The use of contextualised standardised client simulation to develop clinical reasoning in final year veterinary students

Claire EK Vinten; BVMedSci BVM BVS PhD FHEA MRCVS, is Lecturer in Veterinary Education, Royal Veterinary College, The LIVE Centre, Hawkshead Lane, Hertfordshire, AL9 7TA; cvinten@rvc.ac.uk; ORCID: 0000-0002-9395-8431.

Kate A Cobb, BVetMed PGCE MMedSci PhD MRCVS, is Lecturer in Teaching Learning and Assessment, University of Nottingham School of Veterinary Medicine and Science, Loughborough, Leics, LE12 5RA; katy.cobb@nottingham.ac.uk.

Liz H Mossop, BVM&S MMedSci (Clin Ed) PhD MAcadMEd MRCVS, is Associate Professor of Veterinary Education, University of Nottingham School of Veterinary Medicine and Science, Loughborough, Leics, LE12 5RA; liz.mossop@nottingham.ac.uk.

Key words: Clinical reasoning, simulation, veterinary education, veterinary medicine

Word count: 5223

Abstract

Clinical reasoning is an important skill for veterinary students to develop prior to graduation. Simulation has been studied in medical education as a method for developing clinical reasoning in students, but evidence supporting this is limited. This study involved the creation of a contextualised standardised client simulation session aiming to improve the clinical reasoning ability and confidence of final year veterinary students. Sixty-eight participants completed three simulated primary-care consultations, with the client being played by an actor and the pet by a healthy animal. Survey data showed 100% of participants felt the session improved their ability to make clinical decisions. Quantitative clinical reasoning self-assessment, performed using a validated rubric, triangulated this finding – showing an improvement in student perception of several components of their clinical reasoning skill level before and after the simulation. Blinded researcher analysis of the consultation video-recordings found the 'History-taking' and 'Making sense of data' (including differential diagnosis formation) components of the assessment rubric showed a significant increase in ability. Thirty students took part in focus groups investigating their experience within the simulation. Two themes arose from thematic analysis of this data: *Variety of reasoning methods* and '*It's a different way of thinking'.* The latter highlights differences between the decision-making students practice during their time in education, and the decision-making they will use once working in practice. The study findings suggest that simulation can be used to develop clinical reasoning in veterinary students, and demonstrates the need for further research in this area.

Introduction

The use of simulation in veterinary education has grown in the last 10 years. This has been mainly driven by the increasing importance placed on communication training (1) and clinical skills teaching, coupled with the overwhelming acceptance of the pedagogical value of simulation within the fields of human medicine and nursing. It may also be due, in part, by the increasing numbers of veterinary students at universities makes time practicing clinical skills competitive and limited (2,3). However, simulation use within veterinary schools remains very limited compared to other healthcare fields.

Simulated clients (SCs) are commonly used to develop communication skills in veterinary students (1). Actors recreate the experience of conversing with a client so that students may practice the techniques of history taking, dealing with conflict and breaking bad news. Although effective at improving communication (4), SCs are rarely used for any other skill development in veterinary education.

Clinical reasoning is the skill used when veterinary surgeons make a decision regarding the diagnosis, treatment plan or prognosis of a patient (5). There are two cognitive processes a practitioner can employ to solve these clinical problems – known as systems one and two reasoning (6). System one is fast, unconscious and intuitive, whilst system two is slow, logical and analytical (7). Whilst they can be used exclusively, they have been shown to be most accurate when used in combination (8,9) – known as 'dual processing'. As expertise in clinical reasoning develops, students move away from a detailed pathophysiology-based view of disease (system two), and begin to form readily-accessible illness scripts that permit a form of diagnostic pattern-recognition (system one). However, experts retain the ability to to switch back to a slower, logical method of decision making if they wish (dual-processing)(10,11).

