5-2022

Script Concordance Testing in Genetic Counseling Training: A Pilot Study

Yakira Begun
Sarah Lawrence College

Lila Rae Stange
Sarah Lawrence College

SCRIPT CONCORDANCE TESTING IN GENETIC COUNSELING TRAINING: A PILOT STUDY

Yakira S. Begun and Lila R. Stange

May 2022

Submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Human Genetics
Sarah Lawrence College

Abstract

Clinical decision-making has been successfully measured by script concordance testing in various healthcare training programs; it has never been used in genetic counseling education. The aim of this pilot study was to assess script concordance testing in the field of genetic counseling as an objective measure of clinical reasoning in trainees. The script concordance test was administered to 22 second year genetic counseling students in the Joan H. Marks Graduate Program in Human Genetics at Sarah Lawrence College. 12 genetic counselors served on a panel to provide expert judgment responses, and a scoring grid was developed using the Aggregate Scores Method. The utility of the tool was measured using Cronbach's alpha coefficient, and scores of students and the panel were compared using Hedge's *g*. Results revealed statistically significant differences between the scores of panelists and students and good reliability. Concordance and discordance in clinical reasoning was compared across case categories where anchored cases demonstrated the highest overall degree of discordance and ethics cases demonstrated the highest overall degree of concordance. Script concordance testing can be used to measure clinical decision-making skills in genetic counseling trainees in a way that is reliable, standardized, and easy to use.

Keywords: decision-making, education, genetic counseling, script concordance test, uncertainty

ACKNOWLEDGMENTS

We extend a heartfelt thank you to our thesis mentor and advisor Claire Davis, co-program director of the Human Genetics program. Her vision and passion for the project has served as a constant source of inspiration for us and her insight is unparalleled.

Script Concordance Testing in Genetic Counseling Training: A Pilot Study

The Accreditation Council for Genetic Counseling (ACGC) developed the Practice Based Competencies (2019) to define measurable standards and skills that genetic counseling students must be trained in before they graduate (Doyle et al, 2016; Fine et al, 1996). One such component of several of these competencies is clinical decision-making, which is defined by Tiffen et al. (2014) as "a contextual, continuous and evolving process where data is gathered, interpreted and evaluated in order to select an evidence-based choice of action" (Tiffen et al., 2014, p. 401). Clinical decision-making skills have traditionally been a difficult area to measure as there is often not a single right answer in clinical situations where the outcome is uncertain (Ilgen et al., 2019). No such defined measure to evaluate clinical decision-making currently exists in the curriculum at the Joan H. Marks Graduate Program in Human Genetics at Sarah Lawrence College, which will be referred to from this point on as the SLC Human Genetics Program.

Script Concordance Testing (SCT) is a method specifically designed to assess competency in clinical decision-making under conditions of uncertainty by comparing participants' responses against those of a panel of experts (Charlin et al., 1998; Charlin et al., 2000). Graduate training programs across multiple healthcare fields such as nursing (Dawson et al., 2014), radiation-oncology (Lambert et al., 2009), emergency medicine (Steinberg et al., 2020), pharmacology (Kaur et al., 2020), and neurology (Lubarsky et al., 2009) have utilized SCT as a standardized way to measure students' clinical decision-making skills; however, there is no literature on using SCT in genetic counseling training programs.

Therefore, we are interested in studying the use of the script concordance test as a tool to evaluate clinical decision-making skills amongst second year students in the SLC Human

Genetics Program. We aim to uncover the extent to which SCT accurately measures clinical reasoning in the field of genetic counseling, where uncertainty is present in every patient interaction.

Background

As healthcare providers who interact with patients at various points throughout their personal and family health histories, genetic counselors often work under conditions of uncertainty (McCarthy Veach et al., 2018). The uncertainty involved in the patient-provider interaction can come in many forms including dealing with uncertain or conflicting interpretations of health histories or genetic test results, and clinical judgment of multiple factors used to estimate risks and guide testing options. A primary goal of genetic counseling is to prepare the patient to make an informed decision based on the options and information available to them (McCarthy Veach et al., 2018). For these reasons, competence in clinical decision-making under conditions of uncertainty is an important component of genetic counseling training. In the literature, no assessment tool to evaluate genetic counseling students' ability to engage in this type of thinking currently exists.

SCT as a method of assessment for clinical decision-making skills was first proposed by Charlin et al. (2000). The function of the tool in evaluating judgment as a critical component of successful healthcare professional practice is described alongside a detailed guide on item-writing, test construction, expert panel composition, scoring and administration of a reliable and valid script concordance test. Lubarsky et al. (2013) provided an updated, comprehensive guide on test construction and utilization of SCT that can be applied to any clinical setting.

The script concordance test is an objective, unbiased tool used to score students' ability to engage in the clinical decision-making process (Charlin et al., 2000). Script concordance test

cases are designed to represent common clinical encounters in the field (Fournier et al., 2008). Script concordance test questions are constructed in such a way that there is no correct answer, instead reflecting several options that may be reasonable in real-life, ambiguous clinical situations (Charlin et al., 2000; Fournier et al., 2008). This also allows the tool to be used in test-retest situations. Wan et al. (2020) compared medical students' responses on a script concordance test with a written justification of the reasoning they applied (Think-Aloud approach). Script concordance test responses and the written explanations were compared with reasoning of an expert reference panel to determine concordance. The majority of students who selected responses that matched the responses of experts also provided written explanations that matched the thought processes of the experts. The findings of this study provide support for the response process validity of script concordance test scores. This highlights the difference between a typical multiple-choice test, where the answer is correct or incorrect, and the script concordance test, where varied clinical reasoning does not represent an incorrect response as long as it can be justified based on sound knowledge.

