessments that are effective for both discipline-specific and integrated learning," and that "[l]arge scale assessments used for accountability pose the biggest challenges" (p.110). They called for assessment in STEM to be "firmly grounded in research from the learning and measurement sciences" with a shift "to tasks that require integration of reasoning and inquiry in the context of significant, applied problems. In essence, Honey et al. (2014) argue for greater coherence at a systemic level to address the difficulties associated with STEM, including a re-think of the assessment policies.

Assessment as a Vehicle to Build Policy Coherence

Policy-makers tend to prefer summative or standardised forms of assessment, chiefly as an accountability measure (Dolin et al., 2018b; Honey et al. 2014; Levin, 2010). However, these can introduce a lack of coherence in assessment practices, which ultimately works against real teacher change and successful classroom innovation (Cowie & Harrison, 2021; Dolin et al. 2018b). Harlen (2015, p.9) identified assessment as a common source of incoherence in science education:

> *Since what is assessed and reported is assumed to reflect what it is important to learn, it is essential that this is not limited to what can be readily tested. A range of methods should be used to gather and interpret evidence of learning so that students are able to show what they can do in relation to all types of goals.*

Since teacher expertise is developed in a context which is directly relevant and linked to their working environment (Schulman, 1986), assessment can be a way to help teachers to change their practice (Cowie & Harrison, 2021). As opposed to relying on large-scale quantitative research to inform policy, action-learning based Professional Learning (PL) approaches and partnerships, which occur close to classroom contexts, have been shown to support teacher change (Brand & Moore, 2011; Garet et al. 2001; Hargreaves & Fullan, 2012; Luft et al., 2015). PL partnerships enable researchers to work closely with teachers in classrooms to effect change (Cain, 2019; Jones et al. 2016), thereby enabling more coherent policy implementation. This is consistent with the requirements to effect change in complex adaptive systems (Checkland, 2012).

In this case study, we investigate this link between policy and classroom practice through assessment. In the context of Year 9 science, we trialed an assessment rubric to support science teachers to implement a disciplinary-specific pedagogical innovation in their classroom practice and address the following research question:

> *What potential does a rubric that embeds inquiry-based pedagogy have as a support for teachers to implement the pedagogy in their classrooms and to improve coherence between policy and practice?*

We then considered the research implications of these findings and a more general principle, in which supporting teachers with a formative assessment rubric as a way to improve the implementation of systemic educational change and thereby the coherence of an educational reform policy.

Building Coherence in Educational Reform Through Assessment

The principles of educational reform proposed by Kenny and Cirkony (2022) suggest a coherent STEM policy should ensure educational change addresses the practical and pedagogical issues teachers face when implementing inquiry-based teaching approaches. They highlight the importance of addressing the needs of teachers in classrooms and are based on the premise that teachers' expertise is built through action learning PL approaches, with topics relevant to and enacted close to the context of their practice (Faulkner et al., 2019; Garet et al., 2001; Jones et al. 2016; Luft et al. 2015).

Duit and Treagust (2012) called for researchers to consider how to encourage teachers to implement research-based science inquiry teaching ideas in their classrooms. Grangeat et al. (2021) explicitly mention assessment as a pathway to this outcome. Formative assessment is particularly suited for inquiry-based interactive, student-centred teaching; however, these approaches are heavily dependent on teacher expertise and situated in classroom contexts (Cowie & Harrison, 2021; Dolin et al, 2018a; Waldrip et al., 2010).

Importantly, formative assessment needs to support students to understand and be able to explain scientific phenomena instead of just providing the correct responses (Cisterna & Gotwals, 2018). Given the increased focus on formative assessment in the implementation of science inquiry-based reform, there is a need to illustrate what formative assessment looks like in science classrooms (Cisterna, & Gotwals, 2018; Heredia, 2020).

This suggests that appropriately designed PL and assessment resources that empower teachers to implement these inquiry pedagogies, may be effective at promoting the shift to genuine inquiry in science classroom, which has eluded policymakers and science education researchers for decades.

In this study, we set out to support teachers to use formative assessment as a means to build their expertise to implement genuine inquiry practices in their classrooms, by developing a rubric that embeds key aspects of inquiry-based pedagogy.

Background to this Case Study

The ideas in this paper originated from work on two separate PL projects, Project A and Project B, on which the authors were working independently. These projects had in common the introduction of participants to a Representation Construction Approach (RCA) to learning science (Cirkony, 2019; Kenny et al., 2020) as a form of guided-inquiry. Our common experiences brought us together to jointly develop resources to support the teachers (Kenny & Cirkony, 2018a, 2018b).

RCA combines a guided-inquiry approach with interactive representation-rich learning experiences that seeks to emulate a disciplinary specific way of teaching and learning (Tytler et al., 2018). RCA aims to engage students in inquiry through activities in which they generate their own representations to explain their ideas about the phenomenon under study. In developing their Student Generated Representations (SGRs), students are encouraged to make use of a range of representational forms including linguistic, visual, physical and symbolic, to develop and explain their understanding of scientific phenomena. These may include use of everyday language and/or specific technology, text, dialogue, drawings, role plays, 2D and 3D physical models, animations, gestures (Daniel et al., 2018; Knain et al., 2021).

Constructing and critiquing the SGRs enables students to experience how knowledge is built in science, through a creative and collaborative process of refining ideas, reasoning, justifying, and evaluating their own claims or those of others (Ainsworth et al., 2011). When students create their own representations or make choices about which representations are best suited to a given purpose, they become more actively engaged in their own learning (Ainsworth et al., 2011). These activities are believed to activate visual-spatial thinking and ground their actions to thoughts (Goldin-Meadows & Beilock, 2010) and provide a direct and powerful way for students to explain abstract ideas and develop their reasoning processes (Stieff et al., 2005; Tytler et al. 2013). RCA has been gaining attention in the literature because it contributes to quality learning and improved student engagement (Ainsworth et al., 2011; Knain et al., 2021; Tang et al., 2019; Tytler et al., 2013; Xu et al., 2021).

The RCA learning process involves interactive discussions about the SGRs, how well they represent the phenomenon under study, and their strengths and weaknesses. Thus, the process promotes high-level cognitive activity, as students' ideas are challenged, refined and justified based on evidence (Lemke, 2004). Through this process, students can not only build their conceptual understanding, but also develop their representational competence (diSessa, 2004), along with an appreciation of the epistemological basis of how scientific knowledge develops (Tytler et al., 2018; Wilson & Bradbury, 2021; Xu et al., 2021).

Building teacher expertise to plan and assess learning through RCA is crucial for successful implementation in classrooms. However, the interactive nature of RCA presents significant challenges for teachers (Cowie & Harrison, 2021; Dolin et al. 2018a; Duit & Treagust, 2012; Kenny & Cirkony, 2018a, 2018b; Knain et al., 2021; Tytler et al., 2013; Xu et al., 2021). To explore both the teacher and the student experience of RCA assessment, we draw on the following two projects.

Project A involved the second author and a group of 23 secondary teachers of science in Tasmania, Australia (Kenny et al., 2020). The author delivered two iterations of an extended PL program over 10 weeks in 2017-18. Teachers were supported to adapt their individual teaching programs to implement RCA in their classrooms. The participating teachers taught in different schools, therefore the science topics and year levels taught varied for each participant.

Project B involved the first author who was investigating the quality of students' learning in response to RCA implementation in a Year 9 science classroom (Cirkony, 2019). There were 27 students were from a private girls' school in Melbourne, Australia, approximately 14 years of age. The mid-career science teacher had undergone PL for RCA with the research team and jointly developed the unit on Energy from which the case study data came. The investigation took place over 12 lessons over approximately six weeks in 2016. Students completed pre- and post-surveys of their knowledge prior to and after the unit.

Through our common experiences in these respective PL programs, we recognised the extent of the demands RCA placed on the teachers, but also its value to engage students in their learning. This led us to jointly develop a planning framework (see Kenny & Cirkony, 2018a, 2018b). We also recognised the need to provide some practical support for teachers to assess the SGRs as they emerged in their science lessons. We developed a generic rubric for this purpose, and it is this rubric which is the focus of the remainder of this paper (Appendix 1).

A Review of the Research on Assessing Student Generated Representations

While SGRs provide insight into students' current understanding and can inform teachers for subsequent planning, using them to assess students' understanding presents significant challenges. The adequacy of any given representation is difficult to judge because there is no ideal representation; representations need to be understood in relation to a given task (diSessa, 2004). Moreover, no single representation can fully convey conceptual meaning (Lemke, 1998).

RCA is reliant on the ability of teachers to promote an interactive classroom environment, where students engage in the process of refining and re-representing their ideas. Teachers need to be able to make judgements about students' growing understanding and guide their thinking, based on the iterative development of their SGRs (Waldrip et al., 2010).

