March 26, 2020

Scoring the "Integrative Complexity" of Student Responses: A New Strategy for Measuring Student Learning

INTRODUCTION AND BACKGROUND ON AUTOIC 2

Integrative Complexity (IC) 2

Applying Integrative Complexity to Student Learning: A Case Example 4

INSTRUCTIONS FOR USING THE AUTOMATED INTEGRATIVE COMPLEXITY PROGRAM 6

Background 6

Creating Your Qualtrics Survey 7

Downloading the Raw Survey Data 11

Preparing Your Data for AutoIC Analysis 12

Working with the AutoIC Website and Data 14

RESOURCES 15

REFERENCES 16

INTRODUCTION AND BACKGROUND ON AUTOIC

A challenge for all educational institutions, including the Harvard Kennedy School (HKS), is how to measure the extent to which teaching actually enriches and broadens students' thinking. As a professional school, a critical question for HKS is how our curriculum and pedagogy enhance students' capacity to analyze practical problems. Performance on assignments and exams demonstrates command of course material, but it does not measure whether the course experience has actually improved students' capacity to reason through course-relevant practical problems. Particularly in an institution such as HKS, we should actively question whether our students' stellar performance is a credit to our educational practices or to our capacity to recruit the best and brightest.

The current guide was inspired by the path-breaking work of Richard Light, Carl H. Pforzheimer Jr. Professor of Teaching and Learning at the Harvard Graduate School of Education (HGSE), on the improvement of educational experiences at institutions of higher learning. In his HKS Teaching and Learning Working Paper *An Assessment Strategy for Enhancing Sustained HKS Excellence*, Light advocates for measuring student learning using pre-post study designs that compare student responses to course-related practical questions at the beginning and the end of the course. One advantage of the pre-post design is that it controls for variation in students' knowledge and skills at the inception of the course when measuring changes in the quality of their reasoning following the course.

HKS has experimented with pre-post assessment designs in the past, using human coders to analyze students' responses to pre- and post-course assessments. However, as recognized by Light, the financial and time costs of employing qualified human coders to evaluate large numbers of written responses limits the accessibility of this approach. To mitigate these costs in large programs or classes, Light argues for evaluating representative sub-samples of student work.

We propose an alternative to human coding of pre-post learning evaluation by taking advantage of recent technological advances in the automated coding of human language. Specifically, we suggest employing Automated Integrative Complexity (AutoIC) coding software. In the following sections, we introduce the concept of "integrative complexity" and explain its relevance to the assessment of student learning. We also present the results of experiments with this method conducted at HKS. For those interested in experimenting with this approach, we provide detailed "how to" instructions.

Integrative Complexity (IC)

Integrative complexity (IC) is one of various constructs measuring cognitive complexity (i.e., the propensity to think about multiple dimensions of a problem). IC is measured by the extent to which one (a) differentiates among multiple dimensions of or perspectives on a problem ("differentiation") and then (b) draws connections among those dimensions or perspectives in an integrated argument ("integration") (Suedfeld, Tetlock, & Streufert, 1992). These dual measures of differentiation and integration distinguishes IC from other measures of cognitive complexity.

As a construct, IC is rooted in cognitive psychology, reaching back to the development of theories of cognitive style (e.g., Kelly, 1955). For instance, early studies found that the cognitive traits of authoritarianism (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950) and dogmatism (Rokeach, 1960) were negatively correlated with integrative complexity (Suedfeld et al., 1992). Later studies illuminated implications of IC for the management of political conflict. For example, an analysis of official Soviet and American foreign policy statements during the Cold War Era revealed lower integrative complexity (IC) scores prior to hostile initiatives and higher IC scores before peaceful agreements (Tetlock, 1985, 1988). Another study using archival material from leaders of historic revolutionary movements found that leaders who had low integrative complexity scores were most successful in overthrowing their governments. However, after taking power, the revolutionaries whose language increased in cognitive complexity were less likely to be ousted from government (Conway, Suedfeld, & Tetlock, 2001; Suedfeld & Rank, 1976).

In recent decades, more attention has been given to situational factors that influence the complexity of thought. For instance, studies show that the decisions by majority group members tend to score higher in integrative complexity when exposed to minority opinions than when working in homogenous groups (Antonio et al., 2004; Gruenfeld, 1995; Gruenfeld, Thomas-Hunt, & Kim, 1998). This is because, in order to address or challenge minority perspectives, majority group members must reconceptualize their arguments to account for alternative perspectives (Gruenfeld, 1995; Gruenfeld et al., 1998; Nemeth, 1992).

Measuring IC | Until recently, measures of IC have been based on time-intensive human coding systems that required high levels of agreement among multiple formally trained coders (e.g., Suedfeld et al., 1992). In 2014, Conway and colleagues published a study of an automated IC coding system (AutoIC), the results of which correlated well with those of human coders (Conway, Conway, Gornick, & Houck, 2014; see references for a link to this article). Human coders evaluate texts with a degree of nuance that cannot be replicated by automated coding. However, human and AutoIC coding are directionally consistent, and the AutoIC system is vastly more efficient (Conway et al., 2014; Tetlock, Metz, Scott, & Suedfeld, 2014). AutoIC has been lauded for its potential to broaden the application of the IC construct and to further refine IC theory and measurement (Tetlock et al., 2014).