SCT has been used in training programs across a variety of healthcare professions. Dawson et al. (2014) studied the validity and reliability of SCT in nursing students. This study found SCT to be a reliable method of evaluating clinical reasoning in nursing students, which is consistent with findings from previous studies in this area and supports the broad use of SCT as a valid and easy-to-use method of student evaluation. Lambert et al. (2009) assessed the reliability and validity of SCT in radiation oncology students, with special emphasis on how the test was perceived by participants. Reactions to the test were collected following completion. Participants reported that they completed the test with ease and felt that the cases accurately represented real clinical scenarios. This study found that as experience in the field of radiation oncology

increased, scores on the script concordance test also increased, with the authors concluding that students and residents who received lower scores than average may not have sound clinical judgment and may require remedial action. Results from training programs that have utilized SCT support its use as a standardized method for student evaluation.

A key theme that arose through review of the existing literature is the ease with which the script concordance test can be administered and completed by participants. Participants generally felt that the script concordance test was clinically applicable and mimicked the types of decisions they would be required to make in a patient-facing clinical setting (Lambert et al., 2009).

Across studies that have utilized SCT, panel members' scores have differed significantly from students' scores (Dawson et al, 2014; Lambert et al, 2009). The results of these studies indicate that clinical decision-making skills develop as a person gains more patient-facing clinical experiences. While there is a possible inflection point where clinical decision-making skills are gained through years of working experience, there remains a standard of skill that students should possess before graduation from a training program, which is defined as being no more than two standard deviations (2 SD) below the mean participants scores (Lubarsky et al., 2013).

Unlike some of the other medical specialties where SCT has been used, genetic counseling may be unique in that, within the clinical decision-making process, emphasis is distributed across scientific and psychosocial considerations as consistent with patient-centered care (McCarthy Veach et al., 2018). When creating a script concordance test for genetic counseling trainees, care must be taken to account for these considerations. SCT was developed based on script theory (Charlin et al., 2000) and encourages participants to incorporate additional information into a pre-existing knowledge script. On the scientific side, diagnostic cases that

draw heavily on medical knowledge are often utilized for medical professionals to practice clinical judgment. With respect to psychosocial awareness, in a study by Pau et al. (2019), SCT was used instead to evaluate ethical reasoning ability in medical students. Literature has reiterated the importance of clinicians having some degree of ethical reasoning ability in order to resolve medical ethical dilemmas (Pau et al., 2019). This study found that the script concordance test was able to differentiate between students' and panel members' abilities to engage in the process of ethical decision-making and proves that SCT can be used to measure a variety of concepts within the decision-making framework.

Methods

This study explored the utility of SCT as an approach to assessing clinical decision-making skills in genetic counseling trainees in the SLC Human Genetics Program. Because there is not literature on the use of SCT in genetic counseling training, this study piloted all steps of the SCT method: writing cases and questions, utilizing an expert panel, validating test construction, scoring each question and calculating a total test score, and interpreting scores with participants. This study also gathered students' experiences of taking a script concordance test and their impressions of using this tool as an assessment of or supplement to their learning. This study was submitted to the Sarah Lawrence College IRB and was determined to be exempt.

Writing Cases and Questions

A script concordance test for genetic counseling trainees was constructed by two second year genetic counseling students in the SLC Human Genetics Program in conjunction with two certified genetic counselors who also serve as program faculty. Cases were constructed based on common clinical encounters in the three main subspecialties of genetic counseling: prenatal genetics, cancer genetics, and pediatric/adult genetics. The authors' clinical experiences and

example scenarios from McCarthy Veach et al. (2018) and Nussbaum et al. (2016) were used to write the cases and questions. All identifying details have been changed to protect patient identity. Every case incorporated a level of uncertainty as is common and realistic for clinical scenarios in professional practice. To create an assessment tool which is representative of various clinical encounters within genetic counseling, three case categories were created: Diagnostic (Figure 1a), Anchored (Figure 1b), and Ethics (Figure 1c).

Each case begins with a vignette where the participant is introduced to the initial scenario. The vignette is followed by a series of independent questions where the participant is asked to evaluate the effect, if any, that the additional information provided in the question has on their initial judgments or interpretations. By establishing an initial hypothesis based on the vignette and then incorporating additional information and adjusting responses accordingly, participants are prompted to engage in the process of clinical decision-making (Charlin et al., 2000).

In the Diagnostic case category, the initial vignette presented participants with features suspicious of a certain genetic condition. Every vignette was followed by three questions, with each question providing an independent, additional piece of information to incorporate into the vignette. Participants were then asked to evaluate the impact of the new piece of information using a 5-point Likert scale (Figure 1a). This style of item writing is similar to the recommendations proposed by authors of the original SCT method (Charlin et al., 2000; Lubarsky et al., 2013).

Anchored and Ethics cases were created based on the structure of script concordance test questions developed by Tsai (2017) and Pau et al. (2019), which differs slightly from that of the Diagnostic cases. As shown in Figures 1b and 1c respectively, Anchored and Ethics cases

present participants with an initial vignette and in Part A, ask participants to make a clinical or ethical baseline judgment. In Anchored cases, participants are provided with pieces of new and independent information in subsequent questions and are asked to evaluate the effect, if any, of the new information on their initial response from Part A using a 5-point Likert scale. Ethics cases were constructed following the recommendations outlined by Pau et al. (2019), where the baseline judgment in Part A was considered the *decision question* and consisted of a *Yes* or *No* answer to capture the student's position on the ethical dilemma presented in the vignette. All subsequent questions provided an opportunity for participants to demonstrate the extent to which they would change their initial position, if at all, upon the incorporation of a new piece of information using a 3-point Likert scale.