The literature reveals various attempts by researchers to assess students' learning in representation-focused tasks. Some researchers have relied on qualitative methods such as interviews and observations (e.g., Hubber et al., 2010; Waldrip et al., 2013). Others have measured learning gains by using test instruments with items from known national or international assessments (e.g., Lehrer & Schauble, 2006; McDermott & Hand, 2013).

More recently, researchers have developed frameworks to analyse the characteristics of students' drawings (Park et al., 2020; Tang et al., 2019; Wilson & Bradbury, 2021; Xu et al., 2021). For example, Tang et al. (2019) stated that "teachers need to have a clear framework to recognise representational features and meanings made in students' drawings" (p.2320). They developed detailed framework to analyse students' drawings, which they argued might be extended as an assessment rubric for both teachers and students to evaluate the diagrams. However, the complexity of their framework is likely to make it inaccessible to teachers without significant PL. Further, aside from drawings, it did not consider other forms of representations (e.g., models, role-plays).

While these provide useful tools for researchers to interpret SGRs, our intention was to support teachers in the classroom context.

Development of the Assessment Rubric for Teacher Use

Our goal was to develop an assessment tool teachers could easily adapt and use to facilitate formative discussions with their students about their SGRs and monitor changes in their understanding and their representational competence (Waldrip et al., 2010). This is consistent with other researchers who advocate the value of tools, such as rubrics, to support teachers to implement inquiry practices in their classrooms (Cowie & Harrison, 2021; Grangeat et al., 2021, Herbert et al., 2021).

The authors developed a prototype rubric to trial in Project A. This was refined from this experience into the generic rubric contains criteria concerned with formatively assessing students' conceptual understanding from their representations (Appendix 1). The need for a generic rubric to support the teachers was highlighted by the fact that the teachers who participated in Project A taught different year levels and topics. The teachers were also at different schools, so the researcher was unable to attend all classroom lessons. Therefore, the rubric had to be able to be adapted by the teachers to suit a range of teaching situations.

The design of the rubric draws on work by both diSessa (2004) and Kozma and Russell (2005), who put forward the notion of *Representational Competence* (RC), defined as an individual's ability to represent scientific ideas using various representational forms. Each devised developmental scales for which described the students' ability to generate representations, with increasing sophistication evident through the use of multiple representational forms, increasing use of disciplinary symbols and conventions to explain their understanding of scientific concepts.

Kozma and Russell (2005) provided a 5-level framework to classify this progression. Where Levels 1–3 focused on superficial and realistic features, as typically generated by novices, and Levels 4–5 described more the abstract, relational, and reflective use of representations, characteristic of experts. These researchers argued that, with guidance, as students developed their ability to use their SGRs to explain their conceptual understanding, they could also develop their RC.

For the generic rubric, we reasoned the structure should be one that teachers and their students find familiar. Each criterion is assessable at five levels and assigned a numerical value for scoring purposes, where Level 1=1 point to Level 5= 5 points. The generic criteria for *conceptual understanding* are designed to be adapted to specific concepts. The three remaining criteria aim to integrate conceptual understanding with RC by making direct reference to the phenomenon under study and emphasise that the purpose of the representational task is to convey meaning.

The *nature of the representation* refers to the level of sophistication of the SGR(s), considering how well it can explain the phenomenon, including reference to less observable features, and use of symbolic representations and conventions to convey meaning (e.g., use of arrows to represent sunlight). The *coordination of different representational forms* is concerned with the range of representational forms (linguistic, diagrammatic, gestural, etc.) and how well these are coordinated to explain the phenomenon (e.g., text used as labels to explain the parts in a diagram). *Application of Representation* is designed for tasks where students are asked to demonstrate their RC *and* apply their conceptual understanding to explain the outcome of an investigation or solution to a real-world problem.

Methodology

To test our rubric against real data, we selected the following question from the pre-and post-test in Project A: “Scientists make references to the Greenhouse Effect and its relation to climate change. What is the Greenhouse Effect?"

This question required an open-ended response, in which the students were expected to provide some form of representation to explain their understanding. In contrast to the other open-ended questions on the tests, it did not include any expert generated diagrams that might influence students’ choice of representations, so, it was more likely the students would demonstrate their RC as they attempted to explain their understanding of the question.

We adapted the rubric to suit this question by replacing the generic conceptual criteria with three criteria directly addressing the learning outcomes for the Greenhouse question (Appendix 2):

1. Light energy from the Sun is absorbed by the Earth. As the Earth warms, some of this energy is re-radiated as heat/infrared energy.
2. Certain gases in the atmosphere (e.g., carbon dioxide, nitrous oxide, methane) trap this heat energy resulting in a natural Greenhouse Effect.
3. This natural Greenhouse Effect maintains an average Earth temperature that sustains life.

That the three RC criteria did not need to be changed to suit the question, provided the insight that RC maybe a transferrable skill, as discussed further below.

The students’ responses on the pre- and post-test item can be found in Appendix 3. Pseudonyms have been used. As Project A was completed by this time, the analysis was done post-hoc. We purposefully selected 10 students from the 27 in the original class, because they had also participated in a video-stimulated interview following the post-test. This would provide further qualitative data about their ideas and representational reasoning, to assist with assessing their SGRs.

Analysis

To ensure inter-rater reliability, the authors independently applied the rubric to assess the SGRs, then compared results and discussed any discrepancies to jointly decide on scores for each student. The maximum score possible for the six criteria was

30 points, with 15 related to the conceptual and 15 related to the RC criteria (Appendix 2). These agreed raw scores are represented in Table 1 and Table 2

Table 1 shows the overall total score out of a possible 30 points for the test items and the score for conceptual understanding for each student. Table 2 shows the differentiated scores for the RC criteria. From these tables, the overall average gain of 2.4 points consisted of an average of 1.1 points gain in conceptual understanding and an average gain in RC of 1.3. A non-parametric Mann-Whitney U test was used to test for statistical significance due to the small groups size and non-parametric data. It showed the gains in the overall test scores were significant at the 5% level ($p<0.05$) with $p=0.009$. [1]

Six students showed improved conceptual understanding and seven showed improved RC. Only four students, Amel, Jessica, McKinley and Nyla showed improved scores in both sets of criteria. Students with the greatest gains were Amel (+6 points), Jessica (+4 points) and Nyla (+4 points). Two students, Elham and Annisa, showed no overall change in scores.

We then considered each criterion separately. While we found a significant gain in conceptual understanding ($p= 0.04$), the individual RC criteria scores showed no significant change, with Nature ($p= 0.10$), Forms ($p=0.31$) and Application ($p= 0.14$).

Name of student	**Pre-test score**	**Post-test score**	**Change in overall scores**	**Conceptual Understanding**		**Conceptual Change**
				Pre-test	**Post-test**	
Amel	7	13	+6	4	6	2
Clara	9	11	+2	4	4	0
Jessica	10	14	+4	4	7	3
Phoebe	8	10	+2	4	4	0
Komal	10	12	+2	4	6	2
McKinley	8	11	+3	3	5	2
Nyla	6	10	+4	3	4	1
Sofi	6	8	+2	3	3	0
Elham	9	9	0	4	4	0
Annisa	8	8	0	3	4	1
Average score	**8.1**	**10.6**		3.6	**4.7**	
Average gain			**2.5**			**1.1**

Table 1: Pre- and post-test results for overall score and conceptual understanding.

Amel's 6-point gain consisted of a 2-points in conceptual understanding, and a 4-points in her RC. By contrast, three of Jessica's 4-point gain consisted of 3 points for improved conceptual understanding and 2 points for improved RC. Nyla's four points gain showed the reverse, with 1 point for conceptual gain and three for improved RC. Clara's gain was limited to her RC, with no apparent gain in conceptual understanding. Phoebe results showed a similar pattern, while Elham showed no gain in either her conceptual understanding or RC. Annisa showed some improvement in her conceptual understanding, but this was cancelled out by a reduction in her RC score.

While this analysis appears to support the claim by diSessa (2004) that there is not necessarily a link between the students' conceptual understanding and their representational competence, we note this may also be linked to limitations of using a static test item as a way to assess RC (Cirkony, 2019; Xu et al. 2021).

[1] Link to statistical tool used: https://www.socscistatistics.com/tests/mannwhitney/default2.aspx

Name of Student	Nature of Representation		Coordination of Representational Forms/Modes		Application of representation		Change in RC
	Pre-test	**Post-test**	**Pre-test**	**Post-test**	**Pre-test**	**Post-test**	
Amel	1	2	1	3	1	2	4
Clara	1	2	1	2	2	2	2
Jessica	2	3	2	3	2	2	2
Phoebe	1	2	1	1	2	2	1
Komal	2	2	2	2	2	2	0
McKinley	2	2	2	2	1	2	1
Nyla	1	2	1	2	1	2	3
Sofi	1	2	1	1	1	1	1
Elham	2	2	1	1	2	2	0
Annisa	2	1	2	1	1	2	-1
Average score	**1.5**	**2.0**	**1.4**	**1.8**	**1.5**	**1.9**	
Average gain		**0.5**		**0.4**		**0.4**	**1.3**

Table 2: Scoring for representational competence on climate change question.