AutoIC scores integrative complexity based on a search for words and phrases associated with differentiation (e.g., “on the other hand”) and with integration (e.g. “in conjunction with”). Because many words have multiple meanings depending on context, AutoIC uses a probabilistic scoring method such that words and phrases in the AutoIC dictionary are weighted based on the probability that they signal complexity. For instance, Conway et al. (2014) use the example of the phrases “on the other hand” and “apart from” as indicators of differentiation, but “on the other hand” would be scored higher than “apart from” because there are more uses of “apart from” that are unrelated to complex thinking (e.g., “I do not wish to be apart from you” vs. “apart from this reason, there is another reason why”). See Conway et al. (2014), for more details on the AutoIC scoring method.

Applying Integrative Complexity to Student Learning: A Case Example

We propose employing AutoIC technology in pre-post study designs of student learning. More specifically, we propose testing whether students' reasoning about course-relevant practical problems would be more integratively complex following a course than before. For courses that aim to teach students to think in multi-dimensional ways about practical problems, increased IC scores following (as compared to before) a course would be an indication of students' propensity to recognize more dimensions of course-relevant practical challenges and to integrate alternative considerations or perspectives into their prescriptive suggestions.

We applied this method to evaluate learning in a Management, Leadership & Decision Sciences (MLD) course in the Master in Public Policy (MPP) core curriculum at the Harvard Kennedy School. The module-length course (6 weeks) was taught in multiple sections of approximately 60 students. The primary learning objective of the course was to enhance students' capacity to work effectively in diverse and dispersed teams, in large part by helping them to recognize and mitigate the pitfalls, as well as maximize the benefits, of such work groups. In other words, the faculty aimed to help students think in analytically complex ways about work group design and process. Our aim was to test whether students would reason through course-relevant leadership challenges in more cognitively complex ways after taking the course, as compared to immediately before the course began.

We used IC coding systems to evaluate student learning using a pre-post study design. We conducted a pilot test in three of four sections of the course in the fall semester of 2016, and replicated the study in all four sections of the course in the fall semester of 2017. We report a detailed description of these studies in "A Detailed Case Example: Using Integrative Complexity as a Measure of Student Learning," including a fuller explanation of the procedure, analyses, and findings, as well as a comparison of the results automated IC and human IC coding. In short, as part of a pre-course survey distributed immediately before the module began, students briefly explained in writing how they would approach a set of course-relevant leadership challenges (e.g., leading a diverse and dispersed team, advising a high-stakes and heated stakeholder negotiation). Students were then asked to respond to these same problems as one part of the final exam. We coded students' written responses using the AutoIC system (for detailed instructions see p. 7) and compared the integrative complexity of their responses by question and course section.

For most questions in most sections of the course, we observed significant increases in the integrative complexity of the students' analyses of the course-relevant leadership challenges. When we observed a significant increase in students' scores in some but not all sections, this prompted a rich discussion among the faculty teaching different sections about how their course content and delivery had differed. While questions remain about how best to interpret the results and improve the testing procedures (see "A Detailed Case Example: Using Integrative Complexity as a Measure of Student Learning" for details), the faculty gained numerous insights from the process of designing a pre-post study, as well as from assessing the results.

The faculty benefited in at least three respects simply from the process of designing the study and collecting the data. First, it was a useful pedagogical exercise for the faculty to collectively identify three course-relevant practical problems that we hoped the students would analyze in more integratively complex ways following the course than before. Second, reading the students' responses to the pre-course "baseline" survey to those questions provided meaningful insight into students' preparedness for the course material. Third, the course-relevant problems became a tool for connecting academic course content with practical problem solving.

Finally, we gained insights from our analyses of the pre-post results. First, it was heartening that there was enough variation in IC to discern effects even when collecting only short responses from students. Second, when there was variation in the gains in questions and across the sections, those findings prompted constructive conversation among the faculty about how the teaching and course material had varied in delivery and points of emphasis. Finally, testing for correlations between pre-post difference in students' IC scores and students' grades helped us to interpret the results and consider what types of assignments enhance the students' capacity to think in multidimensional ways about course-relevant practical problems. We observed no significant correlations between improvements in students' IC scores and their overall exam performance, which suggested that changes in IC scores were not simply attributable to an exam effect. Testing for correlations between improvements in pre-post IC scores and grade components suggested that relevant paper assignments contributed more than other course requirements (e.g., class participation) to students' capacity to think in integratively complex ways about course-relevant practical problems.

In sum, we propose testing for pre-post gains in the integrative complexity of students' responses to course-relevant practical problems as one avenue for assessing student learning. Use of the AutoIC software enabled faculty to do this assessment more quickly and efficiently than with human coders. More broadly, we found that faculty are likely to benefit simply from the process of attempting to assess student learning, as well as from analyzing data collected. We are encouraged by the results of these studies that advances in automated coding systems will enable more faculty to engage in the learning cycle of pre-post assessment.

Finally, it warrants emphasis that we tested this method in a course in which faculty aimed to help students think in analytically complex ways about practical problems. Pre-post changes in IC would, therefore, be an indication of progress toward learning objectives. Not all courses have this character. For instance, indicators of changes in IC would be less relevant to courses in which progress could be measured by students' capacity to find unique correct answers (e.g., solutions to math problems). Faculty should consider whether increased integrative complexity would be a desired outcome and sign of progress in their courses.

INSTRUCTIONS FOR USING THE AUTOMATED INTEGRATIVE COMPLEXITY PROGRAM

Background

Automated Integrative Complexity (AutoIC) is a free, easy-to-use web-based program in which users upload their data directly to the website.