Figure 1

Examples of concordant cases on the script concordance test from each of the three case types

a) Diagnostic Type: Case 10

A 51 year old woman is returning to clinic to discuss the results of her hereditary cancer panel. She was previously diagnosed with parathyroid cancer at age 50, which was confirmed on pathology to be an adenoma. She is currently undergoing chemotherapy. Genetic testing identified a VUS in MEN1, which is associated with Multiple Endocrine Neoplasia Type 1 (MEN1) syndrome.		
Consider the following recommendation as a next step for this patient:	And then you find out the following new information:	Then this recommendation becomes:
A clinical evaluation for Multiple Endocrine Neoplasia Type 1 (MEN1)	The patient has a family history significant for multiple relatives with MEN1 related cancers.	-2 = contraindicated -1 = less indicated 0 = neither more nor less indicated than before +1 = more indicated +2 = strongly indicated

b) Anchored Type: Case 20

A 37 year old female, G1P0, is returning to discuss the results of her amniocentesis. Karyotype results found 46,XX; however, the fetal ultrasound scan done at 18 weeks GA revealed a normal male fetus. Follow up testing on amniocentesis with FISH found a SRY signal on one X chromosome, confirming a diagnosis of 46,XX testicular disorder of sexual development (DSD).		
Part A: How likely are you to explore the psychosocial concerns and stigmas related to DSDs?	And then you find out the following new information:	Then this discussion becomes:
-2 = very unlikely -1 = not likely 0 = neither likely nor unlikely +1 = likely +2 = very likely	A close friend of the family who also had a DSD recently took his own life after facing years of bullying and discrimination in both his personal and professional life.	-2 = very unlikely -1 = less likely 0 = neither more nor less likely than before +1 = more likely +2 = very likely

c) Ethics Type: Case 3

Your patient is an unaffected 16 year old male with a family history of Huntington's Disease (HD) who would like to pursue his own genetic testing. His father had HD and recently died. His comes to clinic with his mother.		
Part A: Would you order genetic testing for Huntington Disease for this patient today?	And then you find out the following new information:	Then your decision from Part A would:
Yes No	Upon further review of the family history that there are some paternal relatives that showed signs of Huntington's disease in their teens.	-2 = change completely 0 = not change +2 = be strengthened

The initial draft of the script concordance test for genetic counseling trainees included 10 Diagnostic cases, 13 Anchored cases, and nine Ethics cases (32 total cases). This draft was submitted to two certified genetic counselors who serve as program faculty for editing and review prior to sending the tool to expert panel members for completion (Lubarsky et al., 2014). The test was constructed in Google Forms for ease of construction, distribution, and cost-effectiveness.

Utilizing an Expert Panel

As per recommendations by Charlin et al. (2000), the expert panel must consist of professionals with clinical experience in the field that the assessment tool is created for. Panel members are asked to complete the script concordance test independently, then their responses are used to create a scoring grid. In this pilot, 12 genetic counselors – all of whom are connected to the SLC Human Genetics program – formed the expert panel. The panel was representative of the demographic profile of the genetic counseling profession. Panelists were chosen deliberately because their clinical reasoning reflects professional perspectives taught to students throughout the training program (Lambert et al., 2009; Petrucci et al., 2013). 11 of the 12 panelists completed the full script concordance test and one panelist completed cases in one of the three case categories.

Validating Test Construction

A minimum of 10 panelists is recommended for construction of a reliable and valid script concordance test (Fournier et al., 2008; Gagnon et al., 2005). In this pilot study, 10 genetic counselors completed the initial 32-case draft of the test and were encouraged to provide brief rationales of their reasoning to be compiled and shared with the students as example judgments to compare with the students' own thinking. As recommended by Lubarksy et al. (2013) as a step

in the validation process, 10 cases were excluded either because they garnered a large spread of responses from panelists or they appeared to be more focused on knowledge instead of clinical reasoning based on review of panelists' responses and feedback. The revised test included 22 cases. Other steps of test validation included revising and redistributing one case to panelists for new responses (two panelists did not complete this revised case and their answers were counted as missing data for scoring purposes) and asking one panelist to take the assessment in one sitting to estimate the time needed to complete the 22-case test (approximately one hour).

Scoring Participants' Responses

The score assigned to a student's response to each question was calculated based on the proportion of expert panelists who selected that response. The modal response – the response selected by the most panelists – was credited with one point, whereas all other responses selected by any panelist received partial credit based on the proportion of panelists who selected that response (Charlin et al., 2000; Fournier et al., 2008). Responses that were not chosen by any panelists received zero credit or no points. This method was used to score the questions in the Diagnostic case category.

Part A of the Anchored and Ethics cases was scored using the same method as the Diagnostic cases, where students were scored based on their selection of a modal response, partial credit response or no credit response. If a no credit response was selected in Part A, no credit was received on all subsequent questions (Tsai, 2017). This scoring method allows for six possible response combinations for each subsequent question in an Ethics case and 25 possible response combinations for each subsequent question in an Anchored case.

A total score for the script concordance test was calculated by dividing the sum of scores for all questions by the total number of questions and multiplying by 100. There were occasional

questions where a panelist's response was excluded from the scoring grid upon review of the panelist's suspected misrepresentation of information (Gagnon et al., 2011). Panelist scores were calculated using the same scoring grid for continuity.