The pre- and post-test items provided two 'snapshots' of the students' understanding, before and after the learning activities, in the form of their SGRs. However, they provided no information on the learning process that had occurred between the tests. It was difficult to assess overall changes in students' understanding using the diagrams alone. Xu et al, (2021) encountered similar difficulties with assessing SGRs on test items. This challenge consistent other researchers who point out that "students' drawings (a visual mode) must not be interpreted in isolation from other modes of representation, particularly speech or writing" (Tang et al, 2019, p.2319; Xu et al., 2021). To gain further insights about their SGRs, we examined the interview data for each of the ten students.

Amel had achieved the highest learning gain on our analysis so far. Her post-test representation had shown a transition from a single mode (i.e., text) to an annotated multimodal representation. During the interview, Amel elaborated further:

> *So [the radiation] comes from the Sun and some of it, like goes into the Earth's atmosphere. And then they get rebounded because of the greenhouse gases… this Greenhouse Effect it kind of traps the heat in it and like they go back in because the Greenhouse Effect keeps them in. And that causes Earth to warm up, which is like global warming.*

Because she clarified the meaning of the black line in her represented post-test SGR, we adjusted her RC score for the *nature of representation* from 2 to 3, raising her overall score from 13 to 14.

While Clara's post-test scores indicated no improvements in her conceptual understanding, her SGRs transitioned from purely text to an annotated diagram with some labelling and symbolic forms, so her RC score had improved overall. During the interview, she elaborated further on her SGR by explaining the meaning of the relevant symbols: "So this is the Sun and then, the heat coming towards the Earth. And this is the greenhouse gases and then it traps the Sun's heat, so the Earth becomes more hotter." Because she clarified the meaning of the yellow lines in her SGR as representing energy from the Sun, we adjusted her *nature of representation* score from 2 to 3, resulting in an adjusted overall score of 12.

Similarly, Jessica's post-test SGR indicated a gain in conceptual understanding and improved RC. It included coordination between text and symbolic representational forms. During her interview, she added further to the meaning of her post-test diagram:

"I was saying the Sun is directing heat towards the Earth. And some of it reflects… off. But some of it sort of sits between the layer of the Earth and the Universe? The atmosphere? I'm not really sure." Although not a confident response, we adjusted the *nature of representation* score from 3 to 4, raising her overall score from 14 to 15.

Phoebe's post-test SGR showed little change in conceptual understanding while her RC showed some progression from a single text-based SGR to a multimodal symbolic representation. However, her diagram lacked annotations and showed no attempt at coordination between the representational forms. During her interview, Phoebe was able to elaborate that the arrows represented the Sun's rays and that "the rays from the Sun go into the atmosphere. The Sun[light] reflects back out and others warm up the Earth." This information on the meaning of her post-test SGR justified a higher *nature the representation* score, raising her overall score from 10 to 11.

Komal indicated some growth in conceptual understanding but because she gave essentially the same representation in both tests, there was no change in her RC score. During the interview, she indicated some confusion in her explanation: "not all the heat transfers into the Earth, and more comes out than goes in. But I'm not sure about that, though." So, her overall score remained unchanged.

McKinley's post-test response had also indicated some growth in conceptual understanding and RC through her more detailed text-based explanation, and inclusion of symbolic forms to explain the warming of the Earth due to trapping the Sun's energy. During the interview, she revealed a persistent alternative conception about the role of the ozone layer in climate change when she explained that "the ozone layer traps heat inside with the Sun…but then the ozone layer helps to keep it still warm so [the heat is] not just bouncing back off." Her overall score therefore did not change, and this alternative conception was identified as something that would need to be addressed in future lessons.

Nyla's text-based explanation showed a small gain in conceptual understanding and an improved RC score. Her pre-test diagram was largely a depiction of observable features and did not include any annotations. Her post-test SGR was more symbolic and included text but lacked important features such as a representation of the atmosphere. During her interview, Nyla confirmed her limited understanding of the concepts and persistent alternative conceptions in her explanation that "the greenhouse gases come up into the clouds and the ozone layer" and "when the Sun goes down, all the heat will rise back up, so it gets colder at night." Her overall score remained the same, but again the interview revealed an alternative conception that needed to be addressed.

Sofi's representations showed no change in her conceptual understanding, and a persistent alternative conception about the ozone layer. There was some growth in her RC as her initial SGR, which was largely a depiction of observable features, but also included some non-observable features such as linking her depiction of cows with the production of greenhouse gases. Her post-test representation used more symbolic forms to convey these ideas. During the interview she conceded that text in both SGRs said "pretty much the same thing" so her score did not change.

Elham's responses showed no gain in either conceptual understanding or RC. Although her representations included observable and non-observable features, such as land mass, water, and greenhouse gases, there were no linkages or annotations and little difference between the pre- and post-test diagrams. During her interview, she clarified that the red line represented the "blanket" of carbon dioxide but was not able to elaborate further. Her score did not change.

Finally, Annisa showed some limited growth in her conceptual understanding, largely attributed to her not mentioning ozone in the post-test. Her pre-test representation was largely a depiction of observable features, but incorporated text, and some symbolic features to show non-observable greenhouse gases. Curiously, her RC score fell as her post-test SGR consisted of a text-only representation which was less sophisticated than her initial diagram. During her interview, she elaborated on the role of radiation in the Greenhouse Effect:

> *Because [its] part of our daily lives, you don't notice it because [its], everywhere. And like just what they give off, especially with radiation, it just goes into the atmosphere and combines with all the other greenhouse gases and like release and create climate change.*

When prompted to elaborate on how radiation goes into the atmosphere she responded: "I'm not sure…I don't know how to explain it." This confirmed her lack of conceptual understanding about the Sun as a source of radiation and the Greenhouse Effect, so her overall score did not change. The results of this additional analysis are shown in Tables 3 and 4, along with the adjusted marks.

Name of Student	Initial overall score	Overall score post interview	Gain
Amel	13	14	+7
Clara	11	12	+3
Jessica	14	15	+5
Phoebe	10	11	+3
Komal	12	12	+2
McKinley	11	11	+3
Nyla	10	10	+4
Sofi	8	8	+1
Elham	9	9	0
Annisa	8	8	0
Average	**10.6**	**11.0**	**2.8**

Table 3: Change in students' post-test scores following the interviews.

The interviews with students about their SGRs enabled us to gain more insight into their understanding or confirm our original assessment. As a result, the score for the first RC criterion for Amel, Clara, Jessica, and Phoebe increased by one in each case based on the recorded conversation. Consequently, the average score for the group was raised from 2.0 to 2.4 and the gain for this RC criterion was retested and now found to be significant (p=0.026). The overall average RC score remained unchanged.

Name of Student	Nature of Representation		Coordination of Representational Forms/Modes		Application of representation		Change in RC
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	
Amel	1	3	1	3	1	2	4
Clara	1	3	1	2	2	2	2
Jessica	2	4	2	3	2	2	2
Phoebe	1	3	1	1	2	2	1
Komal	2	2	2	2	2	2	0
McKinley	2	2	2	2	1	2	1
Nyla	1	2	1	2	1	2	3
Sofi	1	2	1	1	1	1	1
Elham	2	2	1	1	2	2	0
Annisa	2	1	2	1	1	2	-1
Average score	**1.5**	**2.4**	**1.4**	**1.8**	**1.5**	**1.9**	
Average gain		**0.5**		**0.4**		**0.4**	**1.3**

Table 4: Students' post-test scores for representational competence following the interviews.

Discussion

The aim of this paper was to explore how an assessment tool supports teachers to implement inquiry in their science classrooms, as way to improve coherence between policy and practice. We address this aim first by discussing the findings, then by relating the findings in this case study to the broader issue of policy implementation in STEM and other learning areas.

From our experience with the rubric, we found it relatively easy to adapt to this topic and to assess the students' understanding of the concepts related to climate change, based on their responses to this question. Having explicit criteria to judge RC helped us to focus on the salient features in their representations. For this post-hoc analysis, the recorded interviews helped to clarify meaning. Some students used both verbal and gestural modes of communication to elaborate on aspects of their diagrams, consistent with other studies using representations (e.g., Hackling et al., 2013).