Once registered, users can analyze data in one of two ways, by document or by paragraph. AutoIC for Documents codes an entire document by breaking it up into 75-word sections. It is best for those who would like to analyze multiple files or large amounts of text. AutoIC for Paragraphs analyzes entire paragraphs without breaking them up. It should be used for shorter bits of text, like students' responses to short-answer questions. AutoIC for Paragraphs and for Documents differ in how they decide which words to score (their unit of scoring). See the AutoIC website for more information: http://www.autoic.org/instructions/.

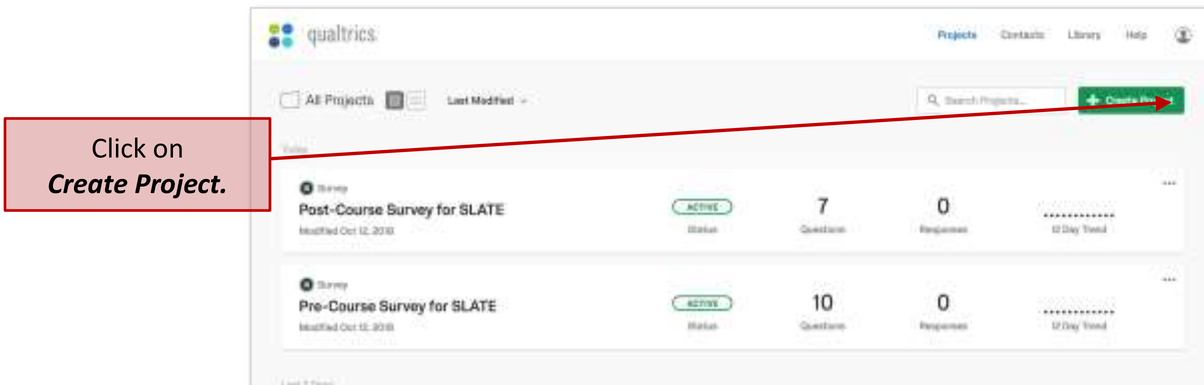
AutoIC allows users to analyze text on the *Integrative Complexity* construct and its subconstructs, *Dialectical Complexity* (i.e., complexity reflecting the recognition of different perspectives on a topic) and *Elaborative Complexity* (i.e., complexity in defense of a single viewpoint). Specifically, the program provides overall scores for *integrative complexity,* its *dialectical complexity* and *elaborative complexity* subconstructs, as well as *differentiation* and *integration* scores for each of them. For more information on scoring refer to Conway et al. (2014).

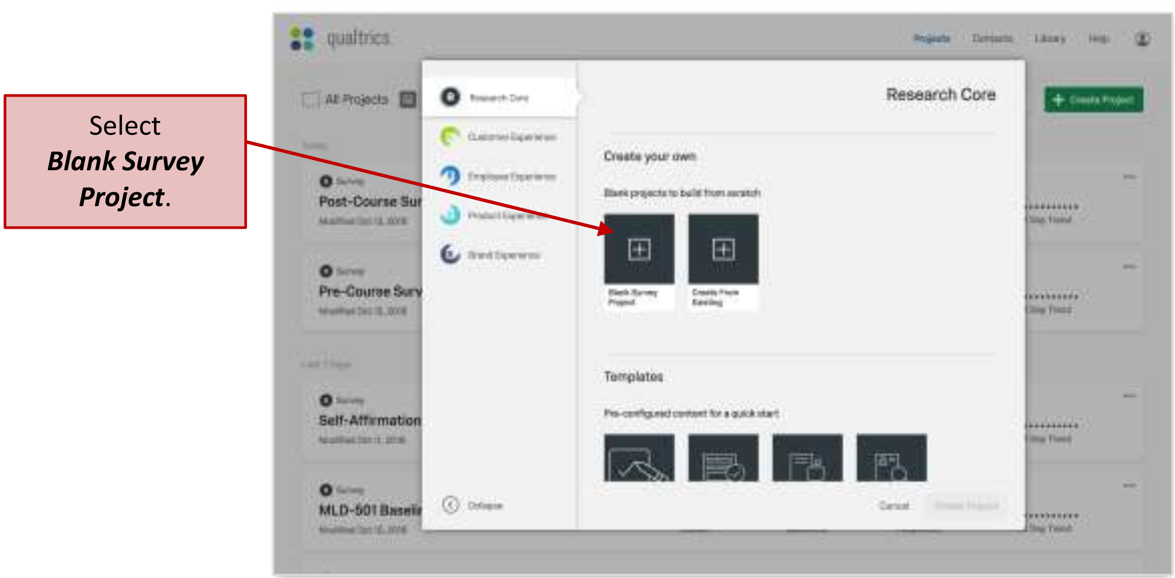
Below are instructions for using AutoIC for Paragraphs. For information on how to use AutoIC for Documents you may visit the AutoIC website (http://www.autoic.org/) or contact Erin Baumann, Associate Director of Professional Pedagogy at Harvard Kennedy School (erin_baumann@hks.harvard.edu) for assistance.

Creating Your Qualtrics Survey

If you would like to view the pre-course and post course surveys we used in 2017, click the following links: Pre-Course Survey, Post-Course Survey. If you would like to download the .qsf survey file for both surveys, click here.

