UNDERGRADUATE STUDENT EXPLANATIONS ABOUT MOLECULAR PROCESSES IN INFORMATION FLOW AND TRANSFER IN BIOLOGY

Juli D. Uhl[1] [https://orcid.org/0000-0003-1569-7122] and Kevin Haudek[12] [https://orcid.org/0000-0003-1422-6038]

[1]CREATE for STEM Institute, Michigan State University, East Lansing MI 48824

[2]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing MI 48824

Abstract

National calls to improve science education include focusing on scientific practices coupled with learning disciplinary core ideas. Among the practices is constructing explanations, which include a mechanism in cellular and molecular biology, and are used by experts to make predictions about phenomena. In this work, we developed an assessment of undergraduate mechanistic explanations about the biology core concept of genetic information flow. We collected and coded 346 undergraduate student written responses. Using a mechanistic framework, we developed a rubric to capture student knowledge in their written responses. The rubric captures whether students provide a fully mechanistic, sub-molecular explanation of transcription, which is a key process in genetic information flow. We used this rubric to compare five item versions and found that one version elicited fully mechanistic explanations from 20% of students, compared to between 2 and 13% from other versions. This version included the element of time, by indicating that a new RNA was formed as part of transcription. We also found that students integrated a median of two related ideas with their explanation of transcription. Our work demonstrates that careful wording is essential for item writing and that students can explain molecular processes like transcription by leveraging sub-molecular mechanisms.

Introduction

National calls to improve science education at the K – 16 levels include a focus on scientific practices linked to disciplinary core ideas (e.g., American Association for the Advancement of Science, 2011; National Research Council, 2012). In the disciplines of molecular and cellular biology, the scientific practice of explaining core ideas of genetic information flow should include a mechanism using entities at lower scalar levels to describe the phenomenon (van Mil et al., 2013). Assessing the complex ideas and connections in mechanistic explanations requires prompts that allow students to respond in their own words, such as constructed response (CR) items. In this paper, we describe the development of a CR item and an associated rubric capable of eliciting and capturing student mechanistic explanations of the process of transcription, a process within the core concept of genetic information flow.

Background

For undergraduate biology, the American Association for the Advancement of Science document, Vision and Change sets a goal for students to be capable of interpreting and communicating about science within the disciplinary practices, as well as the need for students to learn the core biological concepts (American Association for the Advancement of Science, 2011). The three-dimensional learning framework (National Research Council (NRC), 2012) was developed to guide science education at the K-12 level and is applicable to college-level introductory course and assessment design (e.g., Allred et al., n.d.; Cooper et al., 2015; Matz et al., 2018). The NRC framework includes three linked learning dimensions: scientific and engineering practices, disciplinary core ideas, and cross-cutting concepts. The NRC framework includes among its practices, "Constructing Explanations and Designing Solutions," and describes scientific explanations as supporting cause and effect relationships (NGSS Lead States, 2013). Vision and Change implies inclusion of explanations within the practice of applying the scientific process (Clemmons et al., 2020). Based on these frameworks, students must learn to make explanations about core concepts. Further, explanations which draw upon scientific concepts provide students a way to apply their conceptual knowledge and thus represent a potential method for conceptual change (NRC, 2012).

Scientific explanations can draw from multiple philosophies (Braaten & Windschitl, 2011) and experts often incorporate analogies, methods and context as part of their explanation of a molecular phenomenon (Trujillo et al., 2015). However, it is commonly accepted that an explanation of a scientific phenomenon should be a causal account of how and why the phenomenon occurs (NRC, 2012; Sandoval & Millwood, 2005). Furthermore, scientific explanations should apply scientific models or theory to explain the cause of the phenomenon, be consistent with available evidence, and student understanding of science (NRC, 2012 Chapter 3). When experts make explanations, they often use them for the purpose of making predictions and generating new hypotheses, but for students an educational goal might be to make an explanation that accounts for the cause(s) of the phenomenon (NRC, 2012, Chapter 3). Explanations should be more than a definition or description and should include a causal chain linking reasoning that fits the available evidence to a claim (Reiser et al., 2012).

In molecular and cellular biology, scientific explanations include mechanisms as a central feature (Machamer et al., 2000; van Mil et al., 2013). In a mechanistic explanation, one pays attention to entities and their spatial and temporal organization as well as multiple scalar levels of organization. For example, entities interact by binding, which leads to a change in state and a change in activity (van Mil et al., 2013). Furthermore, explanations should transition across scalar levels, from the molecular to cellular level for example (van Mil et al., 2013). In particular, these types of explanations should go down a scalar level from the phenomenon, or to an abstract level (such as energy) and leverage characteristics and activities of entities to describe the phenomenon (Krist et al., 2018).

Work on molecular genetic mechanisms indicates that in addition to physical levels of entities, the level of information must be considered. For example, a DNA sequence has an information level in that it contains genetic information, which must be leveraged alongside a physical level such as the protein for which the DNA codes (Duncan & Reiser, 2007). Thus, in addition to considering a lower scalar level (e.g., to the molecular or sub-molecular level) in their explanations of genetic information flow, students also need to link the information level to the output of the transcriptional process (most commonly protein). Work on genetic mechanisms suggests that students can identify molecular mechanisms involved in biological functions, but may have challenges linking across scalar levels (e.g. from molecular to cellular; see Marbach-Ad & Stavy, 2000; K. Southard et al., 2016; K. M. Southard et al., 2017).

Domain-specific knowledge including explanatory schemas affects student ability to reason about genetic mechanisms (Duncan, 2007). Thus, genetic mechanisms and concepts - when taught together - have the potential to improve learning. Learning about genetic information flow is challenging for both secondary and undergraduate students (see for example, (Briggs et al., 2017; Marbach-Ad & Stavy, 2000) and some misconceptions persist after instruction (Smith & Knight, 2012). Secondary students are unlikely to link DNA to protein by way of mRNA (Gericke & Wahlberg, 2013). At the introductory college biology level, students have misconceptions about the processes of replication, transcription, and translation and the flow of information through the process (Prevost et al., 2016; Smith et al., 2008). Furthermore, students often include additional concepts in their descriptions of each process (Uhl et al., 2020), and transcription can be especially difficult for students to learn and may be considered a threshold concept (Moscarella et al., 2016). Students can have several misconceptions about transcription, including that the DNA undergoes a chemical conversion to become RNA or that the RNA exists before transcription (Wright et al., 2014).

Even when students attain correct understandings of genetics, retaining that learning can be limited, which poses challenges for students when they need to apply introductory genetics concepts to upper-level courses. After completion of introductory biology course, students retain information about the subunits of DNA and protein, and less so on mechanisms (Briggs et al., 2016), and memorizing patterns instead of learning mechanisms can hinder learning (Castro-Faix et al., 2021). Furthermore, it has been proposed that learning genetic mechanisms leads to longer retention of knowledge than rote memorization (Todd & Romine, 2018) and that non-mechanistic conceptions of phenomena can be a barrier to learning (Haskel-Ittah & Yarden, 2018). Thus, some learning challenges may be mitigated by teaching students the skills necessary to mechanistically reason about genetics.

Study Context

We wanted to know if students in introductory biology classes were memorizing the steps of transcription or connecting transcription process to the core concept of information flow from Vision and Change (AAAS; 2011). To analyze student understanding of transcription, we asked students to respond to the item, “When an organism needs a particular gene product (protein), gene expression is activated, resulting in the production of RNA. How is the information contained in a gene transferred from DNA to RNA during transcription?” We found that many students described the process of RNA polymerase synthesizing a new RNA, and/or described that complementary base-pairing occurred between the RNA and DNA molecule. However, few (7%) students link the idea of complementary base-pairing with the idea that the information transferred is in the form of a sequence transferred from the DNA to RNA during transcription (unpublished results).

However, in subsequent focus group interviews related to this item we found that when pressed, students were able to state that the nature of the genetic information is the sequence of DNA and RNA, and that the sequence is transferred during transcription (unpublished results). We concluded that written responses to the original item to be potentially unaligned with student ability to explain and link ideas about transcription. Based on this, we sought to

develop a new item to elicit student mechanistic understanding of transcription and an aligned coding rubric that could identify mechanistic explanations.

Research Questions

We reviewed the item described above and noted that there was nothing that indicated to students that they need to describe how and why the information is transferred, or to make any links between the information transferred and the entities, characteristics, and activities involved in transcription. Thus, we sought to develop a new assessment item and rubric that could elicit and evaluate student explanations about transcription. Our goal was to develop an item that would elicit such explanations as well as a rubric to evaluate the explanations. We asked the following research questions.

1. How do undergraduates explain the phenomena of information transfer during transcription?
2. What are the characteristics of an item that are more likely to elicit a mechanistic explanation about gene expression?
3. What additional ideas do undergraduates use in their explanations of transcription?

Methods

Item Development

We based our assessment design on the NRC assessment triangle (Pellegrino et al., 2001), where cognition is represented by mechanistic links across scalar levels to describe how and why a phenomenon occurs (Machamer et al., 2000). We observed these explanations using student CRs, which closely resemble student verbal responses during interviews (Nehm & Schonfeld, 2008; Weston et al., 2015). Student responses were interpreted with a coding rubric capturing the target phenomenon, identifying activities of entities at scalar level below the phenomenon, and connecting to target phenomenon in the context of genetic information flow during transcription (Krist et al., 2018; Russ et al., 2008).

All items included the context of information as transferred from DNA to RNA during transcription plus one or more prompts for students to describe the process used in information transfer. Items were developed with input from a group of biochemistry, molecular biology, and microbiology experts with the intent that they be useful as formative assessment for biology instructors who wish to understand and respond to students' thinking (Pellegrino et al., 2016). We used written scaffolding in our items (McNeill et al., 2006), which involved using keywords in prompts such as *describe, explain,* and *why* to aid students in writing explanations. We also scaffolded for content to encourage respondents to think about the levels below the phenomena (molecular and sub-molecular), and to include the element of time as a comparison point (as recommended by Krist et al., 2018). Items one, two, and three ask students to use the characteristics of DNA and RNA in their explanations, and items four and five were split into three prompts to instruct students to think about the molecular and sub-molecular level of the process (Table 1). The two multi-part prompts were intended to help students separately focus on characteristics of entities at multiple biological levels.