Using the rubric to assess the students' work revealed a statistically significant conceptual gain for the group, but it was weak. This is consistent with Xu et al. (2021, p.862) who had similar findings, noting "that interpreting students' intended and realised meanings was not straightforward" on pre and post-test items. This exercise also revealed the limitations with using static drawings to assess the RC components of student learning, which was also consistent with Xu et al. (2021). In our case, use of the interview data enabled us to adjust the RC scores for four students. For the six other students, the interviews confirmed our initial assessments, but also revealed the persistent alternative conception of ozone as being involved in the greenhouse effect for two students.

Nevertheless, given this unit was intentionally designed for RCA, and the teacher had attended PL, the low levels of conceptual gain were concerning. In search of an explanation, we turned to the observational notes made during the original project. These revealed that, although the students had been encouraged to produce SGRs as part of the learning, there had been minimal teacher-led discussions of their SGRs (Cirkony et al. 2022). This observation revealed a serious deficiency in the way the RCA teaching process had been implemented. As Kozma and Russel (2005, p.134) argued: "representational skills can best be developed and used within the context of student discourse and scientific investigations. The use of language and representations during the investigative process is also more likely to lead to a deeper understanding".

Further, as diSessa (2004), pointed out, using pictures to convey meaning for a complex idea is not easy. In other words, crucial aspects of the RCA pedagogy had not been applied in Project B classroom. The students had little opportunity to compare and discuss their SGRs with others, how they might be improved to better explain the complex and abstract concepts involved in the phenomenon of climate change (e.g., radiation). This is despite having an experienced teacher who had attended PL on RCA, and who had been assisted by researchers in the planning phase. Lack of discussion would have also seriously limited opportunities for the development of their RC and conceptual understanding.

Given the difficulties teachers face with implementing RCA are already identified in the literature (Tytler et al., 2013; Xu et al. 2021), this observation is not meant as a criticism of the teacher. This suggests a need for more explicit support for teachers to implement a highly interactive pedagogy. We argue that the addition of this rubric to the PL sessions would have supported this teacher to better facilitate the discussion of the SGRs as part of the RCA learning process.

Other researchers also emphasise that teaching using RCA is enhanced by dialogue to make meaning through the SGRs (Stieff et al., 2005; Tytler et al., 2013; Xu et al., 2021). Central to the learning process, dialogue needs to be explicitly planned for and assessed (Kenny & Cirkony, 2018a, 2018b; Cirkony et al., 2022). The formative assessment process is, therefore, a key component of the expertise teachers need to master this pedagogy (Waldrip et al. 2010) and is likely key to their effective use of the approach in science classrooms (Duit & Treagust, 2012; Hubber et al., 2010; Luft et al., 2015).

Even in non-RCA science classrooms, Cisterna and Gotwals (2018) highlighted the complexities inherent in working with students' ideas in the moment. They found that teachers were able to use questioning strategies to elicit student ideas but struggled with responding to exploring them for deeper understanding or to progress learning. This is also consistent with recent research on assessment which identified the need for teachers to be supported when introducing new approaches to assessment and that tools, such as rubrics, can be effective in this process (Cowie & Harrison, 2021; Grangeat et al. 2021; Nielsen et al., 2018).

Further, as we prepared this paper for submission, Herbert et al. (2021) published an article independently of our work. They reported on the efficacy of a rubric to support primary teachers of mathematics to implement mathematical reasoning in their classrooms. They concluded their "rubric encouraged them to design and use teacher actions to elicit and extend students' reasoning, as well as provide informal feedback" (p.15). It built the teachers' confidence with the pedagogy and their ability to design appropriate tasks and provided a useful language framework in conversations with students and each other.

We found this work by Herbert et al. (2021) very encouraging and confirmative of our argument that the provision of a rubric for teachers wanting to implement guided-inquiry through RCA would be useful. Further, what also become evident from our experience of developing and testing this rubric is that the RC criteria did not need to change. In other words, these criteria would remain the same from topic to topic, making the rubric easier to use than we had expected.

This last point suggests that RC may be best considered as a transferrable inquiry skill, which could improve with practice, and means the time invested in building students' and teachers' RC skills should pay-off as they become more proficient with this approach (Duit & Treagust, 2012). This would also address *time* as one of the chief concerns teachers have with inquiry-based approaches in science (Fitzgerald et al., 2019).

Although Herbert et al.'s (2021) rubric was designed for mathematics, we noted that it had crucial design features in common with our rubric: both had been carefully structured to embed key aspects of the underlying pedagogy being implemented. Where our rubric embedded key aspects of RCA, Herbert et al. (2021) had embedded key aspects of *mathematical reasoning* into their rubric. This observation is supportive of our belief that there may be a general principle at play here.

Our contention is that, if the teachers are provided with suitable rubrics, and supported in their use, students' learning gains would have likely been stronger because the rubric would have supported the teachers to implement key elements of the RCA pedagogy as in Herbert et al. (2021).

Finally, our work highlights the need for the direct participation of teachers as a necessary part of the feedback needed to ensure coherence between policy and inquiry-based approaches. Penuel et al. (2008) explored how to improve the implementation of education reforms, such as inquiry-based curriculum materials, and highlighted the need to consider how teachers make sense of these in their local context. They suggest that monitoring how teachers apply innovations would better account for the school-level constraints on teacher cognition and action that impact curriculum implementation, providing valuable feedback for improved coherence between policy and practice.

Along similar lines, Heredia (2020) found that considering how teachers make sense of education reforms in their context reflects the complex and dynamic nature of implementation, extending the limitations of policy coherence. Her claim that formative assessment will become an important lever in the implementation of reforms in science instruction suggests our contention that a rubric for teachers will be a valuable tool to support coherence between policy and practice.

Conclusion

At the outset of this paper, we argued that, despite advances in our understanding teacher expertise and how students learn, real change in classrooms has tended to be elusive. We argue, as a general principle that policymakers and education researchers need to consider more carefully how to increase coherence between educational policy, education research and what teachers do in classrooms if educational reform is to be more effective. We pointed to research indicating the ineffectiveness of STEM as policy to promote engagement through inquiry-based practices in science classrooms and suggested this was due to a lack of coherence between current policy and the experience of teachers.

Drawing on the principles of systemic educational reform outlined by Kenny and Cirkony (2022), we focussed specifically on the role of assessment practices to support classroom implementation of innovations (i.e., pedagogies). We reasoned that supporting teachers to formatively assess students' learning should improve coherence between policy that promotes inquiry-based teaching and its implementation in classrooms. We set out to test this idea by means of a rubric designed for science teachers to implement RCA in their classrooms.

As a form of guided-inquiry, this research emphasised the importance of teacher interactivity with students' ideas through formative assessment. The generic rubric was designed to support teachers by embedding key aspects of the RCA pedagogy with criteria to assess students' Representational Competence (RC), along with criteria that could be adapted to specific conceptual content.

Our post-hoc analysis indicated that the generic rubric was easily adaptable to the chosen topic using the expected learning outcomes. Based on our experience, the RC criteria in the rubric should provide useful prompts for teachers to provide feedback on how their students might refine their representations to better explain the phenomenon under study (Waldrip et al., 2010), as well as with their own professional colleagues (Herbert et al., 2021). We believe that it has the potential to provide useful and practical support for teachers to undertake RCA and thereby help them to develop their expertise with this interactive pedagogy (Heredia, 2020; Nielsen et al., 2018; Penuel et al., 2008; Waldrip et al., 2010;).

We conclude that a rubric that embeds key aspects of RCA will support the classroom-based assessment of SGRs. We encourage our science education colleagues to test and validate these claims by adapting the rubric to a range of topics and year level groups and seeking feedback from teachers on its efficacy. This should build our understanding of the pedagogical demands faced by teachers, along with the link between conceptual growth and RC, leading to the refinement of the rubric. Further, its adaptability may be of use to researchers who could include it in PL programs on RCA. This rubric would be improved with more extensive validation in authentic classroom situations, through direct feedback from teachers.

This knowledge should be considered by researchers in the design and/or conduct of future PL programs and other studies exploring the use of RCA in classrooms. We are supported in this by the independent findings of Herbert et al. (2021), who found, that a rubric was beneficial to primary teachers implementing mathematical reasoning in their classrooms.

This research would complement the growing attention to assessing SGRs evident in the literature (Tang et al. 2019; Wilson & Bradbury, 2021; Xu et al, 2021) by providing prompts and language for teachers as they build their own confidence and facilitate formative conversations with each other and their students with the pedagogy.

However, based on this paper, we propose that a more general principle may be at play: that rubrics can be designed to support teachers to implement real pedagogical change in their classrooms by embedding key aspects of the desired pedagogy, and thereby build greater coherence between policy and classroom practice. Given similar and independent conclusions in science and mathematics disciplines, we anticipate these ideas may be more generally applicable to support teachers in other disciplines. We encourage education researchers in other disciplines to test these ideas in order to refute, confirm or refine them further. Such a principle, if confirmed, would provide both educational researchers and policymakers with a practical way to build coherence in systemic policy implementation.