Although thoroughly researched in medical domains, relatively little is known about veterinary clinical reasoning (12,13). Even less is understood about how to 'teach' clinical reasoning to veterinary students, thus most recommendations have been extrapolated from medical research (14). This is not ideal, as it has been suggested that veterinary surgeons integrate non-clinical factors such as finances and owner preferences to a greater degree than their medical counterparts (12) – indicating different training needs. Vinten et al. (5) conducted a qualitative investigation into clinical reasoning development at one UK veterinary school – finding that graduates faced a steep learning curve when entering practice. This is both supported (15) and refuted (16) by survey data from other authors. Vinten et al. recommended included incorporating contextual factors into decision-making training, and recreating the experience of responsibility for clinical outcomes – without which students rely on clinicians present to prevent any harm to their patients.

Several studies have indicated that simulation might improve clinical reasoning in both medical and nursing students (17–22). However, due to the inherent difficulties in definitively measuring clinical reasoning, no research has provided strong enough evidence to be conclusive on this matter. There has been no research investigating the relationship between simulation and clinical reasoning in veterinary students, to the authors' knowledge.

This study aimed to assess the effect of novel primary care consultation simulation on the clinical reasoning ability of final year veterinary students and explore the student experience of clinical reasoning within a simulation scenario. Ethical approval was granted by the University of Nottingham.

Methods

Simulation session design

The simulation was aimed at final year students, designed to recreate a first opinion small animal consultation as closely as possible. The intended reasoning-based learning outcomes were as follows:

1. Make clinical decisions confidently
2. Formulate differential diagnoses and diagnostic or treatment plans for a range of clinical cases
3. Reflect on clinical decisions that have been made

Key features found to promote effective learning within simulations by Issenberg et al. (23) were incorporated where possible (Table 1). Three cases were developed from genuine patients examined and treated by one of the authors (CV) within a primary care veterinary surgery. These were checked for authenticity by two experienced veterinary surgeons (a summary of the cases is provided in table 2). Clients were played by trained actors and patients played by healthy dogs belonging to the authors.

INSERT TABLES 1 AND 2 HERE

Prior to the simulation, students were provided with a very short description of each case (e.g. 'weight loss') in order to allow them to research the relevant topics, but no information or tuition on clinical reasoning theory or methods. The session took place in a consultation room within a small animal hospital, thus was already fully equipped. After an introduction and familiarisation period, students were given a clinical history for their first patient, detailing only *previous* treatment at the fictional surgery. When ready, the student collected the SC and their pet from the hospital waiting room. The structure of the consultation was controlled by the student and ended with the SC exiting the room. Each simulation lasted roughly 15 minutes. The students were instructed to treat the simulation as if it were a real consultation; responding to the concerns of the client in an appropriate way, discussing possible diagnoses and treatment options and prescribing any necessary medication. A 15-minute debriefing using the model of Good Judgement (24) was then performed by a member of staff who had observed the consultation through a live video feed. Each student participated in all three cases in a randomised order. An overview of the simulation process for each student is shown in Figure 1.

All students that undertook a placement at the small animal hospital during the 10-month study period were required to take part in the simulation, but enrolling in the associated research project was voluntary. Participants were separated into two cohorts – Group A took part in the simulation within the first 6 months of their final year of study, group B within the last 6 months. This was due to the timing of the study, which fell across two academic years, but provided opportunity to observe the effect of the simulation at different points in the curriculum.

Quantitative measurement of simulation impact

Due to the known difficulties objectively measuring clinical reasoning, three methods of data collection were used in order to triangulate any findings. The Lasater clinical judgement rubric (LCJR), developed by Lasater (25) was chosen to grade clinical reasoning ability because a) it is specific for use within high fidelity simulation, allowing grading of physical actions and conduct rather than written answers and b) it could be modified to give a quantitative score of clinical reasoning skill. The components of the rubric were designed to specifically relate to clinical reasoning – for example, the ‘History Taking’ component measures directed questioning relevant to the case, rather than the associated communication skills such as summarising or screening.

As the LCJR was developed to examine clinical judgement in human nursing, rather than veterinary medicine, minor modifications were made to ensure it was suitable for a veterinary application. These included changing of certain words (e.g. ‘patient’ to ‘client’) and the removal of irrelevant areas of assessment (i.e. skills that would not be used).