Table 1. Item versions with characteristics used to encourage students to explain. Keywords underlined in item text.

Version	Item	Biological Level	Comparison
1	During transcription, information contained within a gene is transferred from DNA to RNA. Describe the process used to transfer this information. Be sure to identify and explain the characteristics of DNA and RNA that allow it to hold and transfer information.	Macromolecular (DNA, RNA)	
2	During transcription, information contained within a gene is transferred from DNA to RNA. Identify and explain the characteristics of DNA and RNA that allow it to hold and transfer information. Then, using these characteristics, describe the process used to transfer this information.	Macromolecular (DNA, RNA)	
3	Explain the molecular events that transfer information from DNA to RNA during transcription. Be sure to discuss why the characteristics of DNA and RNA ensure this information is transcribed correctly.	Macromolecular (DNA, RNA)	
4	During transcription, the information from DNA is transferred to a new strand of RNA. a. Describe the process used to transfer this information. b. Identify and explain the characteristics of DNA and RNA that allow it to hold information. c. Identify and explain the characteristics of nucleotides that allow transfer of information.	Macromolecular (DNA, RNA) Molecular (nucleotide)	
5	During transcription, a new strand of RNA is formed, which holds the information from the DNA. a. Describe the process used to transfer this information. b. Identify and explain the characteristics of DNA and RNA that allow it to hold information. c. Identify and explain the characteristics of nucleotides that allow transfer of information.	Macromolecular (DNA, RNA) Molecular (nucleotide)	Time (a new RNA is formed)

Data Collection

We collected 346 undergraduate responses from introductory biology and upper division biochemistry courses at a large research-intensive university in Fall 2019 and Spring 2020. Students randomly received an item version, and responses were collected electronically through a course learning management system as post-instructional homework. In Fall 2019, we piloted and collected 45 responses for version one, 41 for version two, and 48 for version three from an introductory biology course. In Spring 2020, based on responses to versions 1, 2, and 3, we developed two new prompts and piloted versions four and five and collected 74 responses for version 4 and 68 for version 5 from introductory biology courses and 38 responses for version 4 and 32 for version five from an upper-level biochemistry course. Spring 2020 responses were collected after Covid-19 forced a shift to online learning; however, patterns appear to hold in ongoing research. Responses were de-identified and those with multiple text box entries were concatenated into a single block of text for coding. Typical student responses were a paragraph in length. Coders were blind to the item version and course level to avoid biasing the results. Responses presented as examples in this paper were spell-checked for clarity.

Theoretical Framework

We drew upon the mechanistic framework proposed by Krist et al., to develop a rubric to capture student explanations about genetic information flow during transcription (Krist et al., 2018). In the framework, which defines heuristics for a mechanistic explanation, the phenomenon exists at one scalar level, and the mechanism at a second (lower) scalar level. Their framework consists of four codes (italicized) for mechanistic reasoning: *multiple levels* which refers to considering levels above or below the phenomenon, *identifying* and *unpacking factors* at a lower scalar level, and *linking* lower-level interactions and behaviors to the target phenomenon. In *multiple levels*, Krist et al., (2018) indicate that explanations must include a cause at least one scalar level below the phenomenon or consider abstract factors. They also suggest that non-material factors such as energy or forces like gravity, represent a move from the abstract to the concrete and thus can also be categorized as moving a level below. To help identify abstract factors that relate to the phenomenon of sequence transfer during transcription, we drew upon the ideas presented by Duncan and Reiser (2007), who state that students need to move across ontological levels when reasoning about genetics. They described genetic phenomena such as genes affecting traits as a hybrid hierarchical phenomenon that includes an information level (e.g., genes code for proteins) in addition to a physical level (e.g., the molecular processes of replication).

In *identifying* and *unpacking factors*, Krist et al. indicate that students must consider the entities present at the lower scalar level or consider abstract level factors. To unpack those factors, explanations also should consider the behaviors and interactions causal to the phenomenon. Using factors at the molecular level is a part of a mechanistic explanation in molecular biology according to van Mil et al., who state that molecules and their interactions lead to molecular activities (van Mil et al., 2013). In *linking,* Krist et al., 2019 indicate that student explanations need to 'check' that the factors they unpacked are relevant to the target phenomenon in that they can be used to explain the cause of the phenomenon.

Mechanistic Rubric

Informed by Jescovitch et al.'s 2019 work on rubrics, which demonstrated that deconstructing a holistic rubric into analytic components can improve both reliability and validity (Jescovitch et al., 2019), we chose to develop an analytic coding scheme to code the responses and then align combined analytic codes to holistic categories. To ensure our analytic codes (underlined) captured pertinent ideas for mechanistic explanations, each code was associated with the elements of Krist et al. 2018 framework (italicized).

The phenomenon is captured by the analytic code 1. <u>Transfer of Base Sequence</u> from DNA to RNA. The phenomenon is partially included in all five items: though the wording differs slightly, all five items in some manner indicate that information from DNA is transferred to RNA. We recognize that this raises a potential issue, in that students might not recognize a need for an explanation of sequence transfer, and thus will not include any description of the sequence or its transfer in their explanations. Despite this potential limitation, we considered it important for students to recognize that information transfer must include the base sequence; and thus, required that students either state or imply that the sequence is transferred for coding in this category. We chose to capture Krist et al.'s *Multiple Levels* code implicitly in our coding of identification and unpacking of entities a scalar level below. Because the phenomenon is at the macromolecular level (nucleic acids are polymers), the scalar levels below include entities and interactions at the molecular level (nucleotides base-pairing) and the sub-molecular level (non-covalent interactions between bases). We also argue that RNA polymerase activity is at the molecular level on the basis that it catalyzes bond formation between the molecular entity of a nucleotide with the RNA strand. Thus, while RNA polymerase itself is a macromolecule at the same scalar level as DNA and RNA, we consider its activity for identifying a scalar level.

We used a set of analytic codes to capture Krist et al.'s *Identifying and Unpacking* with the following categories. While one might argue that RNA polymerase is a factor at the same scalar level as the nucleic acids, we considered the RNA polymerase behavior in synthesizing a new RNA to be similar to Krist's description of structure-behavior-function and thus refers to a factor that can be unpacked in explanations about information flow during transcription. We captured the unpacked factor RNA polymerase with the analytic code 2. RNA Polymerase Activity (RNA Polymerase synthesizes a new RNA). We capture responses that describe the nature of RNA and DNA nucleotide bases as complementary with the category 3. Complementary Base-Pairing. We found that students often used this idea in two distinct ways, and thus further divided this code into: 3.1-Present and Unlinked for responses that describe RNA and DNA nucleotides as complementary without making connections to sequence transfer or 3.2-Linked to Sequence Transfer for responses that describe that the newly built RNA is complementary to DNA and contains the sequence information from DNA. We capture responses that describe the nature of the base complementarity during transcription with the category 4-Non-Covalent Interactions (e.g., hydrogen bonds or molecular shape, see Supplemental Table 1). Similar to our base pairing code, we further separated this code by whether students linked interactions to other ideas. 4.1-Present and Unlinked for responses that used non-covalent interactions to describe something other than the reason for base-pairing between RNA and DNA including statements that were deemed too vague to accurately categorize, or as 4.2-Linked to Complementarity when the student response used non-covalent interactions as the cause of base pairing specificity between RNA and DNA during transcription, or directly as the cause of sequence transfer. Detailed descriptions and example responses of the analytic codes can be found in Supplemental Table 1.

The final code from Krist et al. 2018 is *Connect to target phenomenon.* This code requires causally linking one or more unpacked factors to the target phenomenon. Our analytic codes do not capture linkages of multiple unpacked factors to the target phenomenon, thus we used combinations of analytic codes to categorize student responses into holistic categories representing the type of explanation (Table 2). Responses that linked the ideas of 1.Transfer of Base Sequence with 2. RNA Polymerase Activity were categorized as Descriptions. We captured two types of mechanistic explanations. First, responses that include the concept of 3. Complementary Base-Pairing linked to 1. Transfer of Base Sequence and an accurate account of 2. RNA Polymerase Activity were considered a Molecular explanation. Second, we considered Molecular explanations that also included 4. Non-covalent Interactions linked to 3. Complementary Base-pairing as *Sub-molecular* explanations. Sub-molecular explanations also include responses in which students directly link 4. Non-covalent Interactions to 1. Transfer of Base Sequence plus an accurate account of 2. RNA Polymerase Activity. Responses that contained two or three analytic mechanism codes but lacked a link to 1. Transfer of Base Sequence and/or to 2. RNA Polymerase Activity were categorized as Unlinked, and responses that contained one or fewer analytic mechanism codes were categorized as Other.

Table 2. Student explanation types and examples. Holistic categories cannot co-occur. All student responses are spell-checked.