To measure the clinical reasoning of genetic counseling trainees, responses from the script concordance test were evaluated for reliability and validity. All data was anonymized prior to analysis to reduce bias and protect confidentiality. Internal consistency of the cases and questions on the test was measured by Cronbach's alpha. Hedges' *g* was calculated to measure effect size for each case type and the total test. The Kolmogorov-Smirnov test was used to evaluate normality of distribution of total test scores.

Concordance in clinical reasoning was measured using the standardized script concordance test scoring method, where the score awarded for a student's response represents the degree of concordance with the panel members' responses (Charlin et al., 2000). Concordant and discordant questions were determined by the amount of no credit responses arising from comparison of panel members and student responses (Figure 2). Concordant questions had zero or few participants who selected a no credit response; discordant questions had a high number of participants who selected a no credit response. Across the three case types, the Anchored cases showed the highest degree of discordance while the Ethics cases showed the highest degree of concordance.

Interpreting Scores with Participants

Participants included current second year students in the SLC Human Genetics Program. The authors themselves were excluded from the study population.

Students were provided with overall scores for the test, scores for each question, and a companion document with detailed explanations of panelists' responses that received credit.

Students were then able to compare the reasoning behind their own clinical decision-making to that of panelists. While a study by Lubarsky et al. (2013) emphasizes that there is no single correct answer on script concordance test questions, the document enabled students to assess where their thinking may have differed from the panelists and why a certain response may have received no credit. Emphasis was placed on the clinical reasoning underlying the selected response as opposed to the score itself.

Students participated in a group session led by the authors before and after completing the test. The goal of the initial session was to review instructions for taking the test and reasoning behind SCT as a method of assessment (Fournier et al., 2008). The goals of the follow-up session were to assist students in interpreting their scores, explain how to compare panelists' reasoning with their own judgments, and engage students in a facilitated conversation about challenging cases included on the test.

Results

Among the eligible participants, 22 out of 24 students voluntarily agreed to participate. While there was no set time constraint, most participants completed the script concordance test in approximately one hour. For the total script concordance test and across all case types, the students' mean score (66%) was found to be lower than the panelists' mean score (78%), and the standard deviation was larger among students (7%) compared to panelists (5%) (Table 1).

Cronbach's alpha ranged from 0.64 to 0.81 depending on case type (Table 1). Cronbach's alpha for the full 22-case test was 0.81, which meets the threshold for reliability and indicates good internal consistency of included items.

Hedges' *g* was specifically selected to account for the different sample size among panel members and students (Table 1). The results are considered statistically significant as shown by

the large effect size for the total test (g=1.88). This indicates that the test is able to differentiate between panelists and students.

As seen in Table 2, results of the Kolmogorov-Smirnov test confirm that the distribution of scores was normally distributed across panelists and students.

Table 1

Descriptive analyses of the distribution of script concordance test scores for panel members and students

Results	Number of Participants (N)	Mean Score (%)	Standard Deviation (SD)	Cronbach's alpha	Hedges' g
Full script concordance test				0.81	1.88
Panel	12	78%	5%		
Students	22	66%	7%		
Diagnostic Type				0.64	0.43
Panel	11	74%	11%		
Students	22	68%	15%		
Anchored Type				0.7	1.91
Panel	11	74%	6%		
Students	22	57%	10%		
Ethics Type				0.75	0.70
Panel	12	81%	8%		
Students	22	74%	11%		

Table 2

Kolmogorov- Smirnov test for normality

Total Score	Statistic	df	p
Panel	0.26	12	0.375
Students	0.08	22	0.281

Amongst the Diagnostic cases, the highest degree of concordance was seen in Case 10 Question 1, where participants were asked about the utility of a clinical evaluation for MEN1 given the history presented in the case (Figure 1a). On this question, all 22 students choose either the modal response or a partial credit response (Figure 2a). The lack of no credit responses demonstrates high concordance in clinical reasoning in this content area. The highest degree of discordance in this case category was observed in Case 1 Question 2, where students were asked to consider the likelihood of non-accidental trauma based on the case indication and new findings on a physical examination (Figure 3a). While the majority of students selected the

modal response or a partial credit response, six students selected a no credit response (Figure 2b) indicating an area where further training and discussion may be beneficial.

Amongst the Anchored cases, the highest degree of concordance was seen on Case 20 Question 2, where participants were asked to determine the likelihood of exploring psychosocial concerns with a patient who has a negative personal experience with a condition (Figure 1b). In Case 20 Part A, all 22 students choose either a modal response or a partial credit response (Figure 2c). In Case 20 Question 2, after the introduction of a new piece of information, 21 students chose response combinations that received credit and only one student chose a no credit response combination (Figure 2c). This demonstrates that the new information introduced in Question 2 had only a slight effect on the initial clinical decision in Part A for most students and panelists. The highest degree of discordance was seen in Case 18 Question 2, where participants were asked to evaluate the utility of single site testing for a familial variant for a patient who was unaware that they were conceived via donor egg (Figure 3b). Case 18 Part A was only slightly discordant with 21 students selecting a response that received credit and one student selecting a no credit response; however, when the additional piece of information was introduced in Question 2, a high degree of discordance was seen with 11 students selecting a response combination that received credit and 11 students selecting a no credit response combination (Figure 2d). This demonstrates the discrepancy in clinical reasoning between panel members and students when evaluating the impact of this new piece of information on the initial case.