The modified Lasater Clinical Judgement Rubric (mLCJR) and the three clinical cases (table 3) were piloted using a test simulation. One experienced veterinary surgeon was video-recorded completing the three simulated consultation cases. The rubric was then used to assess the performance of the participant. No changes were necessary to the simulation cases following the pilot study, but the mLCJR was modified to include a representative example of student performance for each score category (appendix 1).

The mLCJR was used in two ways during the simulation. Firstly, students were asked to score their own clinical reasoning ability pre and post simulation using the rubric. This was performed immediately before the first consultation, and after the debriefing period of the last consultation (self-assessment – SA). Secondly, the participant’s clinical reasoning was scored by a researcher using the rubric (researcher assessment – RA). The first and third consultations each student conducted were video recorded – a process the students were familiar with from communication training earlier in the course. After completion of data collection, these videos were blinded, randomised and scored by researcher CV; a small animal veterinary surgeon experienced in teaching clinical reasoning. Ten percent of the video recordings were also scored by a second researcher, also a veterinary surgeon, allowing the interrater reliability to be calculated. This was done by calculating the Intraclass Correlation Coefficient using SPSS statistics 22 (IBM).

To determine whether the data from groups A and B could be amalgamated, the difference between the pre and post simulation scores of each student were calculated for both the SA and RA. These were input into SPSS statistics 22 (IBM) and A Mann-Whitney U test comparing the improvement of each group was performed on each mLCJR component separately. There was a statistically significant difference in the score-change between the two groups on the SA, so the data sets were not merged. There was not a significant difference for the RA, so the data for groups A and B were combined.

The following methods were performed separately on groups A and B when evaluating the SA and once on the combined data from both groups when analysing the RA.

The pre and post simulation scores were compared using a Wilcoxon Signed Ranks test. Each component of the mLCJR was analysed individually. Median and mean averages were calculated for each component, both pre and post simulation.

To determine whether the components could be summed to create an overall pre/post total for each group, Cronbach's alpha value of internal consistency was calculated. As all alpha-values fell above 0.7, the consistency was accepted within all four categories (Group A SA/RA, Group B SA/RA) and the components summed (26). A Wilcoxon signed ranks test was then performed on the totalled data.

Construct validation

To determine the construct validity of the mLCJR, a cohort of experienced veterinary surgeons were tested using the rubric. A purposive sample of seven university staff members that had over three years' experience as a first-opinion small animal veterinary surgeon and had worked in practice within the last 12 months were selected. All took part in one simulated consultation and were video recorded.

The expert participants' recordings were graded by a researcher (CV) and the median and mean average total score calculated. Blinding was not possible as, due to the age of the experts compared to the students, the identity of the staff was unavoidably clear.

To compare the expert and student performances, all student total scores were combined with the expert total score data set. A Mann-Whitney U test was used to identify any significant ability differences between the two groups.

Survey analysis of simulation impact

A Likert-scale survey was designed to collect student opinions about the simulation. Survey responses were converted to numerical data for analysis, where Strongly disagree = 1 and Strongly agree = 6. A Mann-Whitney U test indicated that group A and B should be further analysed separately.

To determine if the questions could be summed to a total, Cronbach's alpha was performed. As both groups alpha values returned above 0.7 (26) the total score for each student was calculated. For both cohorts, the median and mean averages were determined for each question. The total percentage agreement with each question was then calculated.

INSERT FIGURE 1 HERE

Qualitative insights into simulation impact

Focus groups were conducted with 30 of the 68 students that took part in the simulation. Participants were selected using convenience sampling, due to their busy schedule whilst on final year work placements. Six focus groups were held, each with five participants. Each focus group was held two days after the participants completed the simulation and was optional.