Holistic category	Analytic codes	Description	Example responses
Other	One or fewer analytic codes	No links between any analytic codes. Often describe other processes or includes errors	The RNA is transcribed from the DNA by ribosomes that read off the information necessary to create the new strand of RNA. The fact that DNA is a double stranded helix allows it to hold the bases within its structure to hold DNA. Nucleotides bind to the ribosomes and allow the ribosomes to copy the information. (Introductory student 2624, version 5)
Unlinked	1 or 2 plus 3 and/or 4	Any pair or trio of analytic codes, lacking one or both of RNA polymerase and sequence transfer ideas	The DNA strand is unwound by RNA polymerase at a promoter. Then, the strand of DNA that is being used as the template is read by RNA polymerase (from 3' to 5'), which then begins to build an mRNA (5' to 3') with complementary bases (G to C, A to U (Uracil is used in RNA as opposed to Thymine in DNA)) (the bases are paired correctly through IMFs). When the RNA polymerase reaches a terminator, it is finished making the mRNA. (introductory student 4397, version 1)
Description	1 & 2	RNA polymerase synthesizes an RNA, which results in transfers of the DNA sequence to the RNA	The DNA strand is unwound and used as a template for RNA synthesis. RNA polymerase copies the DNA strand into the RNA strand. DNA is the template that holds information for coding an RNA sequence. The RNA sequence holds the information for coding proteins in terms of codons. The RNA sequence is translated into proteins via three base codons that align with an amino acid. Nucleotides are the foundation for DNA structure that holds information. The order of the nucleotides corresponds with the order of the DNA, RNA and protein sequence. (upper-level student 4254, version 4)
Molecular Mechanism	1, 2, & 3.2	All elements of a description plus complementary base-pairing between DNA and RNA	Some characteristics of DNA that allow it to pass genetic information is base pairing. Base pairs in DNA include A to T, and C to G. These complimentary base pairs are also how the template strand of DNA is read to make RNA. Except for RNA it is not A to T, it is A to U. In transcription the RNA Polymerase starts transcribing at the promoter, then nucleotides are added with their complimentary bases to the growing RNA transcript, and then transcription ends when the RNA Polymerase comes across the terminator. At this point the RNA transcript is released and ready for translation. That is the process of how DNA's information is copied to RNA and to eventually create proteins. (introductory student 3149, version 2)
Sub-molecular mechanism	1, 2, 3.2, & 4.2	All elements of a molecular mechanism plus non-covalent interactions result in complementary base-pairing	During transcription, RNA polymerase reads a strand of DNA from the 3' end to the 5' and links nucleotides that are complementary to the ones in the template strand from the 5' end to the 3' end. The nucleotides of DNA and RNA are arranged in a certain order according to one's genetics, allowing them to act as a blueprint for translating into sequences of amino acids and eventually proteins. The nucleotides are complementary to

	or	or	each other. A is paired with T in DNA or U in RNA. C is paired with G in both. The nucleotides are purines and pyrimidines linked together by hydrogen bonds which can be broken for DNA replication or transcription. (introductory student 2172, version 5)
	1, 2, & 4.2	Direct link between non-covalent interactions and sequence transfer	DNA's nucleotides attach to the nucleotides in RNA. Each type of nucleotide is able to attach to a specific nucleotide in RNA based on chemical interactions. DNA and RNA are made of nucleotides that hold an information sequence that tells the RNA which protein to make. Nucleotides are monomers that make up nucleic acid. Two types of nucleic acid are DNA and RNA. These nucleotides and interact and attach through chemical interactions that allow the RNA to read the information in the DNA (introductory student 4412, version 4)

Jescovitch and colleagues indicated that analytic codes when recombined could lose capacity to accurately categorize responses along a holistic scale (Jescovitch et al., 2019). Thus, we performed a holistic coding check. Two coders, experts in biology and chemistry, who were not familiar with the analytic coding rubric were introduced to the mechanistic framework. The coders independently coded 20 responses holistically as Sub-molecular, Molecular, Description, or no code applied then met with one of the analytic coders to discuss and to resolve discrepancies. Following this process, the holistic codes were compared to analytically determined explanation categories and this comparison was used to further refine the analytic rubric such that the codes from the analytic categories accurately reflected the holistic coding expectations. All responses were reviewed by the coders to ensure final codes were consistent with the final rubric rules. Thus, the final codes presented here have all been agreed upon by two coders, subject to a holistic code check, and a consistency check. For details on analytic rubric categories and final coding rules, see Supplemental Table 1.

Conceptual Rubric

To capture the other ideas and concepts students used to contextualize their responses, we used qualitative analysis software (QDA Miner Version 5.0.31 Copyright 2004-2016) combined with an emergent coding technique similar to that used by Sripathi et al., in 2019. We identified three themes in student responses representing processes involved in gene expression. These are captured in the codes Translation, in which students describe that protein is produced through translation of RNA; Gene Regulation, in which students describe the location of transcriptional initiation as the promoter and/or describe the role of transcription factors in controlling transcription; and RNA Processing, in which students describe one or more of the mRNA post-processing steps such as poly adenylation or splicing. We also identified three themes that did not represent an additional process in gene expression. These were captured in the codes Phase-Naming, in which students name two or more of the phases of transcription (Initiation, Elongation, and Termination); Cellular Structure, in which students describe DNA and/or RNA in relation to their location in the cell; and General Nucleic Acid Characteristics, in which students describe one or more structural or chemical characteristics of DNA, RNA, or nucleotides. A final category, Nature of Genetic Information captures when students state or imply that the sequence of RNA or DNA hold information or that RNA or DNA code for proteins. This category captures a key aspect of student understanding of genetic information flow, that students understand what is meant by the term genetic information, and how nucleic acids can store information. For detailed descriptions of the conceptual rubric, including coding rules and example responses see Supplemental Table 2.

Coding and Inter-Rater Reliability

To code the student responses, two individuals with PhDs in Biology independently coded 20 responses using the analytic rubric, met to discuss discrepancies and to revise and refine the rubric coding rules. This process was repeated three more times with new responses (for a total of 80 responses) until the inter-rater reliability (IRR) of each analytic category measured by Cohen's Kappa exceeded 0.672 (substantial agreement; Supplemental Table 3; Cohen, 1960; Landis & Koch, 1977) and coding rubric rules were considered complete. Following this, the two coders independently coded 40 responses and met to resolve discrepancies. This process was repeated until all 346 responses were consensus coded. It is documented that students often confuse the names of the entities and processes involved in gene expression (Fisher, 1985; Southard et al., 2016; Zukswert et al., 2019). Thus, in cases of uncertainty about the named entities or processes, coders used the entirety of the response when coding to assign the code. That is, when a student used unconventional language to describe a process as well as more typical language, coders used the more typical language in assigning codes. In cases where consensus could not be reached, those responses were brought to a third coder with a PhD in Biology and discussed among all three coders until consensus was reached.

Data and Statistical Analysis

We performed Kruskal-Wallis test to compare numbers of context ideas based on explanation type, and Mann-Whitney tests for post-hoc pairwise analysis. Effect sizes for the Mann-Whitney test were estimated according to (Fritz et al., 2012). For Mann-Whitney tests, p values are reported in article main text and other statistics (i.e., test statistic and effect sizes) are reported in full in the Supplemental Materials. Pearson's Correlation coefficient was calculated to compare the degree to which individual ideas co-occurred with each explanation type. All statistical tests were performed in SPSS, Version 24.

Results

Research Question 1. How do undergraduates explain the phenomena of information transfer during transcription?

We found that over half of undergraduates who responded to the prompts could explain information transfer during transcription by linking ideas related to the role of RNA polymerase in synthesizing an RNA based on the DNA sequence. Without linking other ideas like complementary base-pairing or non-covalent base interactions, to the role of RNA polymerase responses were categorized as Descriptive, which was represented in 12% of student responses. Some students built on this by linking to the Molecular mechanism of complementary base-pairing (25%) and 13% additionally linked the Sub-molecular mechanism of non-covalent bonds. Responses that described either RNA polymerase activity or DNA to RNA sequence transfer, but not both, were categorized as Unlinked, representing 21% of responses. Some Unlinked responses also included ideas about base-pairing and/or non-covalent interactions. The remaining responses (27%) were incomplete, incorrect, or focused on the wrong process, and were categorized as Other. Below, we discuss Descriptions, Mechanistic, and Unlinked explanations.

Descriptive responses

We begin by reviewing descriptions, which we did not consider mechanistic as they did not leverage the properties or activities of the nucleotides bases to describe how the sequence was transferred. These students described that the sequence was transferred by the RNA polymerase synthesizing a new RNA. For example, an introductory student in response to version 3 stated, (underlining added by authors to emphasize coded phrases),

> Transcription is when a DNA sequence is copied via RNA polymerase to create an RNA. In other words, it is genetic information being transferred starting from DNA to RNA. The process functions as DNA>RNA. The first step involved DNA being elongated and is transcribe into RNA. The old segment is transcribed from 3'-5' leading for the complementary to create in the 5'-3' direction. This results in the non-template strand being formed having the same sequence as the synthesized RNA. The end product of translation results in the making of mRNA and later translation.

In the first sentence, the student clearly states that RNA polymerase creates an RNA and that the sequence is copied from DNA to RNA during transcription. While the student later states that translation forms the mRNA, the coders determined that this student likely was still referring to transcription based on the context within the rest of the response. Other students implied that the sequence was transferred from DNA to RNA by stating that the sequence of DNA held information and that information was transferred during transcription. For example, this introductory biology

student indicated that the order of bases holds information (the genetic code, further defined as genetic information), and that the RNA polymerase uses the DNA template to form an RNA in their response to version 2,

> During transcription, the information stored in a gene's DNA is transferred to RNA. DNA is made of molecules called nucleotides. The order of the bases is the thing that determines the genetic code. DNA is better at holding genetic information because it is a stable storage form. RNA contains a ribose sugar that makes it more reactive. Transcription uses DNA as a template to make an RNA molecule. An RNA polymerase enzyme strand comes in to separate the template and its nontemplate strand. then forms an mRNA that is later used in translation to a protein.

These responses demonstrate these students understand that the DNA has a sequence which is transferred to RNA during transcription. However, they either believe complementary base-pairing is implied in the text of their response, do not think it is important to include, or do not recognize or understand the importance of complementary base-pairing in transcription.