Amongst the Ethics cases, there were many highly concordant cases, one of which was Case 3 Question 2, where participants were asked about testing an unaffected minor for Huntington's disease (Figure 1c). In Case 3 Part A, all 22 students selected the modal response and in Case 3 Question 2, all 22 students selected a response combination that received credit

(Figure 2e). This shows significant concordance between students and panelists in this content area and significant concordance among students themselves. One of the most discordant Ethics cases is Case 14 Question 2, where students were asked about looking at the genetic test results of an estranged family member (Figure 3c). In part A of the case, all 22 students selected a response that received credit, but upon the addition of new information in Question 2, 17 students selected a response combination that received credit and five students selected a no credit response combination (Figure 2f).

Figure 2

Descriptive statistics graphs of concordant and discordant case examples

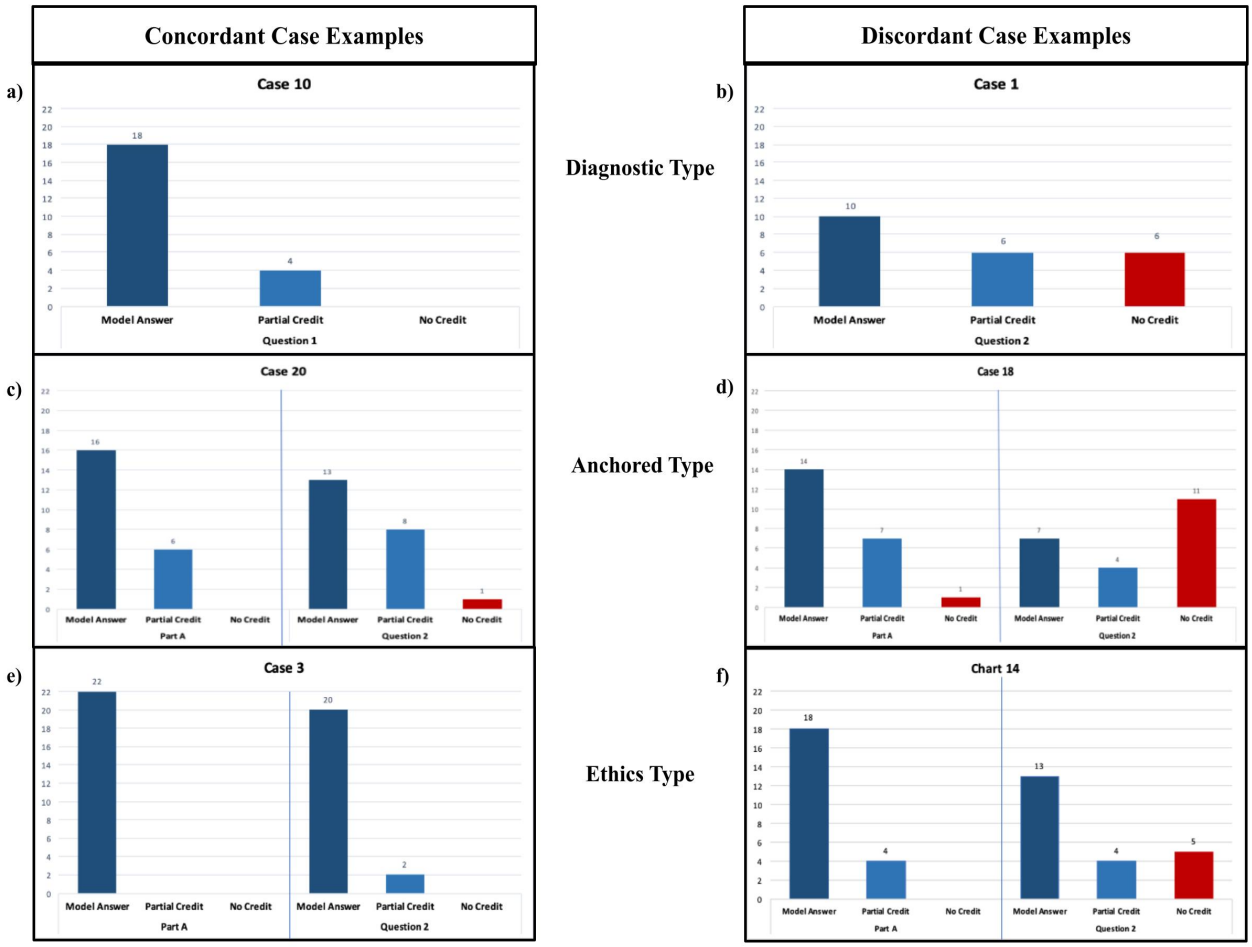

Figure 3

Examples of discordant cases from each of the three types of cases on the script concordance test

a) Diagnostic Type: Case 1

A 10 year old girl is referred to your pediatric clinic with a history significant for multiple skeletal fractures. Patient is accompanied to the appointment by her maternal aunt, who is currently her legal guardian and is in the process of seeking full custody.		
Considering the following differential diagnosis:	And then you find out the following new information:	Then this differential diagnosis becomes:
Abuse/non-accidental trauma	The patient has rhizomelia and extremely short stature on physical exam.	-2 = very unlikely -1 = less likely 0 = neither more nor less likely than before +1 = more likely +2 = very likely

b) Anchored Type: Case 18

A 27 year old female self-refers to genetic counseling after she stumbles upon a test report in one of the locked files in her home revealing that her mother has a pathogenic variant in CHEK2. She presents to clinic to understand her cancer risks and to pursue her own testing. There is no other family history of cancer.		
Part A: Offering single site testing for the familial variant is initially:	And then you find out the following new information:	Then the testing strategy from Part A becomes:
-2 = contraindicated -1 = not indicated 0 = neither more nor less indicated +1 = more indicated +2 = strongly indicated	You are aware that your patient was conceived via donor egg because her mother was a past patient of yours, but she herself is not aware that she and her mother are not biologically related. You know that this means she is not at risk for the familial variant but she is adamant about being tested today.	-2 = contraindicated -1 = less indicated 0 = neither more nor less indicated than before +1 = more indicated +2 = strongly indicated