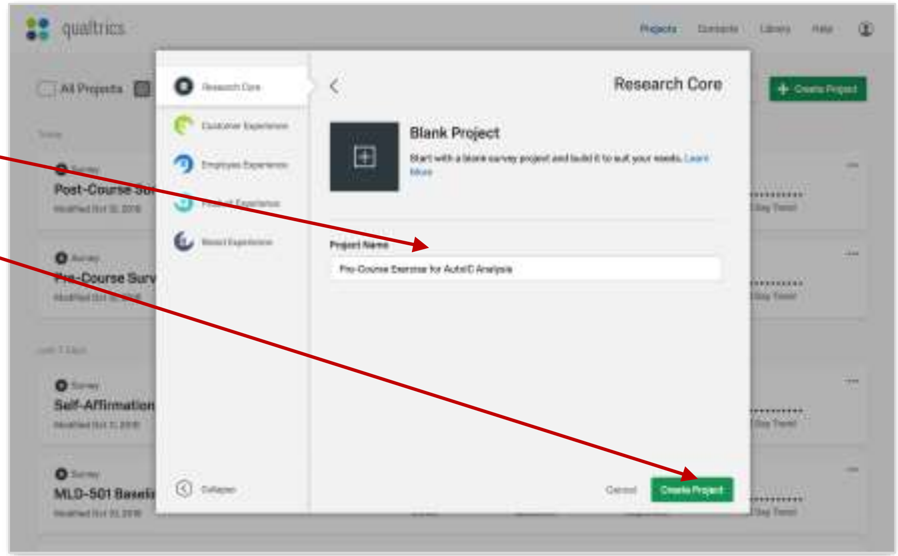
To create a new Qualtrics survey, follow the instructions below:





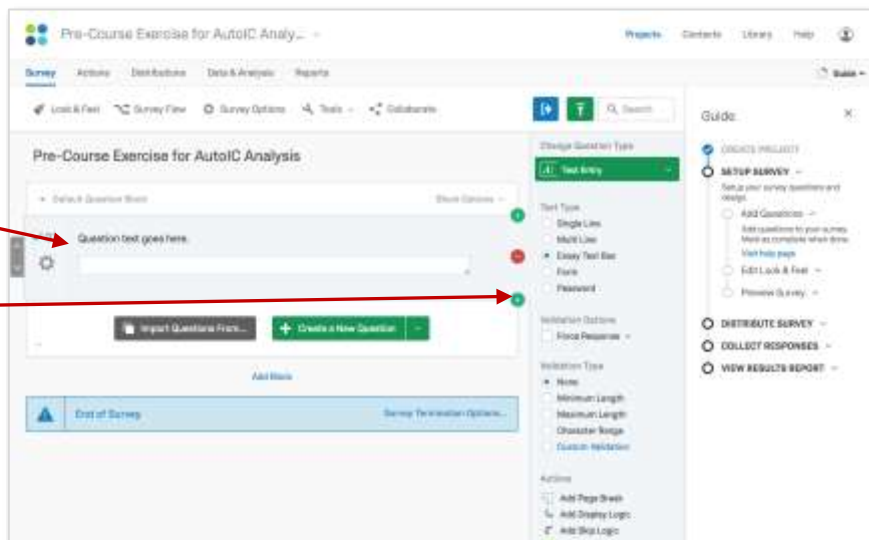
Give your project a name.

Then click ***Create Project***.

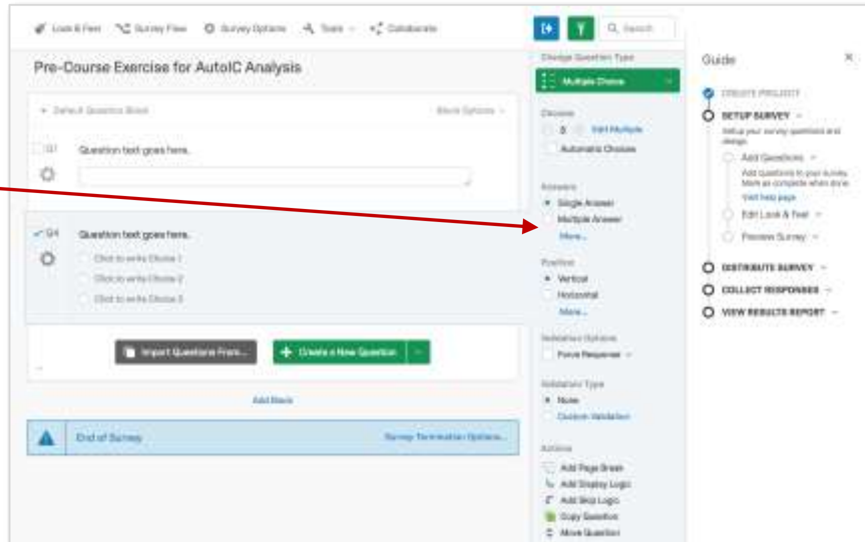


Insert new questions.

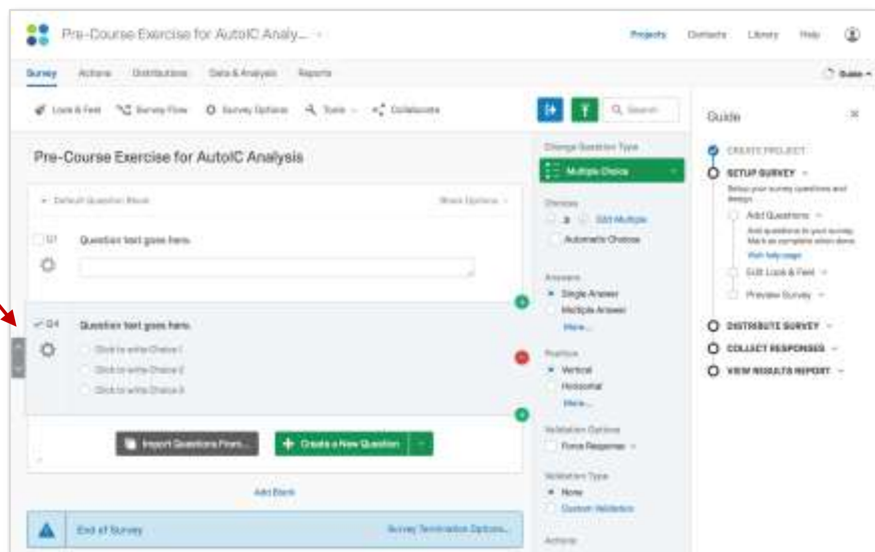
Click on the green plus button to easily add new questions.