Mechanistic responses

We categorized responses that included the concept of complementary base-paring in addition to sequence transfer and RNA polymerase ideas in student descriptive responses as Molecular Mechanism. For example, an upper-level student wrote in response to version 5,

> During transcription, RNA polymerase is used to copy the information from the DNA sequence and create an appropriate mRNA template. The RNA polymerase goes from the 5' end to the 3' end and builds RNA in an antiparallel fashion. DNA and RNA hold information in their nitrogen bases, Adenine, Thymine, Guanine, and Cytosine. Each of these bases has a matching pair that they bond with, which creates the genetic code. The RNA polymerase is able to create the RNA by copying the genetic code and the base pairs match with each other

This student describes that RNA polymerase copies the DNA sequence to the mRNA. They go on to describe complementary base-pairing by stating that each base has a matching pair they bond with, and that this is the reason the code is copied. Other students list the base-pairs (A with T/U and C with G) and/or give a more detailed account of base-pairing during RNA synthesis. For example, an introductory student wrote in response to item version 1,

> During the process of transcription, the DNA double helix is unwound, separating the DNA into a template strand and a noncoding strand for a short series of nucleotides in a sort of transcription bubble. This process starts at a promoter region, where proteins, including RNA polymerase bind. The RNA polymerase takes nucleoside triphosphates (NTPs) and matches them with the complementary nucleotides in DNA. In RNA, A matches with T, G with C, C with G and U with A. As each new nucleotide is added, a phosphodiester bond will link the nucleotides. The base pairing continues along the DNA molecule until the terminator region is reached. At this point, transcription is over, and the new RNA molecule is released from the RNA polymerase complex and the DNA molecule is reconnected back to its double helix shape. The sequence of nucleotides in the RNA molecule will go on to be the code for protein during the process of translation.

This student also wrote a more detailed explanation of the process of RNA synthesis by RNA polymerase, stating that RNA nucleoside triphosphates are matched with complementary nucleotides in DNA, and that phosphodiester

bonds form between the nucleotides. Each of these ideas builds toward the idea that a sequence in DNA is transferred to the RNA, and the student makes that explicit at the end by stating that the sequence of nucleotides in the RNA will go on for translation. While the level of detail a student uses to describe RNA synthesis varies, we did not view a highly detailed account of RNA polymerization as necessary to adequately respond to these items. However, students who provided a highly detailed account of RNA polymerization also tended to be capable of writing a mechanistic explanation.

Responses including a molecular mechanistic explanation that also went to the sub-molecular level by describing hydrogen bonding, purine and pyrimidine interactions, or other non-covalent interactions as the reason that complementary base pairing occurred were categorized as a Sub-molecular Mechanism. For example, this introductory student wrote in response to item version 4,

> An RNA polymerase enzyme attaches to a DNA strand that is targeted for its code for a cell's desired protein. The enzyme "unzips" the DNA double helix, matches up complementary nucleotides to form an mRNA strand, then re-attaches the two strands to DNA. DNA and RNA are able to hold information because their nucleic bases form unique codes for the proper order of amino acids added to a peptide chain during protein synthesis. The information *is* the order of the nucleotides. Each nucleotide forms a hydrogen-bonded base with exactly one other nucleic base... or two bases in the case of adenine. The specific combination requirements (mostly) guarantee that the correct information is passed along the transcription and translation process chains to create properly constructed, structured, and functioning proteins.

This student begins by stating that RNA polymerase forms an mRNA strand and is clear that complementary nucleotides are matched during formation. The student then gives evidence that they understand the concept of sequence transfer from DNA to RNA by defining the information in two ways; that bases form codes for amino acid order as well as stating that the order of the nucleotides is the information. They then complete their description of why base-pairing occurs by explaining that hydrogen bonding is specific to base pairs, and that is the reason for information transfer during transcription. This student response may seem to be a bit out of order, but that is likely a result of the format of this question version, which asked in the final part of the item how the characteristics of nucleotides allows information transfer. A more detailed description of the items follows; however, here we wish to point out that we allowed in our coding for students to link entity behavior and characteristics in a non-linear fashion.

Unlinked explanations

Many students were capable of incorporating some, but not all, of the ideas related to explaining sequence transfer during transcription. Some students provided a reasonable account of the complementary base-pairing while RNA polymerase activity occurs without attending to the relationship between RNA and DNA sequence. For example, an introductory student stated in response to version 3,

> Initially DNA is split by the RNA polymerase and is copied into mRNA. The RNA polymerase reads the DNA starting from promoters and moves from 3' to 5' copying the genetic code with bases except replacing thymine with uracil. Only certain bases are joined together and this ensures that the genetic code is copied correctly although there can be possible mutation which can affect how the RNA is transcribed.

This student describes base-pairing as only certain bases are joined together as well as RNA polymerase 'reads' and copies bases. We consider this type of response similar to a descriptive response, in that students are describing RNA polymerase activity without linking to the full phenomenon of sequence transfer or leveraging molecular or sub-molecular entities properties or activities. We were cautious when coding responses that included statements that DNA is copied into mRNA or used terms like 'genetic code' without further elaboration. We wanted to identify only responses that were less likely to have been learned by rote memory and thus chose not to code DNA copied to mRNA or genetic code copied or similar as sequence transfer.

Another type of response that we encountered was mechanistic base pairing without linking to RNA polymerase or sequence transfer. For example, this introductory student stated in response to version 2,

> The double helix two-stranded structure of DNA allows the information within DNA to be accurately transmitted. Each nitrogen base within the helix is uniformly paired to a complementary base through hydrogen bonding interactions. The purine-pyrimidine pairs are specifically bonded to one another so that when the DNA strands unwind, the template strand that be transcribed to create a complementary pre-mRNA strand. RNA is a single stranded molecule composed of nucleotides lined by phosphodiester bonds. There are also complementary bases but instead of thymine, there is the pyrimidine base uracil. Though RNA is single stranded, it can fold on itself which can be stabilized by small areas of base pairing. This allows the RNA molecule to stabilize throughout the translation process.

This student uses both hydrogen bonds and purine-pyrimidine interactions explain complementary base pairing in the context of DNA then extend those characteristics and activities to RNA. This student included a thorough description of base pairing in DNA, RNA structure, and during transcription. This type of explanation in which students give descriptions of the mechanism of base-pairing without linking back the phenomenon or to RNA polymerase activity is rare (3% of responses).

Research Question 2. Which item characteristics are more likely to elicit a mechanistic explanation about gene expression?

We next asked whether one or more of the item versions were more likely to elicit mechanistic explanations than others. We found that the lowest percent of mechanistic (Molecular and Sub-molecular combined) explanations arose from versions 2 and 3 (29% each) and the highest from version 5 (49%; see Figure 1). Version 1 resulted in 40% mechanistic explanations and version 4 in 33%. We found that version 5 elicited a larger percentage of Sub-molecular explanations than any other version.

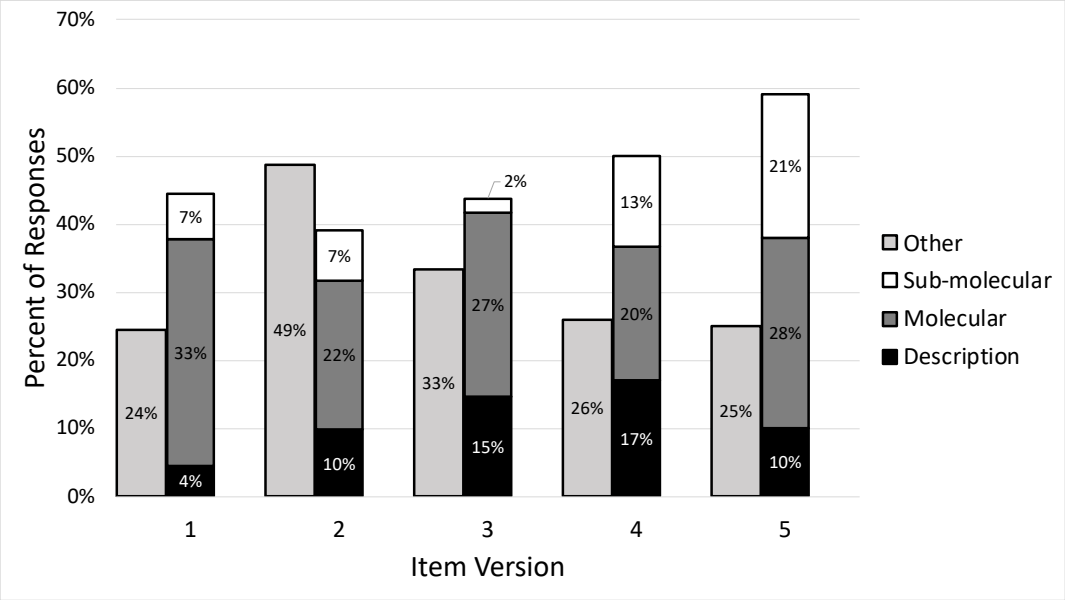


Figure 1. Students use more sub-molecular mechanistic reasoning in response to version 5 than other versions. Responses categorized as Mechanistic Base-Pairing and Descriptive Complementary not included here for clarity; thus, categories may not total to 100%. Responses: Version 1, n = 45; Version 2, n = 41; Version 3, n = 48; Version 4, n = 112, Version 5, n = 100.

Revisiting the wording of the items, all five versions included a prompt for students to explain how information is transferred from DNA to RNA during transcription plus a prompt to describe or use the characteristics of DNA and RNA to help explain information storage. Versions 4 and 5 included three text entry boxes, one for the explanation, a second for DNA and RNA characteristics and information storage and a third in which they were prompted to describe or use the characteristics of the nucleotide to explain information transfer. Many students responded to parts B and C of versions 4 and 5 by naming components of a nucleotide: nitrogenous base, sugar, and phosphate. While this was not strictly necessary for a student to describe information transfer, it did seem to prompt students to think about the molecular level; that is the role of the base and subsequently about the role of non-covalent interactions.