Working in partnership with teachers provides a clear strategy that could be built into future teacher PL. It supports classroom change by empowering teachers to design and formatively assess learning in their own classrooms and providing educational researchers with an opportunity to draw on teachers' experiences concerned with implementing innovative pedagogies. Partnerships also enhance communication and information flows across the levels of the complex educational system and more effectively address the implementation of seemingly intractable teacher change issues often inherent in many top-down systemic educational reform initiatives. Such an outcome would, by definition, improve coherence of the educational policy being implemented.

References

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science, 333*, 1096–1097. https://doi.org/10.1126/science.1204153

Boaz, A., Davies, H., Fraser, A. & Nutley, S. (Eds.). (2019), *What works now? Evidence-informed policy and practice.* Policy Press. https://doi.org/10.51952/9781447345527

Brand, B.R., & Moore, S.J. (2011). Enhancing teachers' application of inquiry-based strategies using a constructivist sociocultural professional development model. *International Journal of Science Education*, *33*(7), 889–913. https://doi.org/10.1080/09500691003739374

Burns, T., & Schuller, T. (2007). The evidence agenda. In OECD Centre for Educational Research and Innovation, *Evidence in education: Linking research and policy* (pp. 15–32). https://www.oecd.org/education/CERI/47435459.PDF https://doi.org/10.1787/9789264033672-2-en

Cain, T. (2019). School-university links for evidence-informed practice. *Education Sciences*, *9*(2), 97. https://doi.org/10.3390/educsci9020097
Checkland, P. (2012). Four conditions for serious systems thinking and action. *Systems Research and Behavioral Science, 29*(5), 465–469. https://doi.org/10.1002/sres.2158
Cisterna, D., & Gotwals, A.W. (2018). Enactment of ongoing formative assessment: Challenges and opportunities for professional development and practice. *Journal of Science Teacher Education*, *29*(3), 200–222. https://doi.org/10.1080/1046560X.2018.1432227
Cirkony, C. (2019). *Students learning science: Representation construction in a digital environment.* [Doctoral thesis, Deakin University]. DRO. https://hdl.handle.net/10536/DRO/DU:30129406
Cirkony, C., Tytler, R., & Hubber, P. (2022). Designing and delivering representation-focused lessons in a digital learning environment. *Educational Technology Research & Development 70,* 881–908. https://doi.org/10.1007/s11423-022-10094-z
Cowie, B. & Harrison, C. (2021). The what, when & how factors: Reflections on classroom assessment in the service of inquiry. *International Journal of Science Education*, *43*(3), 449–465. https://doi.org/10.1080/09500693.2020.1824088
Daniel, K., Bucklin, C., Leone, E., & Idema, J. (2018). Towards a definition of representational competence. In K. Daniel (Ed.), *Towards a framework for representational competence in science education* (pp. 3–11). Springer. https://doi.org/10.1007/978-3-319-89945-9
DeCoito, I. (2016). STEM education in Canada: A Knowledge Synthesis. *Canadian Journal of Science, Mathematics and Technology Education, 16(*2), 114–128. https://doi.org/10.1080/14926156.2016.1166297
diSessa, A.A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293–331. https://doi.org/10.1207/s1532690xci2203_2
Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018a). Exploring relations between formative and summative assessment. In J. Dolin and R. Evans (Eds.), *Transforming assessment through an interplay between practice, research and policy* (pp. 53–82). Springer. https://link.springer.com/book/10.1007%2F978-3-319-63248-3
Dolin, J., Bruun, J., Constantinou, C.P., Dillon, J., Jorde, D. & Labudde, P. (2018b). Policy aspects: How to change practice and in what direction? In J. Dolin and R. Evans (Eds.), *Transforming assessment through an interplay between practice, research and policy* (pp. 249–278). Springer. https://doi.org/10.1007/978-3-319-63248-3_10
Duit, R., & Treagust, D.F. (2012). How can conceptual change contribute to theory and practice in science education? In B.J. Fraser, K. Tobin & C.J. McRobbie (Eds.), *Second international handbook of science education* (pp. 107–118). Springer. https://doi.org/10.1007/978-1-4020-9041-7_9
Farley-Ripple, E., Karpyn, A.E., McDonough, K., & Tilley, K. (2017). Defining how we get from research to practice: A model framework for schools. In M. Eryaman & B. Schneider (Eds.), *Evidence and public good in educational policy, research and practice. Educational governance research*, vol 6. Springer. https://doi.org/10.1007/978-3-319-58850-6_5

Faulkner, F., Kenny, J., Campbell, C., Crisan, C. (2019). Teacher learning and continuous professional development. In L. Hobbs & G. Törner (Eds.), *Examining the phenomenon of "Teaching Out-of-field"* (pp. 269–308). Springer. https://doi.org/10.1007/978-981-13-3366-8_11

Fitzgerald, M., Danaia, L., & McKinnon, D.H. (2019). Barriers inhibiting inquiry-based science teaching and potential solutions: Perceptions of positively inclined early adopters. *Research in Science Education, 49*(2), 543–566. https://doi.org/10.1007/s11165-018-9812-x

Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Goldin-Meadows, S., & Beilock, S. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science, 5*, 664–674. https://doi.org/10.1177/1745691610388764

Grangeat, M., Harrison, C., & Dolin, J. (2021). Exploring assessment in STEM inquiry learning classrooms. *International Journal of Science Education, 43*(3), 345–361. https://doi.org/10.1080/09500693.2021.1903617

Halverson, K.L. & Friedrichsen, P.M. (2013). Learning tree thinking: Developing a new framework of representational competence. In D. Treagust and C.Y. Tsui (Eds.) *Multiple representations in biology education* (pp. 185–201). Springer. https://doi.org/10.1007/978-94-007-4192-8_11

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school.* Teachers College Press.

Harlen, W. (Ed.). (2015). *Working with the big ideas of science education*. Science Education Programme of the Interacademies Press. https://www.interacademies.net/node/48225

Heredia, S.C. (2020). Exploring the role of coherence in science teachers' sensemaking of science-specific formative assessment in professional development. *Science Education*, *104*(3), 581–604. https://doi.org/10.1002/sce.21561

Herbert, S., Vale, C., White, P., & Bragg, L.A. (2021). Engagement with a formative assessment rubric: A case of mathematical reasoning. *International Journal of Educational Research, 111*, 101899. https://doi.org/10.1016/j.ijer.2021.101899

Hilton, A., & Nichols, K. (2011). Representational classroom practices that contribute to students' conceptual and representational understanding of chemical bonding. *International Journal of Science Education, 33*(16), 2215–2246. https://doi.org/10.1080/09500693.2010.543438

Honey, M., Pearson, G., & Schweingruber, H. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research*. The National Academies Press. https://doi.org/10.17226/18612

Hubber, P., Tytler, R., & Haslam, F. (2010). Teaching and learning about force with a representational focus: Pedagogy and teacher change. *Research in Science Education, 40*(1), 5–28. https://doi.org/10.1007/s11165-009-9154-9

Jones M., Hobbs L., Kenny J., Campbell C., Chittleborough, G., Gilbert, A., Herbert, S. & Redman, C. (2016). Successful university-school partnerships: An interpretive framework to inform partnership practice. *Teaching and Teacher Education, 60,* 108–120. https://doi.org/10.1016/j.tate.2016.08.006

Kania, J., Kramer, M., & Senge, P. (2018). *The water for systems change.* FSG. https://www.fsg.org/publications/water_of_systems_change

Kenny, J. & Cirkony, C. (2018a). Teaching using Student Generated Representations (SGRs) in science. In G. Woolcott & R. Whannell (Eds.), *Science teaching theory and practice: Engaging with scientific thinking, problem solving and real-world contexts* (pp. 141–67). Cambridge. https://doi.org/10.1017/9781316882535.009

Kenny, J. & Cirkony, C. (2018b). Using representations in the science classroom. In G. Woolcott & R. Whannell (Eds.), *Science teaching theory and practice: Engaging with scientific thinking, problem solving and real-world contexts* (pp. 348-74). Cambridge.