c) Ethics Type: Case 14

A G0P0 patient presents to clinic for a preconception counseling session with her partner. When collecting the family history, she reports multiple family members with history of severe pediatric onset illness. When you ask the patient more specific details about these illnesses, she insists that although she does not know details, they were all treated in the same institution where you work and you should have access to their medical records. She is willing and ready to give you the names and date of birth of her relatives who have been treated at your rather small institution. She asks you specifically about her chances to have an affected child given her family history.		
Part A: Would you look up the medical records of your patient's affected family members to be able to give the patient a more informed estimate of her risk for an affected offspring?	And then you find out the following new information:	Then your decision from Part A would:
Yes No	Your patient is estranged from her parents and siblings. She knows that they receive care at the same institution because of their posts that she has seen on social media. She reports having no other way to get access to this information.	-2 = change completely 0 = not change +2 = be strengthened

Following the submission of the assessment, both students and panelists shared that cases were representative of common clinical encounters. Students enjoyed participating in the study and found that the script concordance test could serve as both a learning and assessment tool. Panelists found that cases were thought provoking and reminiscent of clinical decision making encountered in their professional roles. They also noted that completing the test took longer than they anticipated.

Discussion

In this study, we piloted a script concordance test to evaluate the feasibility of using SCT to assess clinical decision-making in genetic counseling trainees. The script concordance test presents students with common clinical scenarios in conditions of uncertainty and asks them to determine the impact of additional information on initial actions or impressions.

Clinical decision-making is a skill that is often difficult to measure in a standardized way (Ilgen et al., 2018). The script concordance test is scored by measuring the degree of concordance between a student's response and the responses from expert panelists. The higher the score on the script concordance test, the greater the degree of concordance, therefore indicating clinical decision-making skills similar to those of experts in the field. The results of the current pilot study established the reliability and validity of SCT in measuring clinical decision-making in genetic counseling trainees, similar to results from previous studies in other healthcare professions (Dawson et al., 2014; Lambert et al., 2009). The script concordance test revealed differences in clinical decision-making of students and panelists, and highlighted a wider spread of responses in students than in panelists. Both findings align with the point made by Charlin et al. (2000) that clinical decision-making skills develop over time and would be more developed in working genetic counselors compared to students. By creating a standardized assessment that is representative of clinical experiences, students' clinical decision-making can be evaluated by program leadership who may not observe students utilizing these skills in clinical settings, and may offer a particularly useful direct comparison to how this competency is applied by practicing genetic counselors. A study done by Lambert et al., (2009) determined that students that score two standard deviations below the student mean may require remedial action. In the field of genetic counseling, this can provide an opportunity during training to identify students who may require additional support in this area.

Through the development and implementation of a script concordance test, this pilot study demonstrated the feasibility of using SCT to assess clinical decision-making in genetic counseling trainees. SCT provides a reliable, standardized way to measure clinical decision-

making skills. It is easier to administer and score compared to alternative methods such as multiple choice questions, oral examinations or assignments.

Diagnostic Cases

While genetic counselors often construct a list of differential diagnoses to guide testing strategy, they are not diagnosticians (McCarthy Veach et al., 2018). By incorporating elements of uncertainty into the cases, the Diagnostic cases on the script concordance test for genetic counseling trainees ask participants to evaluate the impact of new information on their lists of differential diagnoses (Figure 1a and Figure 3a). A challenge to construction of cases in this category is that much of the ability to engage in the diagnostic process requires medical knowledge, making clinical reasoning difficult to tease apart from the medical knowledge upon which it relies (Lubarksy et al., 2011). Across both panelists and participants, standard deviation (SD) was largest in this case category (Table 1), which could be explained by variation in the knowledge base required to engage in clinical decision-making (Lubarksy et al., 2011). While genetic counselors receive generalized training, many specialize in one field of genetics once they begin working. Having an area of expertise in genetics may account for the larger standard deviation among panelists' scores when, for example, a genetic counselor who specializes in reproductive genetics was asked to respond to a case about pediatric genetics. Creating cases which are representative of clinical decision-making skills required for all professional subspecialties is an important goal for future iterations of the script concordance test.

Anchored Cases

The largest effect size between panelists' and students' responses was seen in the Anchored cases where Hedge's *g* was 1.91 (Table 1), indicating a substantial difference between the clinical decision-making skills of these two groups. The Anchored cases also received the

students' lowest mean score of 57%, whereas the panelists' mean score of 74% was equivalent to their mean score for the Diagnostic cases (Table 1). There are various explanations that may contribute to the increased rate of discordance seen in the Anchored cases. One notable explanation is that with a 5-point Likert scale for Part A and all subsequent questions, the Anchored cases contained the greatest number of possible response combinations across all case categories. This, in turn, created more opportunities for no credit responses as a smaller proportion of answer combinations could receive credit using this question format.