For example, in response to part A, an introductory student stated in response to version 5, “Copying part of a nucleotide sequence of DNA into a complementary sequence in RNA. This process is called process of transcription. It includes enzymes like DNA polymerase.” In response to part B, they stated “Two nucleotides are bound together by phosphate bond. Two strands of DNA and RNA are held together by hydrogen bonds between two nitrogen bases.” In response to part C they stated, “Nucleotides have nitrogen bases on one of their ends and those nitrogen bases keep the encoded information in them. These Nucleotides are held together by phosphate bond which do not break easily but two nitrogen bases are held together by hydrogen bond which are easy to break, thus can help in uncoiling of DNA while no breaking the whole structure.” This student described hydrogen bonding as related to complementary base pairing and sequence transfer in response to part B and went on to explain the role of the base vs. the role of the nucleotide in part C.

Versions 1, 2, and 3 were also developed with the intent to elicit sub-molecular and molecular ideas. Specifically, version 3 prompted students to think about the ‘molecular events,’ which should prompt a molecular level explanation, but student responses were similar in proportion of molecular explanations to versions 1 and 2, and very few (2%)

students were categorized as Sub-molecular Mechanism in response to version 3 (Figure 1). Thus, simply calling attention to molecular level isn't enough to elicit molecular thinking. It appears that the scaffolding in the form of multiple prompts and entry boxes, which draw attention first to the process, then names and asks for explanations about the macromolecular (DNA and RNA) and molecular (nucleotides) appear to be more productive ways to prompt students to provide a mechanistic explanation. The wording of item version 5 indicated that a new RNA was formed, which was intended to provide the element of time as a point of comparison for students (as suggested by Krist et al., 2018). As version 5 elicited the most sub-molecular explanations, the addition of the time element may have enhanced the effect of scaffolding.

Research Question 3. What additional resources or ideas do undergraduates use when writing their explanations of transcription?

We were interested in other ideas or related processes student drew upon in their responses and thus applied the conceptual rubric to categorize and quantify context. We captured seven other concepts that students used in their explanations: Gene Regulation, which includes the role of the promoter or transcription factors; mRNA processing, which includes post-transcriptional modifications to RNA such as splicing; Translation, where students state the RNA will be translated to a protein; Phase Naming, in which students say there are or name the three phases of transcription; Cellular Structure, in which students state where in the cell a process or molecule occurs; General Nucleic Acid Characteristics, in which students explain DNA and/or RNA structure such as describing the base, sugar, and phosphate components of a nucleotide or the double-helical nature of DNA; and Nature of Genetic Information, in which responses state genetic information is present in the sequence or that nucleic acids code for proteins. We found that the most frequent context idea was Nature of Genetic Information, which occurred in 63% of responses and the least common context idea was mRNA Processing, which occurred in 10% of responses. (See Supplemental Table 2 for details and examples).

We asked whether student responses categorized as Molecular or Sub-molecular would include similar or more context ideas than student responses categorized as Descriptions or as Other. We found that students included a median of two context ideas in responses categorized as Other or Descriptions and a median of three context ideas in responses categorized as Molecular or Sub-molecular (Figure 2). We performed a Kruskal-Wallis test and found a difference in the numbers of ideas included based on explanation type (Kruskal-Wallis $H = 41.590$, $df = 3$, $p < 0.005$). To more closely examine the differences in number of ideas included based on description type, we performed pairwise Mann-Whitney tests and found a significant difference between the median ideas in responses categorized as Other and Molecular ($p < 0.005$) and Sub-molecular ($p < 0.005$). Descriptions differed slightly from Molecular ($p < 0.065$) and from Sub-molecular ($p = 0.013$; see Supplemental Table 6 for effect size). Response lengths differed; the median word count of a Sub-molecular explanation was 171.5 words, Molecular was 140 words, Descriptions 89 words, and Other 79.5 words (Supplemental Figure 1). Thus, students who wrote a Molecular or Sub-molecular explanation included more content as measured by word count and more context as measured by number of ideas included than students who wrote a Description or Other responses.

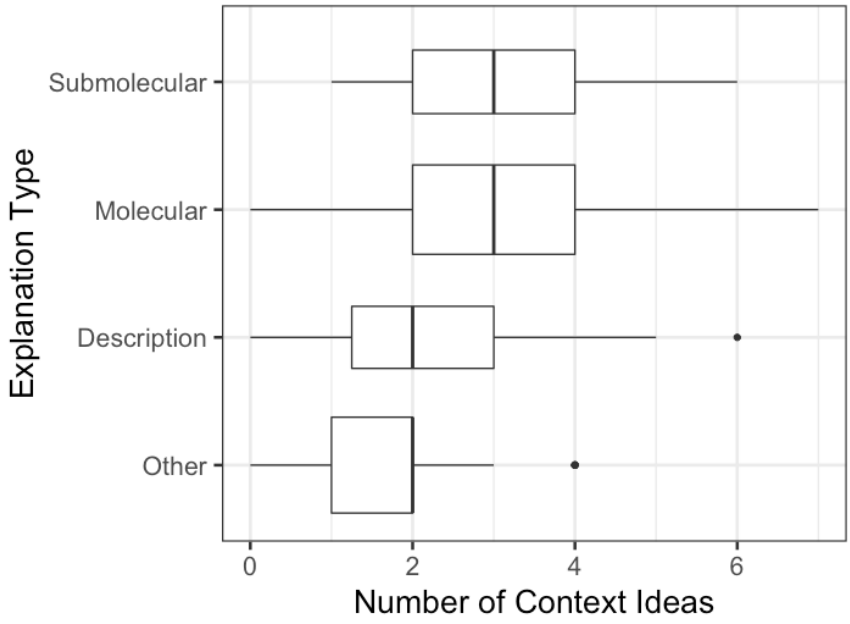


Figure 2. Students include different numbers of context ideas depending on the type of explanation they write. Width of boxes is proportional to the number of responses: Other, n = 100; Description, n = 42; Molecular Mechanism, n = 87; Sub-molecular Mechanism, n = 44. There is evidence of a significant difference in number of ideas included in student responses based on type or explanation (Kruskal-Wallis H = 41.590, df = 3, $p < 0.005$).

To explore the context that students included alongside their explanations, we identified the most frequent context idea in student responses categorized as a Description, Molecular, or Sub-molecular, which was the idea Nature of Genetic Information (86%, 78%, 89% respectively; Supplemental Table 5). This idea is positively correlated to each of the three response types based on the Pearson correlation (Supplemental Table 6). The frequency of Nature of Genetic Information in these categories is partially an effect of the coding rubric and rules as the rubric allowed combining Nature of Genetic Information ideas with the concept that RNA is made by sequential addition of bases into the phenomenon of Sequence Transfer. However, presence of the Nature of Genetic Information idea was insufficient for phenomenon in coding rules, and 43% of responses categorized as Other also included this idea. The second most frequent context idea was General Nucleic Acid Characteristics, which occurred in 40% of Descriptions, 60% of Molecular Mechanisms, and 64% of Sub-molecular Mechanisms. General Nucleic Acid Characteristics was also the most frequent context idea occurring in Other Reponses (50%). However, there was a negative correlation between General Nucleic Acid Characteristics with Description, and no correlation with other response types. The context of Gene Regulation occurred frequently in, and was positively correlated with, both Molecular and Sub-molecular explanations but not responses categorized as Other Information. Finally, we highlight that no context ideas are positively correlated with responses categorized as Other, context ideas are either negatively or not correlated with Other responses. It appears that students using Molecular or Sub-molecular explanations include more context ideas related to ideas about gene regulation and genetic information than students using other explanation types.

Discussion

In this study, we found that 38% of responses from undergraduate biology students include complementary base pairing as the reason RNA polymerase synthesizes an RNA with a sequence based on the DNA. About a third of those responses (13% of the total) add to these ideas that non-covalent nucleotide interactions drive complementary base pairing. We found that items which separately prompted students to respond about how DNA holds information and the role of nucleotides in transferring information were most successful at eliciting explanations that included a sub-molecular activities and interactions. Further, we found that students who provided a mechanistic explanation included more context than non-mechanistic explanation and were more likely to include ideas about gene regulation and genetic information.

Students who use a description about transcription (e.g., RNA polymerase transfers the sequence), or similarly, students who give an Unlinked response like those describing that the RNA polymerase synthesizes a complementary RNA, without linking to sequence transfer or non-covalent interactions may have mechanistic conceptions about transcription, meaning they recognize the need for a mechanism of sequence transfer. Reasons these students may not include the full mechanism include they do not recognize which additional entities interact to produce an effect. We suggest that instruction focus on helping students build on the ideas they hold by encouraging discussion about the application of ideas they may have memorized such as applying the idea of base-pairing to the phenomenon of sequence transfer. Some students CRs were categorized as Unlinked because they had mis-assigned the role of RNA polymerase. For example, some students stated that the mRNA itself transcribes RNA or claimed that another entity (e.g., the ribosome or helicase) synthesizes RNA. These students might be considered to have mechanistic conceptions of how transcription works and have simply mistaken terms, which is a documented learning challenge (Zukswert et al., 2019). In these cases, instructors might develop or utilize tools to help clear up confusion between the three Central Dogma processes (e.g. (Pelletreau et al., 2016).

Chemical concepts underlie many molecular biological phenomena, but students struggle to make connections between the two fields (e.g., Allred et al., n.d.; Kohn et al., 2018; Loertscher et al., 2014). We found this in our work about information transfer as well; many students could provide an explanation that included complementary base-pairing as a mechanism, but they less frequently included the non-covalent interactions between bases. This is a concern for instruction, as memorizing a set of base pairs (A-T or A-U and G-C) makes it difficult for an instructor to know whether the student considers these 'letters' to represent molecules with interactions or if the student only uses them procedurally. That is, a student can use the base pairs to produce a DNA-RNA sequence or even a protein sequence using heuristics without understanding the cause or effect behind the rules. The underlying cause of base-pairing is non-covalent interactions, which underlie many other molecular and cellular phenomena, such as protein-protein interactions. Loertscher et al., (2014) identified that students struggle with understanding non-covalent interactions, rather they memorized types of interactions. The authors also found that when students understand these interactions, they can begin to fully understand and make connections between structure and function, a core biological and scientific concept (Loertscher et al., 2014). We suggest that specifically prompting for nucleotide characteristics, as in item versions 4 and 5, may promote students making connections between chemical interactions and their relationship to biological phenomena like base-pairing.