Kenny, J. D., & Cirkony, C. (2022). Using systems perspectives to develop underlying principles for educational reform. *Australian Journal of Teacher Education, 47*(1), 80–99. https://doi.org/10.14221/ajte.2022v47n1.6

Kenny, J., Hobbs, L., & Whannell, R. (2020). Designing professional development for teachers teaching out-of-field. *Professional Development in Education, 46*(3),500–515. https://doi.org/10.1080/19415257.2019.1613257

Klein, P.D. (2006) The challenges of scientific literacy: From the viewpoint of second-generation cognitive science. *International Journal of Science Education*, *(28)*2–3, 143–178. https://doi.org/10.1080/09500690500336627

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. Gilbert (Ed.), *Visualization in science education* (pp. 121–146). Springer. https://doi.org/10.1007/1-4020-3613-2_8

Knain, E., Fredlund, T., & Furberg, A. (2021). Student reasoning and representation construction in school science through the lenses of social semiotics and interaction analysis. *Research in Science Education, 51*, 93–111. https://doi.org/10.1007/s11165-020-09975-1

Lehrer, R., & Schauble, L. (2006). Developing modeling and argument in elementary grades. In T. Romberg, T. Carpenter & F. Dremock (Eds.), *Understanding mathematics and science matters* (pp. 29–53). Lawrence Erlbaum Associates.

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. Martin & R. Veel (Eds.), *Reading science* (pp. 87–113). Routledge.

Lemke, J. (2004). The literacies of science. In E.W. Saul (Ed.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 33–47). International Reading Association. https://doi.org/10.1598/0872075192.2

Levin, B. (2010). Governments and education reform: Some lessons from the last 50 years. *Journal of Education Policy*, *25*(6), 739–747. https://doi.org/10.1080/02680939.2010.523793

Luft, J.A., Dubois, S.L., Nixon, R.S., & Campbell, B.K. (2015). Supporting newly hired teachers of science: Attaining teacher professional standards. *Studies in Science Education, 51*(1), 1–48. https://doi.org/10.1080/03057267.2014.980559

Marginson, S., Tytler, R., Freeman, B., & Roberts, K. (2013). STEM: Country comparisons: International comparisons of science, technology, engineering and mathematics (STEM) education. Final report. Australian Council of Learned Academies.

McDermott, M. A., & Hand, B. (2013). The impact of embedding multiple modes of representation within writing tasks on high school students' chemistry understanding. *Instructional Science, 41*, 217–246. https://doi.org/10.1007/s11251-012-9225-6

Michalchik, V., Rosenquist, A., Kozma, R., Kreikemeier, P., & Schank, P. (2008). Representational resources for constructing shared understandings in the high school chemistry classroom. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 233–282). Springer. https://doi.org/10.1007/978-1-4020-5267-5_11

Mishra, C., Clase, K.L., Bucklin, C.J., & Daniel, K.L. (2018). Improving students' representational competence through course-based undergraduate research experience. In K. Daniel (Ed.), *Towards a framework for representational competence in science education* (pp. 177–201). Springer. https://doi.org/10.1007/978-3-319-89945-9

Nielsen, J.A., Dolin, J., & Tidemand, S. (2018). Transforming assessment research: Recommendations for future research. In J. Dolin and R. Evans (Eds.), *Transforming assessment through aniInterplay between practice, research and policy* (pp. 279–290). Springer. https://doi.org/10.1007/978-3-319-63248-3_11

Park, J., Chang, J., Tang, K., Treagust, D.F. & Won, M. (2020). Sequential patterns of students' drawing in constructing scientific explanations: Focusing on the interplay among three levels of pictorial representation. *International Journal of Science Education*, *42*(5), 677–702. https://doi.org/10.1080/09500693.2020.1724351

Penuel, W., Fishman, B.J., Gallagher, L.P., Korbak, C., & Lopez-Prado, B. (2008). The mediating role of coherence in curriculum implementation. https://repository.isls.org/handle/1/3150

Sharma, J. & Yarlagadda, P.D.V. (2018). Perspectives of 'STEM education and policies' for the development of a skilled workforce in Australia and India. *International Journal of Science Education*, *40*(16), 1999–2022. https://doi.org/10.1080/09500693.2018.1517239

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.2307/1175860

Stieff, M., Bateman, R. & Uttal, D. (2005). Teaching and learning with three-dimensional representations. In J. Gilbert (Ed.), *Visualization in science education* (pp. 93–120). Springer. https://doi.org/10.1007/1-4020-3613-2_7

Tang, K., Won, M. & Treagust, D. (2019) Analytical framework for student-generated drawings, *International Journal of Science Education*, *41*(16), 2296–2322. https://doi.org/10.1080/09500693.2019.1672906

Timms, M., Moyle, K., Weldon, P., & Mitchell, P. (2018). Challenges In STEM learning in Australian schools: Literature and policy review. ACER. https://research.acer.edu.au/policy_analysis_misc/28/

Tytler, R., Hubber, P., Prain, V., & Waldrip, B. (2013). A representation construction approach. In R. Tytler, V. Prain, P. Hubber, & B. Waldrip, B. (Eds.), *Constructing representations to learn in science* (pp. 31–50). Sense Publishers. https://doi.org/10.1007/978-94-6209-203-7_3

Tytler, R., Prain, V., & Hubber, P. (2018). Representation construction as a core disciplinary literacy. In K-S. Tang & K. Danielsson (Eds.), *Global developments in literacy research for science education* (pp. 301–317). Springer. https://doi.org/10.1007/978-3-319-69197-8_18

Waldrip, B., Prain, V., & Sellings, P. (2013). Explaining Newton's laws of motion: Using student reasoning through representations to develop conceptual understanding. *Instructional Science, 41*(1), 165–189. https://doi.org/10.1007/s11251-012-9223-8

Waldrip, B., Prain, V. & Carolan, J. (2010). Using multi-modal representations to improve learning in junior secondary science. *Research in Science Education*, *40*(1), 65–80. https://doi.org/10.1007/s11165-009-9157-6

Wilson, R.E. & Bradbury, L.U. (2021). Assessing early primary students' growth in a science unit using multiple modes of representation: investigating the promise of explicit drawing instruction. *International Journal of Science Education*, *43*(8), 1341–1364. https://doi.org/10.1080/09500693.2021.1909774

Xu, L., Prain, V., & Speldewinde, C. (2021). Challenges in designing and assessing student interdisciplinary learning of optics using a representation construction approach. *International Journal of Science Education*, *43*(6), 844–867. https://doi.org/10.1080/09500693.2021.1889070

Acknowledgements

This research associated with Cirkony (2019) was supported by the Australian Research Council as part of the Digital Pedagogies project (LP130100233).

Appendix 1. Generic assessment rubric for conceptual understanding and representational competence of student generated representations.

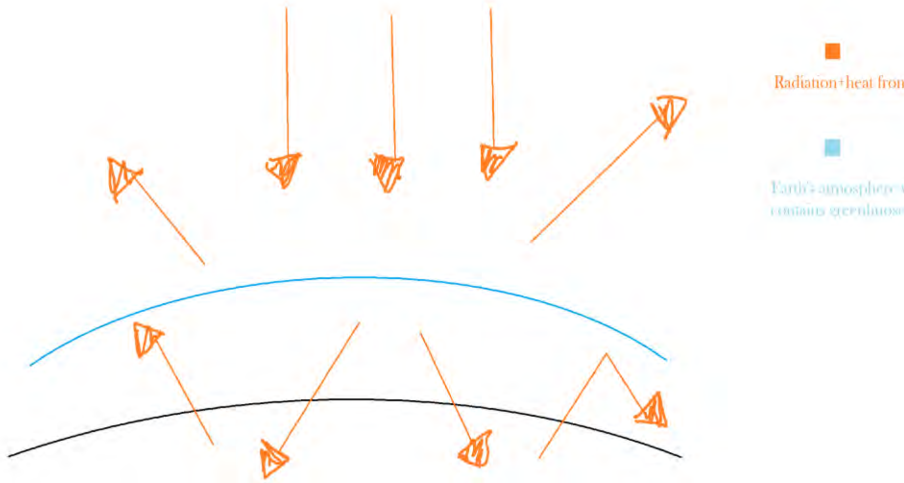
Criteria	Level 1	Level 2	Level 3	Level 4	Level 5
Conceptual Understanding	The representation shows many alternative conceptions and little or no understanding of the concept(s).	The representation shows one or more alternative conceptions and a partial understanding of the key concept(s).	The representation shows a developing scientific understanding of the key concept(s).	The representation shows a correct understanding of the key concept(s).	The representation shows a deep understanding of the key concept(s).
Nature of Representation	The representation depicts only obvious/observable features.	The representation depicts both observable and less obvious entities or processes (e.g., motion, time). It shows a rudimentary use of symbolic forms (e g., arrows, motion, time) without necessarily adding meaning to the explanation.	The representation depicts both observable and less obvious entities or processes. It shows some use of symbolic forms that add meaning to the explanation.	The representation depicts both observable and less obvious entities or processes. It shows the correct use of conventional symbolic forms to enhance meaning and explain the phenomenon.	The representation depicts both observable and less obvious entities or processes. It combines the correct conventional symbolic forms to show relationships between/among entities to clearly explain the phenomenon.
Coordination of Representational Forms/Modes	The representation shows little or no attempt to link between representational forms.	There is an attempt to link two representational forms, but the focus is only explaining observable features of the phenomenon.	There is an attempt link two or more representational forms, but the focus is on explaining observable features of the phenomenon.	There is clear link across two or more representational forms, focusing on both observable and non-obvious features to explain the phenomenon.	There is purposeful link across two or more representational forms and a description of how each explain distinct aspects of the phenomenon.
Application of Representation	The representation shows an incorrect application of key concepts to solve the problem or explain the phenomenon.	The representation shows the partial application of key concepts to solve the problem or explain the phenomenon.	The representation shows the correct application of key concepts to solve the problem or explain the phenomenon. It does not generalise ideas to similar situations.	The representation shows the correct application of key concepts to solve the problem or explain the phenomenon. It shows an attempt to generalise ideas (e.g., apply it to similar situations, make predictions or recognise patterns).	The representation shows the correct application of key concepts to solve the problem or explain the phenomenon. It generalises ideas or describes the strengths and limitations of each representation.