Ethics Cases

The Ethics cases consisted of common ethical dilemmas encountered in the field of genetic counseling, such as genetic testing of a minor and patient confidentiality (Figure 1c and Figure 3c). While students have learned about principles of ethics and policy in their coursework and clinical rotations, SCT provides an opportunity to navigate these considerations unsupervised. Across all case categories, the students' mean score was the highest (Table 1) in the Ethics cases, indicating a higher percentage of credit responses compared to the other case types. This is likely explained by the wide range of responses from panelists in the Ethics case category, where there were fewer answer options on the scoring grid deemed as no credit. Pau et al. (2019) theorize that ethical reasoning is an area of extreme ambiguity across clinical practice and is therefore expected to vary among working professionals. Additionally, ethical reasoning skills develop over time and are related to the years spent in clinical practice. The panel members in this study consisted of a combination of new graduates and certified genetic counselors who have been practicing in the field for many years, which may have accounted for a wider spread of possible responses that received credit and an overall higher mean score for the students.

Both students and panelists found the cases to be representative of scenarios that would be encountered in genetic counseling clinical practice and affirmed that the questions were realistic and applicable. SCT can also serve as an opportunity for learning when accompanied by a written explanation of panelists responses and a follow-up group session where student's responses can be compared. Participants in this study reported completing the test with ease, and overall impressions were consistent with findings from a previous study where participant impressions were recorded (Lambert et al.. 2009). In this pilot study, the extended time requirement noted by panelists could be explained by the fact that they took the original version of the test, which consisted of 32 cases instead of the standard student version which consisted of 22 cases, and that they were providing written justification for their responses.

Limitations

In this study, all panelists were genetic counselors connected to, or involved in, the SLC Human Genetics program, and have established clinical decision-making skills that are embedded within the teaching framework and ideologies of the program. While this was intentional for this pilot, it likely impacted the degree of concordance seen across a variety of questions as the students in this study were instructed directly by many of the panel members. Additionally, panel members consisted of some clinical genetic counselors, some genetic counselors who are primarily focused on genetic counseling education, and some who are involved in both. For this reason, there were panel members who had not been involved in one, or many, of the specialties referenced on the script concordance test for many years, leading to possible discrepancies between their responses and current clinical practices. Consistent with a pilot study, the sample size was small and confined to one training program.

Implications for Future Use

The universality of common clinical encounters in genetic counseling allows for widespread use of SCT for training programs. Genetic counseling graduate programs in North America utilize the Practice-Based Competencies set by the ACGC (2019) to receive and maintain accreditation. Thus additional research can be done to confirm that the script concordance test can be successfully incorporated into the learning and evaluation of genetic counseling trainees across the country.

SCT is a relatively flexible method of assessment that could be routinely updated to reflect shifts in professional practice. Case scenarios can be refined and updated as testing technology and guidelines change. When this occurs, a new scoring system must be created. While this would require panel members to retake the script concordance test, it would not require any prior test preparation on the part of the working genetic counselor, as the scenarios are designed to be ones commonly encountered in practice (Fournier et al., 2008). The biggest barrier to future iterations would be the time commitment required by expert panelists.

SCT has the potential to be used as a method of assessment not only for clinical decision-making skills in genetic counseling students, but also to assess competence in practicing genetic counselors for the purposes of ongoing credentialing. This serves as a possible area for future research.

Conclusion

SCT has proven to be a valid, reliable, and successful method for assessing clinical decision-making skills in healthcare providers in situations that are ambiguous or contain a high degree of uncertainty. This pilot study demonstrates the feasibility and ease with which SCT can be used to assess clinical decision-making skills in genetic counseling trainees. The use of SCT

will enable genetic counseling training programs to more effectively evaluate clinical decision-making skills, thereby tracking the development of critical practice-based competencies.

References

Accreditation Council for Genetic Counseling. (2019). Practice-Based Competencies for Genetic Counselors. ACGC. Retrieved April 4, 2022, from https://www.gceducation.org/practice-based-competencies/

Charlin, B., Brailovsky, C., Leduc, C., & Blouin, D. (1998). The Diagnosis Script Questionnaire: A New Tool to Assess a Specific Dimension of Clinical Competence. *Advances in Health Sciences Education: Theory and Practice*, 3(1), 51–58. https://doi.org/10.1023/A:1009741430850

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & van der Vleuten, C. (2000). The Script Concordance test: a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 12(4), 189–195. https://doi.org/10.1207/S15328015TLM1204_5

Charlin, B., Tardif, J., & Boshuizen, H. P. (2000). Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Academic Medicine: Journal of the Association of American Medical Colleges*, 75(2), 182–190. https://doi.org/10.1097/00001888-200002000-00020

Dawson, T., Comer, L., Kossick, M. A., & Neubrander, J. (2014). Can script concordance testing be used in nursing education to accurately assess clinical reasoning skills?. *The Journal of Nursing Education*, 53(5), 281–286. https://doi.org/10.3928/01484834-20140321-03

Doyle, D. L., Awwad, R. I., Austin, J. C., Baty, B. J., Bergner, A. L., Brewster, S. J., Erby, L. A., Franklin, C. R., Greb, A. E., Grubs, R. E., Hooker, G. W., Noblin, S. J., Ormond, K. E., Palmer, C. G., Petty, E. M., Singletary, C. N., Thomas, M. J., Toriello, H., Walton, C. S., & Uhlmann, W. R. (2016). 2013 Review and Update of the Genetic Counseling Practice Based Competencies by a Task Force of the Accreditation Council for Genetic Counseling. *Journal of Genetic Counseling*, 25(5), 868–879. https://doi.org/10.1007/s10897-016-9984-3

Fine, B. A., Baker, D. L., Fiddler, M. B., & ABGC Consensus Development Consortium (1996). Practice-based competencies for accreditation of and training in graduate programs in genetic counseling. *Journal of Genetic Counseling*, 5(3), 113–121. https://doi.org/10.1007/BF01408656