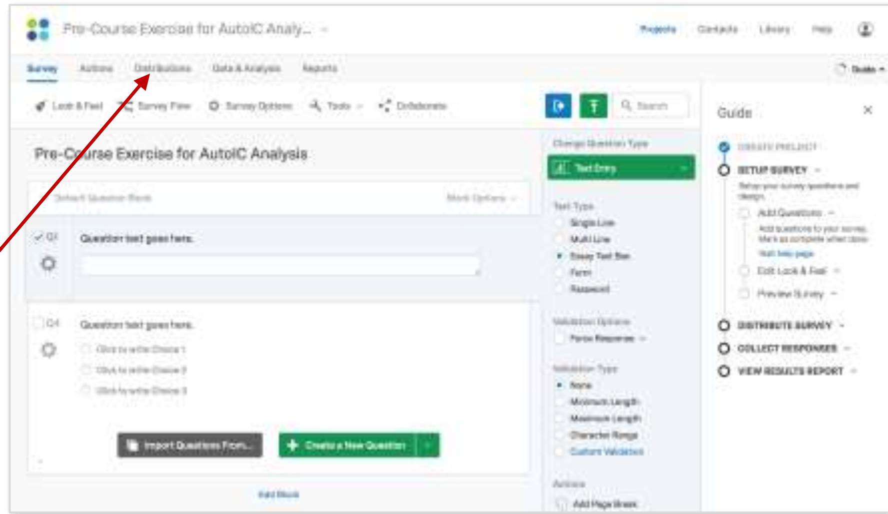
Use the sidebar to change question types, add page breaks, force responses, and set minimum length rules.



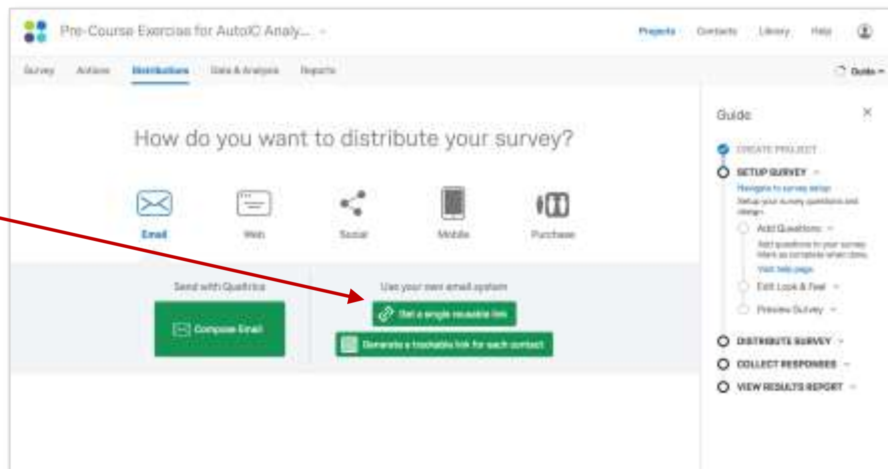
Use the arrow buttons next to each question to move questions around.



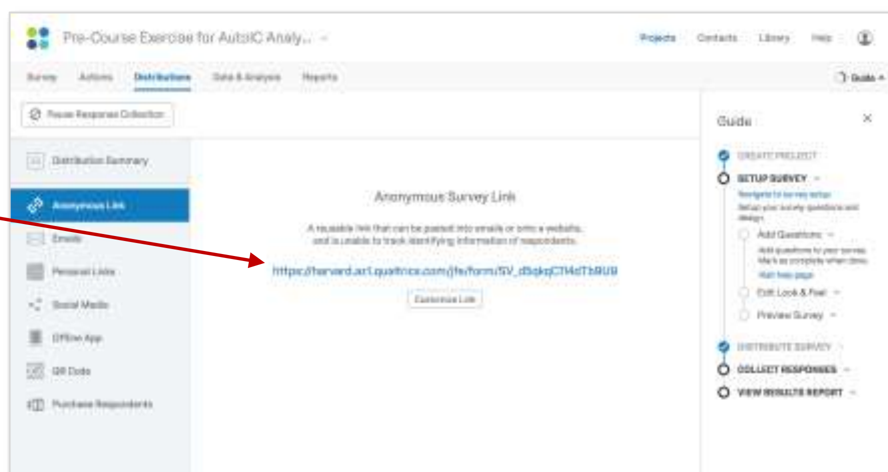
To get the survey link to distribute to students, click the ***distributions*** tab.



Click ***Get a single reusable link***.