Appendix 2: Applied Rubric: Assessment of conceptual understanding and representational competence of student generated representations about the Greenhouse Effect.

Conceptual Understanding					
Criteria	**Limited (1)**	**Partial (2)**	**Sound (3)**	**Good (4)**	**Excellent (5)**
Light energy from the Sun is absorbed by the Earth. As the Earth warms, some of this energy is re-radiated as heat/infrared energy.	The representation shows many alternative conceptions and little or no understanding of the concept(s).	The representation shows one or more alternative conceptions and a partial understanding of the key concept(s).	The representation shows a developing scientific understanding of the key concept(s).	The representation shows a correct understanding of the key concept(s).	The representation shows a deep understanding of the key concept(s).
Certain gases in the atmosphere (e.g., carbon dioxide, nitrous oxide, methane) trap this heat energy resulting in a natural Greenhouse Effect.	The representation shows many alternative conceptions and little or no understanding of the concept(s).	The representation shows one or more alternative conceptions and a partial understanding of the key concept(s).	The representation shows a developing scientific understanding of the key concept(s).	The representation shows a correct understanding of the key concept(s).	The representation shows a deep understanding of the key concept(s).
This natural Greenhouse Effect maintains an average Earth temperature that sustains life.	The representation shows many alternative conceptions and little or no understanding of the concept(s).	The representation shows one or more alternative conceptions and a partial understanding of the key concept(s).	The representation shows a developing scientific understanding of the key concept(s).	The representation shows a correct understanding of the key concept(s).	The representation shows a deep understanding of the key concept(s).

Representational Competence					
Nature of Representation	The representation shows many alternative conceptions and little or no understanding of the concept(s).	The representation shows one or more alternative conceptions and a partial understanding of the key concept(s).	The representation shows a developing scientific understanding of the key concept(s).	The representation shows a correct understanding of the key concept(s).	The representation shows a deep understanding of the key concept(s).
Coordination of Representational Forms/Modes	The representation depicts only obvious/observable features.	The representation depicts both observable and less obvious entities or processes (e.g., motion, time). It shows a rudimentary use of symbolic forms (e g., arrows, motion, time) without necessarily adding meaning to the explanation.	The representation depicts both observable and less obvious entities or processes. It shows some use of symbolic forms that add meaning to the explanation.	The representation depicts both observable and less obvious entities or processes. It shows the correct use of conventional symbolic forms to enhance meaning and explain the phenomenon.	The representation depicts both observable and less obvious entities or processes. It combines the correct conventional symbolic forms to show relationships between/among entities to clearly explain the phenomenon.
Application of Representation	The representation shows little or no attempt to link between representational forms.	There is an attempt to link two representational forms, but the focus is only explaining observable features of the phenomenon.	There is an attempt link two or more representational forms, but the focus is on explaining observable features of the phenomenon.	There is clear link across two or more representational forms, focusing on both observable and non-obvious features to explain the phenomenon.	There is purposeful link across two or more representational forms and a description of how each explain distinct aspects of the phenomenon.

Appendix 3. Student Generated Representations: Responses to the pre-unit survey (left) and post-unit survey (right).

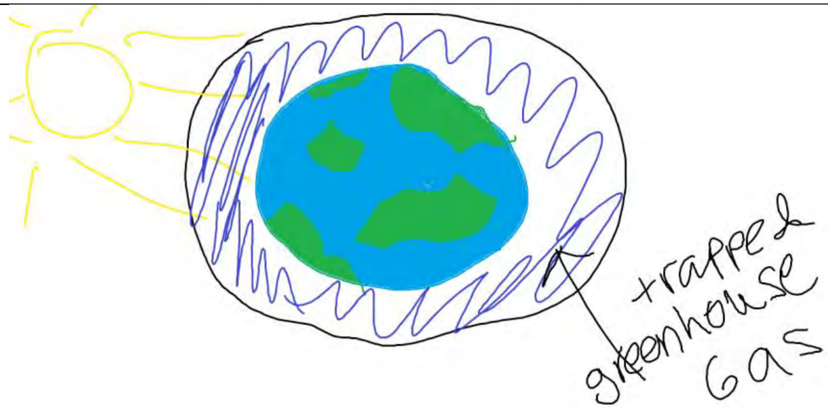
<table>
<tr><th colspan="2">Question: Scientists make reference to the Greenhouse Effect and its relation to climate change. What is the Greenhouse Effect? (Energy transfer, radiation)</th></tr>
<tr><th colspan="2">Amel</th></tr>
<tr><td>Green house effect is trapping the suns heat and energy in the atmosphere and it keeps the Earth warm, and it gets more warm gradually since they are trapped there.</td><td>

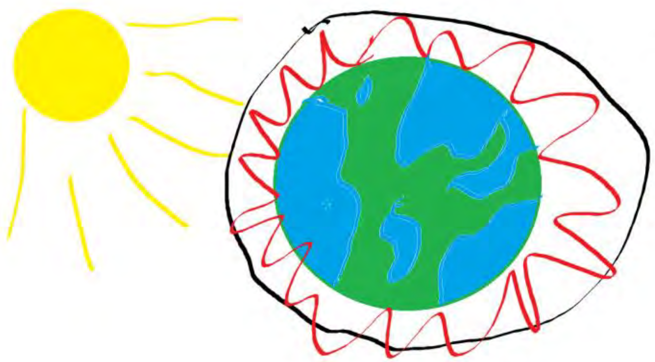
Greenhouse gasses in the Earth's atmosphere traps the radiation and heat from the sun. A part of the energy would return back into space while the other goes into earth's lower atomosphere and warms up the surface. With more greenhouse gasses, more heat is trapped within the lower atmosphere and less can return back into space, this causes the Earth to heat up more, which is what "glaobal warming" is.</td></tr>
</table>

Clara

the greenhouse effect is the process of the sun heating up he suface of the sun. When the sun's energy touches the atmosphere, the energy is returned back to space but the rest of the energy is absorbed by the greenhouse gases