Fournier, J.P., Demeester, A. & Charlin, B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak* 8, 18 (2008). https://doi.org/10.1186/1472-6947-8-18

Gagnon, R., Charlin, B., Coletti, M., Sauvé, E. and Van Der Vleuten, C. (2005), Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test?. *Medical Education*, 39: 284-291. https://doi.org/10.1111/j.1365-2929.2005.02092.x

Gagnon, R., Charlin, B., Lambert, C., Carrière, B., & Van der Vleuten, C. (2009). Script concordance testing: more cases or more questions?. *Advances in Health Sciences Education: Theory and Practice*, 14(3), 367–375. https://doi.org/10.1007/s10459-008-9120-8

Gagnon, R., Lubarsky, S., Lambert, C. et al. Optimization of answer keys for script concordance testing: should we exclude deviant panelists, deviant responses, or neither?. *Adv in Health Sci Educ* 16, 601–608 (2011). https://doi.org/10.1007/s10459-011-9279-2

Humbert, A. J., & Miech, E. J. (2014). Measuring gains in the clinical reasoning of medical students: longitudinal results from a school-wide script concordance test. *Academic medicine: Journal of the Association of American Medical Colleges*, 89(7), 1046–1050. https://doi.org/10.1097/ACM.0000000000000267

Ilgen, J. S., Eva, K. W., de Bruin, A., Cook, D. A., & Regehr, G. (2019). Comfort with uncertainty: reframing our conceptions of how clinicians navigate complex clinical situations. *Advances in Health Sciences education: Theory and Practice*, 24(4), 797–809. https://doi.org/10.1007/s10459-018-9859-5

Kaur, M., Singla, S., & Mahajan, R. (2020). Script concordance test in pharmacology: maiden experience from a medical school in India. *Journal of advances in medical education & professionalism*, 8(3), 115–120. https://doi.org/10.30476/jamp.2020.85163.1168

Lambert, C., Gagnon, R., Nguyen, D. et al. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol* 4, 7 (2009). https://doi.org/10.1186/1748-717X-4-7

Lubarsky, S., Chalk, C., Kazitani, D., Gagnon, R., & Charlin, B. (2009). The script concordance test: a new tool assessing clinical judgement in neurology. *The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques*, 36(3), 326–331. https://doi.org/10.1017/s031716710000706x

Lubarsky, S., Charlin, B., Cook, D. A., Chalk, C., & van der Vleuten, C. P. (2011). Script concordance testing: a review of published validity evidence. *Medical education*, 45(4), 329–338. https://doi.org/10.1111/j.1365-2923.2010.03863.x

Lubarsky, S., Dory, V., Duggan, P., Gagnon, R., & Charlin, B. (2013). Script concordance testing: from theory to practice: AMEE guide no. 75. *Medical teacher*, 35(3), 184–193. https://doi.org/10.3109/0142159X.2013.760036

Lubarsky, S., Dory, V., Meterissian, S., Lambert, C., & Gagnon, R. (2018). Examining the effects of gaming and guessing on script concordance test scores. *Perspectives on Medical Education*, 7(3), 174–181. https://doi.org/10.1007/s40037-018-0435-8

McCarthy Veach, P., LeRoy, B., & Callanan, N. P. (2018). *Facilitating the genetic counseling process practice-based skills (2nd ed.)*. Springer International Publishing.

Nussbaum, R. L., McInnes, R. R., Willard, H. F., & Hamosh, A. (2016). *Thompson & Thompson Genetics in Medicine* (8th ed.). Elsevier.

Pau, A., Thangarajoo, S., Samuel, V.P. et al. Development and validation of a script concordance test

(SCT) to evaluate ethical reasoning ability among first and fifth year students in a medical school. *J Acad Ethics* 17, 193–204 (2019). https://doi.org/10.1007/s10805-019-09327-3

Petrucci, A. M., Nouh, T., Boutros, M., Gagnon, R., & Meterissian, S. H. (2013). Assessing clinical judgment using the script concordance test: the importance of using specialty-specific experts to develop the scoring key. *American Journal of Surgery*, 205(2), 137–140. https://doi.org/10.1016/j.amjsurg.2012.09.002

Steinberg, E., Cowan, E., Lin, M. P., Sielicki, A., & Warrington, S. (2020). Assessment of emergency medicine residents' clinical reasoning: validation of a script concordance test. *The Western Journal of Emergency Medicine*, 21(4), 978–984. https://doi.org/10.5811/westjem.2020.3.46035

Subra, J., Chicoulaa, B., Stillmunkès, A., Mesthé, P., Oustric, S., & Rougé Bugat, M. E. (2017). Reliability and validity of the script concordance test for postgraduate students of general practice. *The European Journal of General Practice*, 23(1), 208–213. https://doi.org/10.1080/13814788.2017.1358709

Tiffen, J., Corbridge, S. J., & Slimmer, L. (2014). Enhancing clinical decision making: development of a contiguous definition and conceptual framework. *Journal of Professional Nursing: Official Journal of the American Association of Colleges of Nursing*, 30(5), 399–405. https://doi.org/10.1016/j.profnurs.2014.01.006

Tsai T. C. (2017). Twelve tips for the construction of ethical dilemma case-based assessment. *Medical teacher*, 39(4), 341–346. https://doi.org/10.1080/0142159X.2017.1288862

Wan, M., Tor, E., & Hudson, J. N. (2020). Examining response process validity of script concordance testing: a think-aloud approach. *International Journal of Medical Education*, 11, 127–135. https://doi.org/10.5116/ijme.5eb6.7be2