In our work, we found that students were more likely to elaborate on the idea that the information transfer between DNA and RNA is in the form of a nucleotide sequence when we prompted separately for how DNA and RNA hold information. Duncan and Reiser (2007) described genetic information flow between genes and proteins as a genetic hybrid hierarchical phenomenon, which requires understanding and connecting between two organizational levels:

information level (genes) and the physical level (hierarchical entities like proteins, cells, tissues). The authors suggested that this is one of the main challenges of learning about molecular genetics (Duncan & Reiser, 2007). In our assessment items, getting students to think about the nature of information by prompting in part two is related to students providing a mechanistic explanation. Thus, we suggest that using this item as a formative assessment could aid in driving student thinking about the information level in addition to the physical/mechanical aspects of transcription and improve student ability to construct a mechanistic explanation.

Limitations

Responses collected for versions 4 and 5 were collected shortly after COVID-19 forced instruction to move online, thus we hesitate to draw broad conclusions about student ability to write a mechanistic explanation of transcription, as this rapid shift to remote learning may have affected student ability to participate in class, access materials, or focus on the face of potential personal concerns. We are currently working on a study to characterize student explanations under the current circumstance of planned remote learning (rather than emergency remote learning). Because this study took place at one large research university, student demographics may not be representative of the larger student population.

Acknowledgements

The authors thank the Automated Analysis of Constructed Response (AACR) collaboration for helpful conversations while developing the item and coding rubric, especially Megan Shiroda, Jenifer Saldanha, Leonora Kaldaras, and John Merrill. Details about the prompt “DNA to RNA Information Transfer” can be found at beyondmultiplechoice.org. This material is based upon work supported by the National Science Foundation (DUE grant 1323162).

References

Allred, Z. D. R., Farias, A. J., Kararo, A. T., Parent, K. N., Matz, R. L., & Underwood, S. M. (n.d.). *Students' use of chemistry core ideas to explain the structure and stability of DNA*. 14.

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action. Final report. Washington, DC.*

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, *95*(4), 639–669. https://doi.org/10.1002/sce.20449

Briggs, A. G., Hughes, L. E., Brennan, R. E., Buchner, J., Horak, R. E. A., Amburn, D. S. K., McDonald, A. H., Primm, T. P., Smith, A. C., Stevens, A. M., Yung, S. B., & Paustian, T. D. (2017). Concept Inventory Development Reveals Common Student Misconceptions about Microbiology. *Journal of Microbiology & Biology Education*, *18*(3). https://doi.org/10.1128/jmbe.v18i3.1319

Briggs, A. G., Morgan, S. K., Sanderson, S. K., Schulting, M. C., & Wieseman, L. J. (2016). Tracking the Resolution of Student Misconceptions about the Central Dogma of Molecular Biology. *Journal of Microbiology & Biology Education*, *17*(3), 339–350. https://doi.org/10.1128/jmbe.v17i3.1165

Castro-Faix, M., Duncan, R. G., & Choi, J. (2021). Data-driven refinements of a genetics learning progression. *Journal of Research in Science Teaching*, *58*(1), 3–39. https://doi.org/10.1002/tea.21631

Clemmons, A. W., Timbrook, J., Herron, J. C., & Crowe, A. J. (2020). BioSkills Guide: Development and National Validation of a Tool for Interpreting the Vision and Change Core Competencies. *CBE—Life Sciences Education*, *19*(4), ar53. https://doi.org/10.1187/cbe.19-11-0259

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cooper, M. M., Caballero, M. D., Ebert-May, D., Fata-Hartley, C. L., Jardeleza, S. E., Krajcik, J. S., Laverty, J. T., Matz, R. L., Posey, L. A., & Underwood, S. M. (2015). Challenge faculty to transform STEM learning. *Science*, *350*(6258), 281–282. https://doi.org/10.1126/science.aab0933

Duncan, R. G. (2007). The Role of Domain-Specific Knowledge in Generative Reasoning about Complicated Multileveled Phenomena. *Cognition and Instruction*, *25*(4), 271–336.

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, *44*(7), 938–959. https://doi.org/10.1002/tea.20186

Fisher, K. M. (1985). A misconception in biology: Amino acids and translation. *Journal of Research in Science Teaching*, *22*(1), 53–62. https://doi.org/10.1002/tea.3660220105

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Gericke, N., & Wahlberg, S. (2013). Clusters of concepts in molecular genetics: A study of Swedish upper secondary science students understanding. *Journal of Biological Education*, *47*(2), 73–83. https://doi.org/10.1080/00219266.2012.716785

Haskel-Ittah, M., & Yarden, A. (2018). Students' Conception of Genetic Phenomena and Its Effect on Their Ability to Understand the Underlying Mechanism. *CBE—Life Sciences Education*, *17*(3), ar36. https://doi.org/10.1187/cbe.18-01-0014

Jescovitch, L. N., Scott, E. E., Cerchiara, J. A., Doherty, J. H., Wenderoth, M. P., Merrill, J. E., Urban-Lurain, M., & Haudek, K. C. (2019). Deconstruction of Holistic Rubrics into Analytic Rubrics for Large-Scale Assessments of Students' Reasoning of Complex Science Concepts. *Practical Assessment, Research & Evaluation*, *24*(7). https://par.nsf.gov/biblio/10112967

Krist, C., Schwarz, C. V., & Reiser, B. J. (2018). Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences*, 1–46. https://doi.org/10.1080/10508406.2018.1510404

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. JSTOR. https://doi.org/10.2307/2529310

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, *67*(1), 1–25. https://doi.org/10.1086/392759

Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, *34*(4), 200–205. https://doi.org/10.1080/00219266.2000.9655718

Matz, R. L., Fata-Hartley, C. L., Posey, L. A., Laverty, J. T., Underwood, S. M., Carmel, J. H., Herrington, D. G., Stowe, R. L., Caballero, M. D., Ebert-May, D., & Cooper, M. M. (2018). Evaluating the extent of a large-scale transformation in gateway science courses. *Science Advances*, *4*(10), eaau0554. https://doi.org/10.1126/sciadv.aau0554

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *Journal of the Learning Sciences*, *15*(2), 153–191. https://doi.org/10.1207/s15327809jls1502_1

Moscarella, R., Haudek, K., Knight, J., Mazur, A., Pelletreau, K., Prevost, L., Smith, M., Steele, M. M., Urban-Lurain, M., & Merrill, J. (2016). *Automated Analysis Provides Insights into Students' Challenges Understanding the Processes Underlying the Flow of Genetic Information*. National Association for Research in Science Teaching.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education.* National Academies Press. https://doi.org/10.17226/13165

Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, *45*(10), 1131–1160. https://doi.org/10.1002/tea.20251

NGSS Lead States. (2013). *Next Generation Science Standards; For States, By States*. The National Academies Press. https://www.nextgenscience.org/

Pellegrino, J. W., DiBello, L. V., & Goldman, S. R. (2016). A Framework for Conceptualizing and Evaluating the Validity of Instructionally Relevant Assessments. *Educational Psychologist*, *51*(1), 59–81. https://doi.org/10.1080/00461520.2016.1145550

Pelletreau, K. N., Andrews, T., Armstrong, N., Bedell, M. A., Dastoor, F., Dean, N., Erster, S., Fata-Hartley, C., Guild, N., Greig, H., Hall, D., Knight, J. K., Koslowsky, D., Lemons, P., Martin, J., McCourt, J., Merrill, J., Moscarella, R., Nehm, R., … Smith, M. K. (2016). A clicker-based case study that untangles student thinking about the processes in the central dogma. *CourseSource*, *3*. https://doi.org/10.24918/cs.2016.15

Prevost, L. B., Smith, M. K., & Knight, J. K. (2016). Using Student Writing and Lexical Analysis to Reveal Student Thinking about the Role of Stop Codons in the Central Dogma. *CBE—Life Sciences Education*, *15*(4), ar65. https://doi.org/10.1187/cbe.15-12-0267

Reiser, B. J., Berland, L. K., & Kenyon, L. (2012). Engaging Students in the Scientific Practices of Explanation and Argumentation. *Science and Children*, *49*(8), 8–13.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525. https://doi.org/10.1002/sce.20264

Sandoval, W. A., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction*, *23*(1), 23–55. https://doi.org/10.1207/s1532690xci2301_2

Smith, M. K., & Knight, J. K. (2012). Using the Genetics Concept Assessment to Document Persistent Conceptual Difficulties in Undergraduate Genetics Courses. *Genetics*, *191*(1), 21–32. https://doi.org/10.1534/genetics.111.137810

Smith, Michelle K., Wood, W. B., & Knight, J. K. (2008). The Genetics Concept Assessment: A New Concept Inventory for Gauging Student Understanding of Genetics. *CBE—Life Sciences Education*, *7*(4), 422–430. https://doi.org/10.1187/cbe.08-08-0045

Southard, K. M., Melissa R. Espindola, Zaepfel, S. D., & Bolger, M. S. (2017). Generative mechanistic explanation building in undergraduate molecular and cellular biology. *International Journal of Science Education*, *39*(13), 1795–1829. https://doi.org/10.1080/09500693.2017.1353713

Southard, K., Wince, T., Meddleton, S., & Bolger, M. S. (2016). Features of Knowledge Building in Biology: Understanding Undergraduate Students' Ideas about Molecular Mechanisms. *CBE—Life Sciences Education*, *15*(1), ar7. https://doi.org/10.1187/cbe.15-05-0114