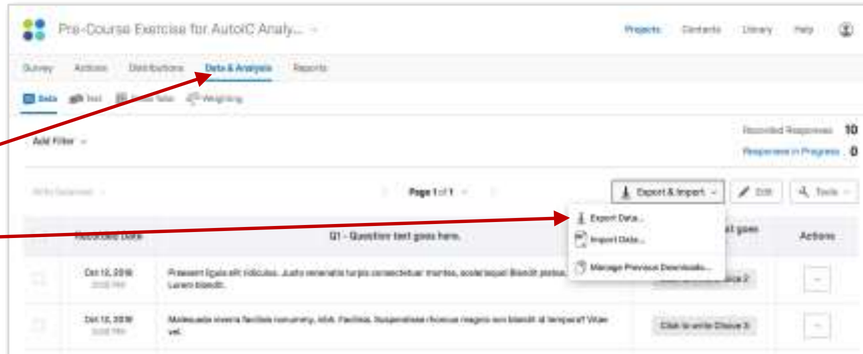
Copy the link.



Downloading the Raw Survey Data

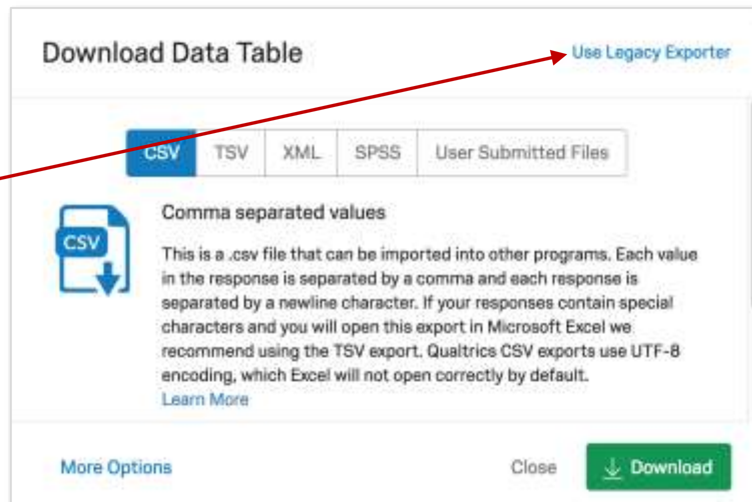
Navigate to the ***Data & Analysis*** tab.

Click the ***Export & Import*** dropdown box and select ***Export Data***.

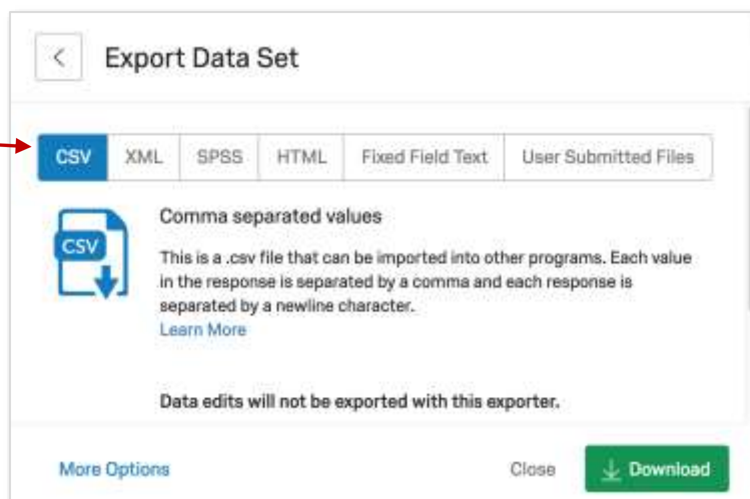


Use the Legacy Exporter option to prevent special characters from being inserted in the text (e.g., `Æ`and $).

Select ***Use Legacy Exporter***.



Download the data as a CSV file.



For more help on using Qualtrics, visit their website or call their help line at 1-800-340-9194.

Preparing Your Data for AutoIC Analysis

Students' responses that you are analyzing should be in a column labeled *Paragraph*. You may include three columns labeled *Var1*, *Var2*, and *Var3* containing extra data you would like in the output but not analyzed (e.g., participant numbers). Leave these columns blank if you do not need them.

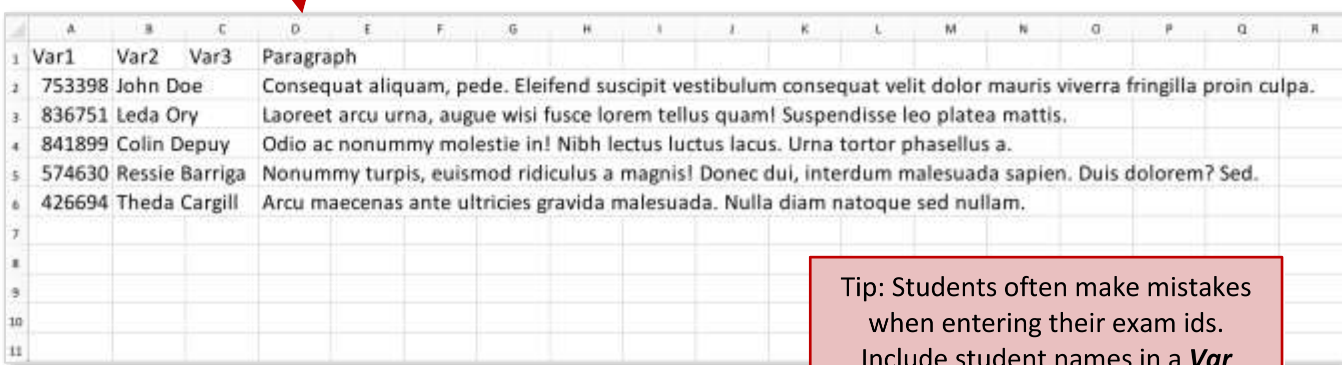
This is what a final data set of 5 responses would look like.