The focus groups followed a semi-structured format and lasted between 30 to 60 minutes. Questions focused on the experience of the students during the simulation; how the experience differed from other experiences of decision-making during their training and how participants felt they reasoned through the cases. All focus groups were audio recorded, transferred electronically to a computer and then transcribed verbatim, by either an external source or a researcher. Where transcription was done by an external source, the document was checked by the researchers for accuracy.

The transcriptions for all focus groups were merged into one data set for thematic analysis. Thematic analysis was performed using guidelines developed by Braun & Clarke (27). Complete inductive code generation was performed by one researcher (CV), managed through NVIVO (QSR, version 10). One focus group transcript was coded by a second researcher (LM) and agreement reached in order to ensure consistent approach to coding. Codes were then interpreted and grouped together by that researcher to form subthemes and themes. These themes were iteratively revised and edited. Once complete, the themes were reviewed by the remainder of the research group (KC, LM) and changes were made, which prompted another round of iterative revision and editing. When finished, the group reviewed the final themes once more and agreed on their interpretation.

Results

Sixty-eight students took part in the simulation – 32 in group A and 36 in group B. A confidence interval of 95% was selected as a measure of statistical significance.

Student self-assessment

Group A reported significant improvement in all components of the mLCJR (Table 3). Group B showed significant improvement in four out of eight components: History-taking, Identifying abnormalities, Making sense of data and Well planned intervention (table 4).

INSERT TABLES 3 AND 4 HERE

Cronbach's alpha showed acceptable reliability to sum total 'before' and total 'after' scores ($\alpha=>0.7$). A Wilcoxon Signed-Ranks Test indicated that post-simulation scores were statistically significantly higher than pre-simulation scores for both groups (A: $Z=-4.61$, $p=<0.001$; B: $Z=-3.44$, $p=0.001$). A Mann-Whitney test then showed that the level of improvement was greater for group A (Mdn=1, Mn=1.88) than for Group B (Mdn=1, Mn=1.26), $U=340.0$ $p=.0006$).

Researcher assessment

The two assessors reached an ICC of 0.894 ($p=<0.05$) after marking 10% of the video recordings, indicating 'almost perfect' inter-rater reliability (28).

None of the mLCJR components showed a statistically significant difference in score between groups A and B, so datasets were combined for further analysis. Within this combined data, two mLCJR components showed significant improvement as a result of the simulation: History taking and Making sense of data (table 5).

INSERT TABLE 5 HERE

Cronbach's alpha showed acceptable reliability to sum total 'before' (α=67) and total 'after' scores (α=0.75). The Wilcoxon signed ranks test showed no significant difference between total scores (table 5).

Construct validation of mLCJR

Seven expert participants took part in the validation simulation. A Mann-Whitney test indicated that total scores were higher for the expert group (Mdn=31.00) than the student group (Mdn=27.00), U=43.00 P=0.003. This suggests the mLCJR has an acceptable construct validity.

Survey

A Mann-Whitney test showed the two groups answered nine questions significantly differently, therefore data was not merged (table 6). In both groups, 100% of students reported feeling more confident in making decisions, reaching a diagnosis and forming a treatment plan. The median and mean averages show group A answered all questions with a higher level of agreement than group B (table 6).

INSERT TABLE 6 HERE

Cronbach's alpha showed excellent internal consistency of both group A (α=0.84) and group B (α=0.86). A Mann-Whitney test indicated that the total level of agreement (table 6) was greater for group A (Mdn=82.00) than for group B (Mdn=77.00), U=284.50 P=0.001.

Qualitative data

Two key themes emerged from the focus group data, which are described below with supporting quotes from the transcripts. Each focus group has been assigned a code - FG1, FG2, FG3, FG4, FG5 or FG6.

Theme one: 'It's a different way of thinking…'

During the analysis, it became clear that the clinical reasoning taking place within the simulation had many differences from other decision-making experiences students had had in the curriculum. Three key factors were described as being novel. Firstly, the students described the simulation as being their first experience of making clinical decisions alone. They spoke of using the clinician usually present in consultations as a 'safety-net'; ensuring that any mistakes they make are corrected before they have consequences. Thus, they felt their decisions were always 'checked' and approved.