Todd, A., & Romine, W. (2018). The Learning Loss Effect in Genetics: What Ideas Do Students Retain or Lose after Instruction? *CBE—Life Sciences Education*, *17*(4), ar55. https://doi.org/10.1187/cbe.16-10-0310

Trujillo, C. M., Anderson, T. R., & Pelaez, N. J. (2015). A Model of How Different Biology Experts Explain Molecular and Cellular Mechanisms. *CBE—Life Sciences Education*, *14*(2), ar20. https://doi.org/10.1187/cbe.14-12-0229

Uhl, J. D., Sripathi, K. N., Saldanha, J. N., Moscarella, R. A., Merrill, J., Urban-Lurain, M., & Haudek, K. C. (2020). Introductory biology undergraduate students' mixed ideas about genetic information flow. *Biochemistry and Molecular Biology Education*, 1–11. https://doi.org/10.1002/bmb.21483

van Mil, M. H. W., Boerwinkel, D. J., & Waarlo, A. J. (2013). Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Science & Education*, *22*(1), 93–118. https://doi.org/10.1007/s11191-011-9379-7

Weston, M., Haudek, K. C., Prevost, L., Urban-Lurain, M., & Merrill, J. (2015). Examining the Impact of Question Surface Features on Students' Answers to Constructed-Response Questions on Photosynthesis. *CBE—Life Sciences Education*, *14*(2), ar19. https://doi.org/10.1187/cbe.14-07-0110

Wright, L. K., Fisk, J. N., & Newman, D. L. (2014). DNA → RNA: What Do Students Think the Arrow Means? *CBE—Life Sciences Education*, *13*(2), 338–348. https://doi.org/10.1187/cbe.cbe-13-09-0188

Zukswert, J. M., Barker, M. K., & McDonnell, L. (2019). Identifying Troublesome Jargon in Biology: Discrepancies between Student Performance and Perceived Understanding. *CBE—Life Sciences Education*, *18*(1), ar6. https://doi.org/10.1187/cbe.17-07-0118

Supplemental Materials

Supplemental Table 1. Mechanistic Coding Rubric and Example Student Responses. All student responses were spell-checked. Underlining by authors to highlight coded phrases.

Category	Description	Example Student Responses
1. Transfer of Base Sequence	Student states or implies that the base sequence is transferred from DNA to RNA. Student may use terms like 'copy' to describe transfer of the RNA sequence. Student responses may imply sequence transfer if they combine a description of a DNA strand as having a sequence with the idea that the strand is copied to the new RNA; that the sequence has information with RNA being made base by base; state that the RNA is identical except Ts and Us; or provide an example DNA sequence plus a correct corresponding RNA sequence.	"RNA polymerase copies a DNA sequence to create an RNA sequence. They have nucleic acids that hold information that can be used to transfer information from one. DNA and RNA contain nucleotides (nucleic acids that hold the genetic information). They transfer information by binding to their complementary base pairs." Response 1406, item version 5. Also coded as 2, 3.2 "A strand of DNA is copied into a new molecule called mRNA. The DNA is read from the 3 end to the 5 end and is replicated. DNA and RNA hold information through its sequence of nucleotides. DNA is held within the cells nucleus. The sequence of nucleotides makes up the DNA and is a genetic code for amino acids. They store information. The nucleotides on the DNA can be read and copied to be transcribed as RNA." Response 5572, item version 5. "DNA strands are split and transcribed into mRNA strands by RNA polymerase. The nucleotides (DNA: ATGC; RNA: AUGC) code for certain proteins. The nitrogenous base pairs are arranged in differing sequences allowing for different proteins to be synthesized." Response 9357, item version 4. Also coded as 2.
2. RNA Polymerase Activity	Students correctly describe the function of RNA Polymerase in synthesizing a new RNA molecule, including student language like 'moving along', 'coding' or 'reading' the DNA.	"Transcription is when a DNA sequence is copied via RNA polymerase to create a RNA. In other words, it is genetic information being transferred starting from DNA to RNA. The process functions as DNA>RNA. The first step involved DNA being elongated and is transcribe into RNA. The old segment is transcribed from 3'-5' leading for the complementary to create in the 5'-3' direction. This results in the non-template strand being formed having the same sequence as the synthesized RNA. The end product of translation results in the making of mRNA and later translation." Response 4380, item version 3. Also coded as 1. "RNA Polymerase opens the double helix and then synthesizes an RNA strand from the DNA. This will create an mRNA which can further be synthesized to

		make a protein DNA and RNA are long strands of bases which provide genetic code for genes. Nucleotides contain an OH group which allow them to form phosphodiester bonds elongating the DNA strand. These nucleotides can be transferred by the RNA binding complementary to the DNA template and thus the code is transferred" Response 8553, item version 4. Also coded as 3.1
3.1 Complementary base-pairing, Present and Unlinked	Students describe RNA as complementary to DNA but do not use it to explain how sequence transfer occurs between DNA and RNA.	"During transcription, DNA passes the information it possesses to mRNA. DNA has two strands, which become separated by a single RNA strand. This single strand reads the nucleotides of the one DNA strand and uses them to create complementary nucleotides, with the base uracil instead of thymine. When this is completed, the messenger RNA (mRNA) then carries the information out into the cytoplasm for translation to occur." Response 2754, item version 1. Also coded as 1.
3.2 Complementary base-pairing, Linked to Sequence Transfer	Students describe the process of sequence transfer by describing the RNA nucleotides as complementary. These responses must link the idea that the RNA is complementary to a sequence being transferred from DNA to RNA.	The DNA coding strand codes the RNA in transcription. The DNA is complimented by RNA base pairs and so the information is transferred in terms of complimentary bases, that are reflecting the template strand. The nitrogenous bases in particular are the defining factors of DNA and RNA that hold specific information, the letter for the DNA and RNA. The pattern of these nucleotides then creates information as a sequence. This is then valued as codons, 3 letter sequences, that code for specific amino acids. The characteristics that hold information are the 2' carbon attachments and the nitrogenous base. The 2' carbon substituent has an H (Ribose) or OH (Deoxyribose). The nitrogenous base defines what will pair and is the defining factor of the nucleic acid as it is the letter represented." Response 2639, item version 5. Also coded as 1.
4.1 Non-covalent Interactions, Present and Unlinked	Students name hydrogen bonding, or another non-covalent interaction but do not use these ideas to explain DNA/RNA complementary base-pairing. For example, students may explain that hydrogen bonding is involved in some process other than transcription or describe some other function of noncovalent bonds.	"After replication of DNA, DNA is transcribed into RNA using RNA polymerase. Initiation of transcription starts RNA polymerase finding a promoter. The promoter tells RNA polymerase where to start transcription. The DNA is unwound, and RNA polymerase starts transcribing. Once RNA polymerase reaches a terminating sequence, the mRNA is released from RNA polymerase. DNA is double stranded deoxyribose which doesn't have a hydroxyl on carbon 2. DNA can form a double helix with its bases in the center and phosphates on the outside making it very stable. RNA is single stranded and is a ribose. RNA contains uracil and DNA contains thymine. Both are made up of nucleotides linked by phosphodiester bonds. These nucleotides help with transferring information for cell functions. Nucleotides are made up of a deoxyribose/ribose, phosphate

		group, and a base. In DNA/RNA nucleotides are liked up through phosphodiester bonds which connect to form strands of DNA/RNA. The bases in nucleotides can form hydrogen bonds which make them super strong. This allows them to carry information for replication/transcription/translation." Response 1754, item version 4. Also coded as 2.
4.2 Non-covalent Interactions, Linked to Complementarity	Students describe hydrogen bonding or other non-covalent interactions between bases as the reason complementary base-pairing can occur. Examples of interactions include shape or conformation, functional group interactions, IMFS, and purine/pyrimidines interact.	"In the first stage of transcription, the RNA polymerase binds to the promoter region of the DNA template strand and opens the helix. The RNA polymerase adds nucleotides complementary to the template strand, reading 3' to 5'. Through hydrogen bonds, the nucleic acids pair with their counterparts; uracil is placed as adenine's complement, as opposed to thymine. The RNA is transcribed 5' to 3', creating the mRNA sequence." Response 2587, item version 3. Also coded as 1, 2, 3.2. "DNA helicase unzips the DNA, allowing for RNA polymerase to bind at the transcription start site, as well as transcription factors. RNA pol moves down the strand, creating a RNA strand that is complementary and anti-parallel to the DNA strand. Both DNA and RNA contain a phosphate backbone that holds on it nucleotide bases. These bases, depending on their sequence, code for specific proteins in the body. The different kinds of RNA complete different responsibilities, some aiding in the transfer of information and some aiding in the process of that transfer. DNA is double stranded, held together by hydrogen bonding between bases. This allows for great stability, unlike RNA which is single stranded. Nucleotides can bind via hydrogen bonds to one another, however due to their conformations, they can only bind to specific other nucleotides. This allows for nucleotide matching when creating RNA from DNA, or replicating DNA from DNA. By matching nucleotides based on their conformations, information is able to be transfered." Response 2727, item version 4. Also coded as 1, 2, 3.2

Supplemental Table 2. Rubric categories capturing context in student responses. Responses were spell checked. Underlining by authors to highlight coded phrases. All context categories can co-occur with each other, or with mechanistic rubric categories.