Var1 = exam ids
Var2 = names
Paragraph = essay responses

You must remove line breaks from text in the *Paragraph* column.

Use the following formula (in column E for the example below):
=SUBSTITUTE(D2,CHAR(13)," ").

Cell D2 contains students' responses from which you are removing line breaks.



Var1	Var2	Var3	Paragraph
753398	John Doe		Consequat aliquam, pede. Eleifend suscipit vestibulum consequat velit dolor mauris viverra fringilla proin culpa.
836751	Leda Ory		Laoreet arcu urna, augue wisi fusce lorem tellus quam! Suspendisse leo platea mattis.
841899	Colin Depuy		Odio ac nonummy molestie in! Nibh lectus luctus lacus. Urna tortor phasellus a.
574630	Ressie Barriga		Nonummy turpis, euismod ridiculus a magnis! Donec dui, interdum malesuada sapien. Duis dolorem? Sed.
426694	Theda Cargill		Arcu maecenas ante ultricies gravida malesuada. Nulla diam natoque sed nullam.

Tip: Students often make mistakes when entering their exam ids. Include student names in a ***Var*** column to have a backup identifier.

Save your final data set as a CSV file for the AutoIC analysis.

A Note for Pre/Post and Multi-Question Surveys

AutoIC can only code one column of text data in each analysis.

If you would like to analyze pre- and post-course responses and/or responses to multiple questions in the same data set, you will need to include <u>all text responses in one ***Paragraph*** column.</u>

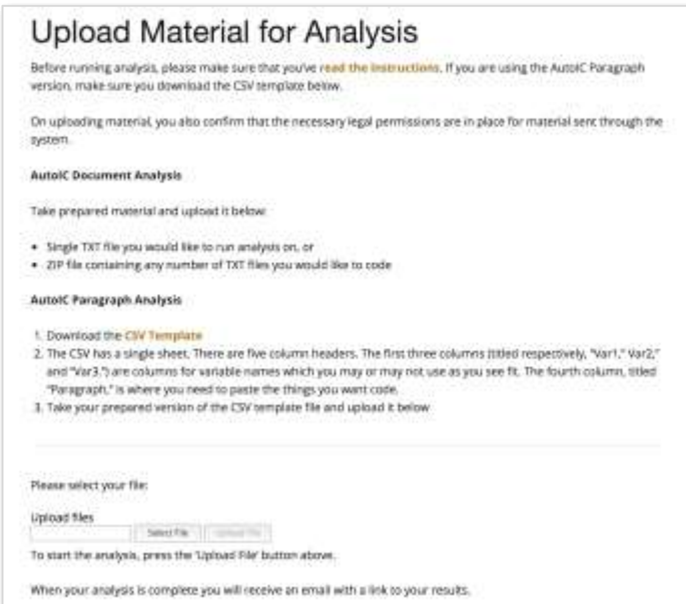
To do this, you must reshape the data to include time (pre or post) and question number in one of the ***Var*** columns.

Var3 (Column C) indicates the **question number** (q1) and **time** (pre or post)

	A	B	C	D	E	F	G	H	I	J
1	Var1	Var2	Var3	Paragraph						
2	753398	John Doe	q1_pre	Consequat aliquam, pede. Eleifend suscipit vestibulum consequat velit dolor mauris viverra						
3	836751	Leda Ory	q1_pre	Laoreet arcu urna, augue wisi fusce lorem tellus quam! Suspendisse leo platea mattis.						
4	841899	Colin Depuy	q1_pre	Odio ac nonummy molestie in! Nibh lectus luctus lacus. Urna tortor phasellus a.						
5	574630	Ressie Barrig	q1_pre	Nonummy turpis, euismod ridiculus a magnis! Donec dui, interdum malesuada sapien. Duis						
6	426694	Theda Cargil	q1_pre	Arcu maecenas ante ultricies gravida malesuada. Nulla diam natoque sed nullam.						
7	753398	John Doe	q1_post	Orci. Aliquet, ullamcorper odio posuere. Vestibulum lorem, sapien sed posuere montes euis						
8	836751	Leda Ory	q1_post	Ac porttitor et vulputate dictumst magna. Vulputate. Tempora cras neque! Pede risus. Nunc						
9	841899	Colin Depuy	q1_post	Porttitor diam. Molestie eleifend cursus neque lorem turpis, dignissim! Dapibus.						
10	574630	Ressie Barrig	q1_post	Culpa consectetuer vitae integer. Sagittis. Proin elementum at, ac augue maecenas tellus se						
11	426694	Theda Cargil	q1_post	Praesent augue vehicula turpis sodales rutrum ultricies sit tortor natoque.						
12										
13										
14										

Working with the AutoIC Website and Data

In order to utilize the AutoIC system you must first go to the system's website (http://www.autoic.org/) and create a free academic user account.



Upload Material for Analysis

Before running analysis, please make sure that you've read the instructions. If you are using the AutoIC Paragraph version, make sure you download the CSV template below.

On uploading material, you also confirm that the necessary legal permissions are in place for material sent through the system.