Jessica



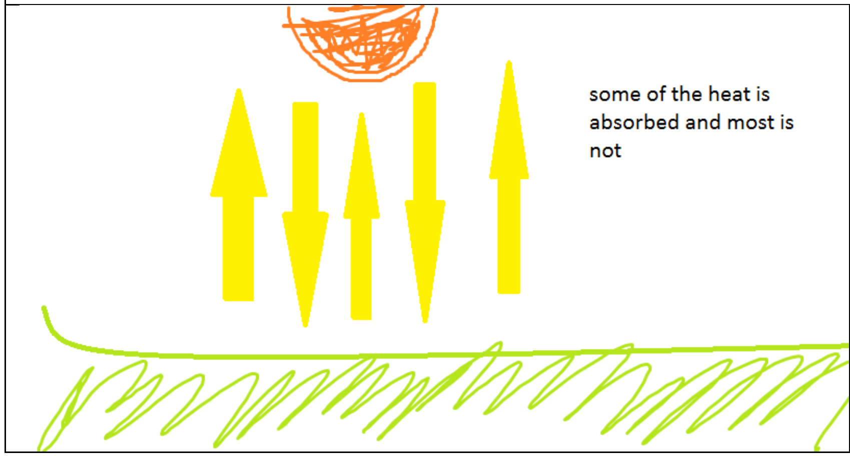


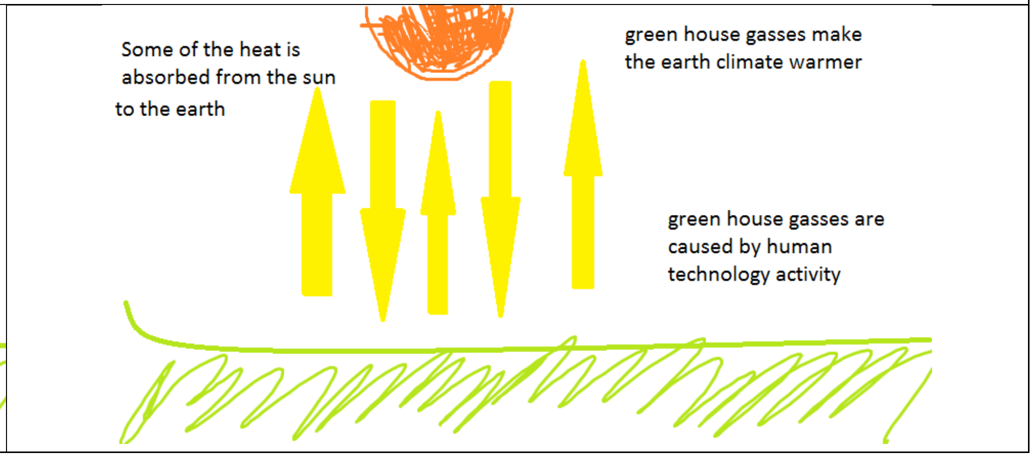
The sun directs heat towards the earth, and some of it reflects but some of it sits in the layer between the earth. This is where the red represents the cO2 and the other gases in the atmosphere which sit here and heat up the earth to an unhealthy amount

Phoebe

the green house effect is when the suns energy makes contact with the atmosphere. it warms up the earth.

Komal

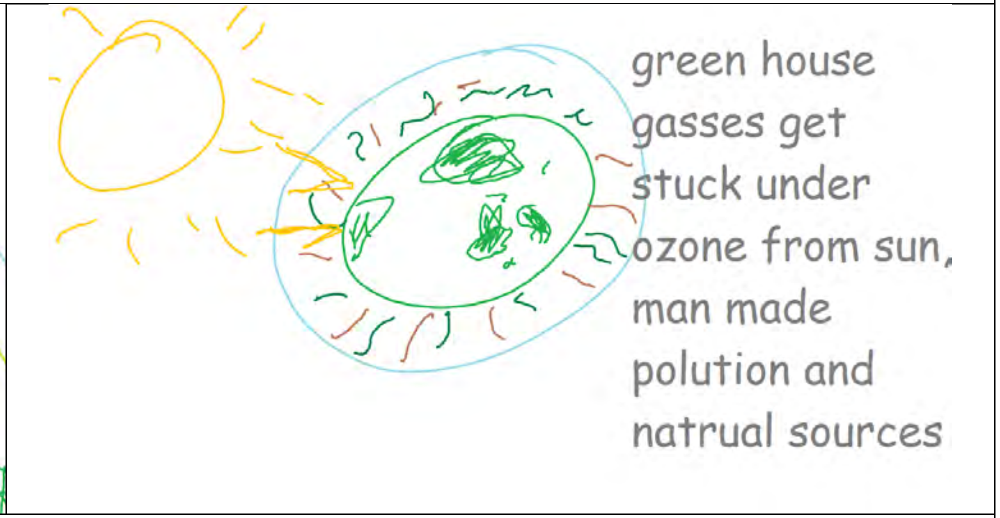




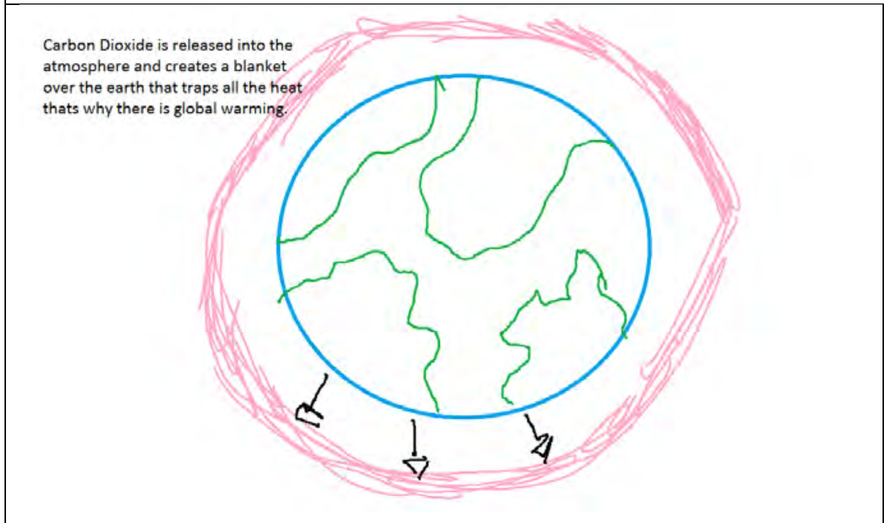
<table>
<tr><th colspan="2">McKinley</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><th colspan="2">Nyla</th></tr>
<tr><td></td><td>
</td></tr>
</table>

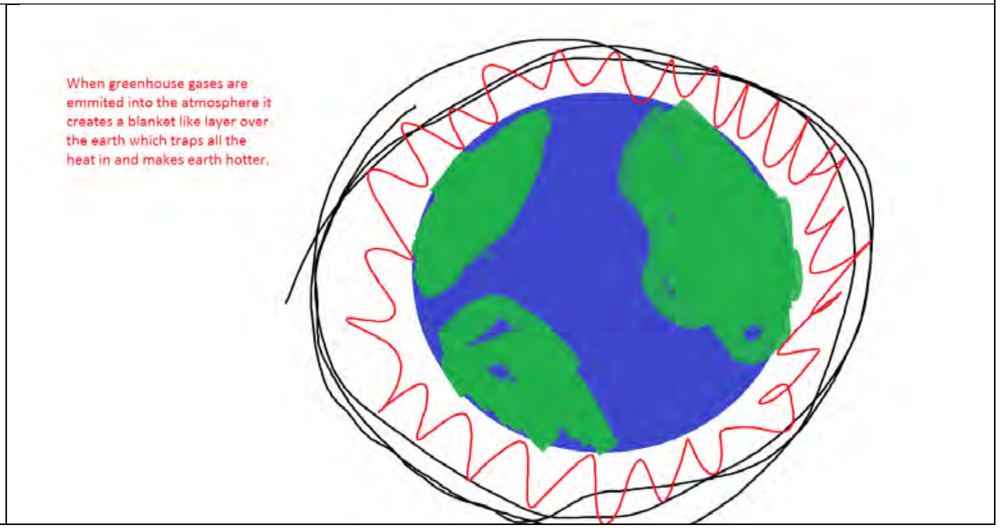
Sofi

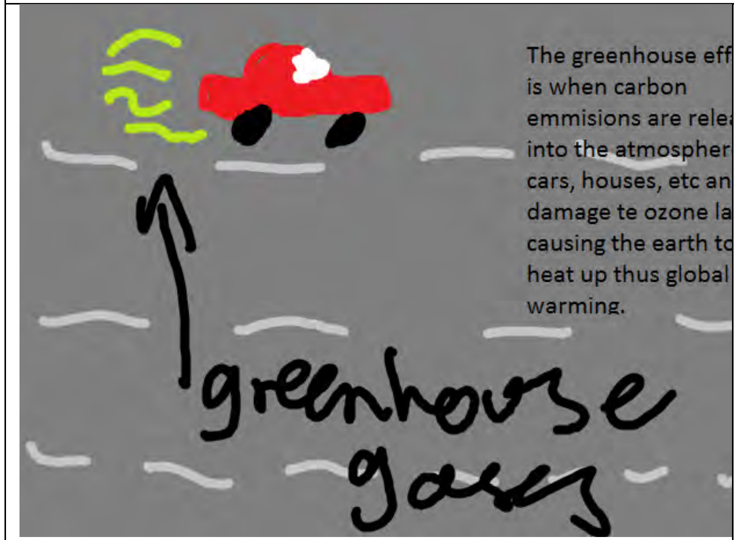

the ozone layer traps
the green house gas
caused by naturual
causes, like cows and
unatural, like factories.
this causes poloution
and melts icecaps
green house gas


green house
gasses get
stuck under
ozone from sun,
man made
polution and
natrual sources

Elham


Carbon Dioxide is released into the
atmosphere and creates a blanket
over the earth that traps all the heat
thats why there is global warming.


When greenhouse gases are
emmited into the atmosphere it
creates a blanket like layer over
the earth which traps all the
heat in and makes earth hotter.

<table>
<tr><th colspan="2">Annisa</th></tr>
<tr><td>


</td><td>The greenhouse effect the Earth's ever increasing climate due to human activity producing greenhouse gases. When human go about their daily lives using up so much energy that the radition and watse of those products go into the air, getting trapped in the Earth's atmosphere thus heating it up causing climate change.</td></tr>
</table>