Category (percent responses)	Description	Example Student Responses
5. Gene Regulation (36%)	Students use the control of transcriptional activation to help explain their response. Students discuss the DNA or proteins involved in transcriptional regulation. For example, a student might describe the role of an operon, promoter, or transcription factors.	"During the process of transcription, the DNA double helix is unwound, separating the DNA into a template strand and a noncoding strand for a short series of nucleotides in a sort of transcription bubble. This process starts at a promoter region, where proteins, including RNA polymerase bind…" (excerpt from response 4302) "The process of transcription starts in the initiation phase. There, RNA polymerase ad other various proteins are brought towards DNA in the promoter region. The double helix structure of DNA is significant to this phase because both strands of DNA are needed to bring the initiation proteins. In order to start transcription, at least six general transcription factors and a least transcriptional activator protein is needed at the promotor region. These proteins bind with enhancer DNA sequences that attract a mediator complex of proteins and then attracts the RNA polymerase (pol II)." (excerpt from response 4351) "During transcription factors or operons pull RNA Pol to the DNA where transcription starts with the help of free nucleotides." (excerpt from response 4695)
6. mRNA Processing (10%)	Students describe post-transcriptional processing of mRNA as part of their response. They may state that mRNA undergoes splicing, polyadenylation, or capping, or imply that there is a post-transcriptional modification.	"Simultaneously, the RNA polymerase begins to add complementary RNA nucleotides onto the growing primary RNA strand. This primary strand eventually becomes modified into the mRNA strand that is used during translation." (excerpt from response 9082) "…which will release the RNA molecule from the RNA polymerase. After transcription the RNA molecules then are spliced and have a 5' cap and poly-A tail put on their ends." (excerpt from response 1292)
7. Translation (28%)	Students state that the mRNA produced by transcription will be used in translation to synthesize a protein.	"The mRNA is then translated into a polypeptide of amino acids that form a protein using a ribosome." (excerpt from response 0293) "The mRNA sequence is used to assemble-- in order-- the chain of amino acids that form a protein." (excerpt from response 2982)

8. Phase Naming (17%)	Students say there are three phases in transcription or name two or more of the three major phases of transcription: 1-initiation, 2-elongation, and 3-termination. They may also use the verb forms of these phases, e.g., initiated, elongated, terminated. This category can co-occur with any other category, even when students provide a complete, detailed response and simply use the three phases to frame their response	"Transcription is the first step in gene expression. It begins by copying a gene's DNA sequence to make an RNA molecule. Transcription is performed by enzymes called RNA polymerases, which link nucleotides to form an RNA strand. Transcription has 3 stages. Long linear polymers called nucleic acids can hold information and pass it down to other generations. These chains can make proteins. Different combinations of nucleotides allow them to store so much information in many different ways. Base pairing between nucleotides creates a chain which can signal for different molecules" (Response 0249) "Transcription has 3 parts: Initiation, elongation, and termination. In the first part Polymerase is formed and the sigma portion of the enzyme joins to a promoter end of DNA. Polymerase splits the DNA double helix strand to begin transcribing it into RNA. In the next part RNA begins to be synthesized from the template strand of DNA. RNA moves from the 3' to the 5' of DNA and synthesizes from 5' to 3'. Complimentary base pairs form with the DNA making mRNA. Once the polymerase reaches the termination codon it breaks off from the protein synthesizing site. DNA is a double helix and contains the subunit T that is better at holding Genetic information than the Uracil subunit in RNA." (complete response 3369)
9. Cellular Structure (27%)	Students refer to the basic aspects of cell structure as part of their response. They may describe or name a cellular structure in which the process they describe occurs.	"This new strand of RNA is singe stranded and unstable and is transferred through the cytoplasm to be processed" "DNA has the ability for long-term storage of genetic information and the transmission of genetic information to make other cells. RNA is used for short term storage and to transfer the genetic code from the nucleus (in eukaryotes) to the ribosome to make proteins. They can transfer genetic information through their codes (ATCG/AUCG)." (excerpt from response 2972)
10. General Nucleic Acid Characteristics (55%)	Students include molecular characteristics of DNA or RNA structure in their response. Responses may include one or more structural terms such as: sugar phosphate backbone, phosphodiester, double-helix.	"Essentially, each monomer or nucleotide is a piece of information/coding. These monomers within DNA/RNA can connect due to the negatively charged phosphate groups and free -OH group on 3' carbons on the appropriate sugar. This allows for the complex nature of DNA (genetic material) because there are so many potential combinations. A nucleotide consists of a sugar (ribose or deoxyribose), a phosphate group, and a nitrogenous base (purine or pyrimidine.) The monomers can bond together via phosphodiester bonds in the 5' to 3' direction." (excerpt from response 3728)
11. Nature of Genetic Information (63%)	Student responses state genetic information as present in the sequence or order of DNA or RNA bases. Or student	"The central dogma describes the flow of information from DNA to RNA to Protein The specific nucleotide sequence allows DNA and RNA to hold specific information. Nucleotides hold genetic information and can use it to make specific proteins." (response 2475)

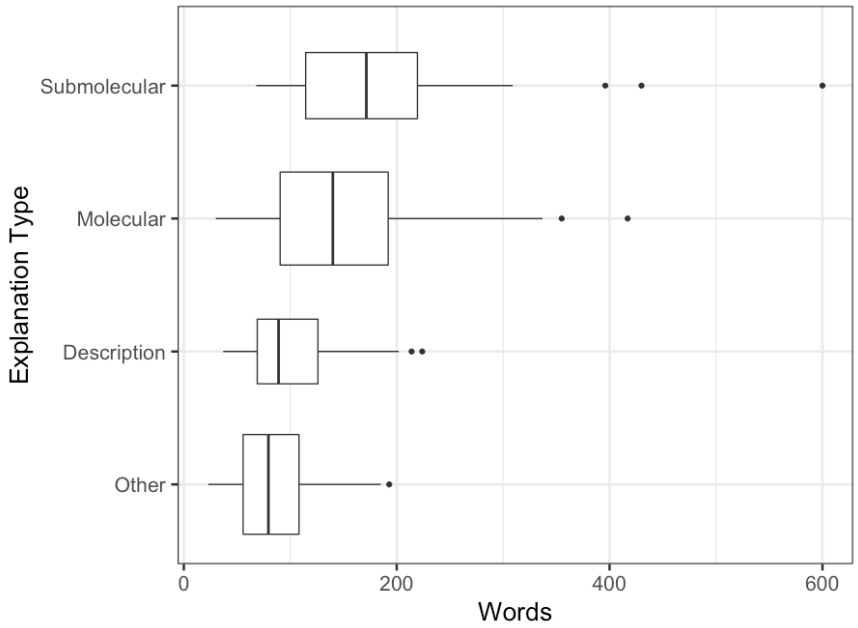
	describes genetic information by stating or implying that DNA and/or RNA codes for protein.	"The nitrogenous bases in particular are the defining factors of DNA and RNA that hold specific information, the letter for the DNA and RNA. The pattern of these nucleotides then creates information as a sequence. This is then valued as codons, 3 letter sequences, that code for specific amino acids." (excerpt from response 3141)

Supplemental Table 3. Inter-rater reliability for rubric categories after calibration.

Category	Coding	Cohen's Kappa	Percent Agreement
Nature of Genetic Information	Present/Absent	0.816	91%
RNA Polymerase Activity	Present/Absent	0.827	92%
Transfer of Base Sequence	Present/Absent	0.697	85%
DNA and RNA Complementarity	Linked/Present/Absent	0.697	80%
Non-covalent Interactions	Linked/Present/Absent	0.698	88%
General Nucleic Acid Characteristics	Present/Absent	0.746	87%
Phase Naming	Present/Absent	0.809	95%
Gene Regulation	Present/Absent	0.849	94%
mRNA Processing	Present/Absent	0.867	98%
Cellular Structure	Present/Absent	0.790	93%
Translation	Present/Absent	0.672	86%

Supplemental Table 4. Results of Mann-Whitney pairwise comparisons for number of context ideas included in each explanation type.

Holistic Category	Description (median = 2, n = 42)	Molecular (median = 3, n = 87)	Sub-molecular (median = 3, n = 44)
Other (median = 2, n = 100)	U = 1536.5 p = 0.009 Effect size, r = 0.22	U = 2457.5 p < 0.005 effect size, r = 0.39	U = 999.5 p < 0.005 effect size, r = 0.45
Description (median = 2, n = 42)		U = 1468.5 p = 0.065 effect size, r = 0.16	U = 644.0 p = 0.013 effect size, r = 0.27



Supplemental Figure 1. Student responses categorized as Mechanistic (Molecular or Sub-molecular) write longer responses than student responses categorized as Descriptions or Other. Width of boxes is proportional to number of responses in each category: Other, n = 100; Description, n = 42; Molecular Mechanism, n = 87; Sub-molecular Mechanism, n = 44.

Supplemental Table 5. Frequency of ideas included in student CRs based on explanation type. Responses: Other, n = 100; Unlinked, n = 73; Description, n = 42; Molecular, n = 87; Sub-molecular, n = 44.

	Explanation type				
Context Category	Other	Unlinked	Description	Molecular	Sub-molecular
5. Gene Regulation	21%	27%	33%	48%	61%
6. mRNA Processing	9%	8%	10%	13%	11%
7. Translation	19%	26%	40%	30%	39%
8. Phase Naming	11%	19%	17%	24%	14%
9. Cellular Structure	26%	25%	14%	33%	32%
10. General Nucleic Acid Characteristics	50%	59%	40%	60%	64%
11. Nature of Genetic Information	43%	42%	86%	78%	89%

Supplemental Table 6. Pearson's correlation between context ideas included in student CRs and explanation type. Responses: Other, n = 100; Unlinked, n = 73; Description, n = 42; Molecular, n = 87; Sub-molecular, n = 44. * $p < 0.04$, ** $p < 0.005$

	Explanation Type				
Context Category	Other	Unlinked	Description	Molecular	Sub-molecular
5. Gene Regulation	-.197**	-.091	-.019	.150**	.203**
6. mRNA Processing	-.024	-.033	-.007	.049	.016
7. Translation	-.132*	-.026	.100	.020	.087
8. Phase Naming	-.103	.029	-.004	.109*	-.035
9. Cellular Structure	-.013	-.026	-.106*	.084	.043
10. General Nucleic Acid Characteristics	-.063	.041	-.108*	.057	.067
11. Nature of Genetic Information	-.260**	-.217**	.177**	.185**	.205**