AutoIC Document Analysis

Take prepared material and upload it below:

- Single TXT file you would like to run analysis on, or
- ZIP file containing any number of TXT files you would like to code

AutoIC Paragraph Analysis

1. Download the CSV Template
2. The CSV has a single sheet. There are five column headers. The first three columns (titled respectively, "Var1," Var2," and "Var3.") are columns for variable names which you may or may not use as you see fit. The fourth column, titled "Paragraph," is where you need to paste the things you want code.
3. Take your prepared version of the CSV template file and upload it below

Please select your file:

Upload files

To start the analysis, press the 'Upload File' button above.

When your analysis is complete you will receive an email with a link to your results.

Once logged in, navigate to the ***Academic User*** tab and click ***Upload***.

Follow the instructions to upload your CSV file.

AutoIC results will be emailed to you as an excel file.

AutoIC provides nine scores: *integrative complexity, dialectical complexity,* and *elaborative complexity* (subconstructs of integrative complexity*)* scores, and *differentiation* and *integration* scores for each of them.

The overall integrative complexity score is most informative.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Var1	Var2	Var3	Paragraph	Words	IC	DIAL	ELAB	IC_Differentiation	IC_Integration	DIAL_Differentiation	DIAL_Integration	ELAB_Differentiation	ELAB_Integration	
2	753398	John Doe		Consequat a		7	6.5	3.5	2	4	2	3.5	2	0.5	
3	836751	Leda Ory		Laoreet arcu		3.5	3.5	3	2	0.5	2	0.5	2	0	
4	841899	Colin Depuy		Odio ac non		4	3.5	3	2	1	1.5	1	2	0	
5	574630	Ressie Barriga		Nonummy ti		2.75	2.75	2.75	1.25	0.5	1.5	0.25	1.75	0.25	
6	426694	Theda Cargill		Arcu maecei		6	6	3	2	3	2	3	2	0	
7															
8															

RESOURCES

Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. "Automated Integrative Complexity." *Political Psychology* 35 (2014): 603–624. Permalink

Houck, S. C., Conway, L. G., & Gornick, L. J. "Automated Integrative Complexity: Current Challenges and Future Directions." *Political Psychology* 35, no. 5 (2014): 647-659. Permalink

Light, R. "Assessment Strategy for Enhancing Sustained HKS Excellence." SLATE HKS Faculty Working Paper Series on Teaching & Learning. Cambridge, MA: Harvard Kennedy School, 2017. Permalink

Tetlock, P. E., Metz, S. E., Scott, S. E., & Suedfeld, P. "Integrative Complexity Coding Raises Integratively Complex Issues." *Political Psychology* 35*,* no. 5 (2014): 625-634. Permalink

For additional assistance you may contact Erin Baumann, Associate Director of Professional Pedagogy at Harvard Kennedy School (erin_baumann@hks.harvard.edu).

REFERENCES

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. *The Authoritarian Personality* (1st ed.). New York, NY: Harper, 1950.

Antonio, A. L., Chang, M. J., Hakuta, K., Kenny, D. A., Levin, S., & Milem, J. F. "Effects of Racial Diversity on Complex Thinking in College Students." *Psychological Science* 15 (2004): 507–510.

Conway, L. G., Suedfeld, P., & Tetlock, P. E. "Integrative Complexity and Political Decisions That Lead to War or Peace." In *Peace, Conflict, and Violence: Peace Psychology for the 21st Century,* D. J. Christie, R. V. Wagner, & D. A. Winter (Eds.), 66–75. Englewood Cliffs, NJ: Prentice-Hall, 2001.

Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. "Automated Integrative Complexity." *Political Psychology* 35, (2014): 603–624.

Gruenfeld, D. H. "Status, Ideology, and Integrative Complexity on the U.S. Supreme Court: Rethinking the Politics of Political Decision Making." *Journal of Personality and Social Psychology* 68 (1995): 5–20.

Gruenfeld, D. H., Thomas-Hunt, M. C., & Kim, P. H. "Cognitive Flexibility, Communication Strategy, and Integrative Complexity in Groups: Public versus Private Reactions to Majority and Minority Status." *Journal of Experimental Social Psychology* 34 (1998): 202–226.

Kelly, G. *A Theory of Personality: The Psychology of Personal Constructs* ([1st ed.]). New York, NY: Norton, 1955.

Nemeth, C. J. "Minority Dissent as a Stimulant to Group Performance." In *Group Process and Productivity,* S. Worche, W. Wood, & J. A. Simpson (Eds.), Vol. 95, 95–111. Newbury Park, CA: Sage, 1992.

Rokeach, M. *The Open and Closed Mind*. New York, NY: Basic Books, 1960.

Suedfeld, P., & Rank, A. D. "Revolutionary Leaders: Long-term Success as a Function of Changes in Conceptual Complexity." *Journal of Personality and Social Psychology* 34 (1976): 169–178.

Suedfeld, P., Tetlock, P. E., & Streufert, S. "Conceptual/Integrative Complexity." In *Motivation and Personality: Handbook of Thematic Content Analysis,* C. P. Smith (Ed.), 393–400. Cambridge, England: Cambridge University Press, 1992.

Tetlock, P. E. "Integrative Complexity of American and Soviet Foreign Policy Rhetoric: A Time-Series Analysis." *Journal of Personality and Social Psychology* 49 (1985): 1565–1585.

Tetlock, P. E. "Monitoring the Integrative Complexity of American and Soviet Policy Rhetoric: What Can Be Learned?" *Journal of Social Issues* 44 (1988): 101–131.

Tetlock, P. E., Metz, S. E., Scott, S. E., & Suedfeld, P. "Integrative Complexity Coding Raises Integratively Complex Issues: Complexities of Integrative Complexity Coding." *Political Psychology* 35 (2014): 625–634.