'(In other consultations) you have got that safety net behind you… if you say 'I'm thinking about this' and they say 'Well maybe, but think about...' you have always got someone there pointing you in the right direction.' FG1

'(In the simulation) all the responsibility is on you – it's the first time we have properly had it all on us in a way… because you have always got a clinician as a back-up in every other case we've been doing.' FG2

Students felt that having a clinician present in other consultations has removed their sense of case responsibility. Being alone in the simulation helped to create the experience of having sole charge over decision-making – despite the fact the clients and patients were not real.

'I just found it quite generally daunting taking on the consult and prime responsibility…where you did not have anyone to rely upon for the first time.' FG6

Secondly, the students were not used to making clinical decisions in pressurised situations. They felt that having a client in the consultation room forced them to make decisions faster. Students described the consultations they perform with clinicians (which are normally given triple the standard appointment time allowance) as slow-paced, and thus the skill of thinking under pressure is not practiced.

'You have to make quite a quick decision (in the simulation)... Where I think with (clinicians) you can have a nice chat and discuss your different options and then decide which ones are sensible to go with.' FG1

'(In the simulation) you have got to make the decision there and then, you haven't got time to go away and think about it…' FG3

Students also commented that the pressure of the consultation did not allow for the same reasoning processes they have developed on paper through case based learning and assessment. It was suggested that thinking 'in your head' is harder than reasoning on paper or similar, and thus the opportunity to practice it was valuable.

'It's a different way of thinking though, isn't it, because when you've, when you write it on paper you're working through in stages, whereas if you're in conversation you have to skip half of that stuff' FG5

'It is one thing being able to write on a piece of paper what you are thinking and sit there and look at what you have put down, but it is another thing processing it all in your brain and your head and thinking about what you need to ask and then thinking of what other possible things it could be.' FG6

The integration of situational factors was the third aspect of clinical reasoning within the simulation that students found novel. This involved combining their decision-making skills with communication, considering the owners needs and administrative tasks.

'You are multi-tasking in the simulation because you are also thinking what am I projecting to the client? How am I going to explain it to the client? Am I being clear?' FG1

'We've never, ever had to deal with money before, we've never had to think about prices, or trade names…' FG4

'On paper you could be like 'Go home on a bland diet, whatever' – but (in the simulation) there is a client, waiting, stood there, probably expecting antibiotics or something… so that's different because you have to manage client expectations.' FG3

Students appear to process information differently to draw conclusions within the simulation compared to case-based learning sessions, examinations and clerkship consultations. They are learning to think in different way to cope with the time pressures and multi-tasking required.

Theme two: Variety of reasoning methods

Students reported using both system one and system two reasoning. They were not consciously aware of the difference, but it became clear through their discussion that this was the case.

'Sometimes I find it hard to explain how I came up with the solution, sometimes it does just ping there like 'Oh I think this is what I should do'.' FG3

'My brain doesn't just go like (clicks fingers) … it always takes me a longer time for some reason.' FG4

It was also clear from the data that there was a degree of case specificity affecting the ability to make clinical decisions. Students disagreed on which case was most complicated, and their opinions generally reflected their level of knowledge about each pathology.

'I felt that the one consult that I did better in was the one that I knew more about and you felt more comfortable with.' FG5

Discussion

The effect of standardised client simulation on clinical reasoning development

The RA showed improvement in only two of the components of the mLCJR – history taking and making sense of data. The latter of these focuses on the formation of differential diagnoses, arguably a key aspect of clinical reasoning and one the session aimed to improve. The former, history taking, is a skill that the fifth year students involved in the simulation were already expected to be proficient in. One explanation for the noticeable improvement in history taking may be that the task actually required the formation of differential diagnosis in order to ask the necessary question to rule each in/out. Although students have practiced the communicatory tasks of history taking previously, they had limited opportunities to combine this with a diagnostic task. This theory is supported by the work of Nendez et al. (29). They found that the diagnostic accuracy of students, residents and practitioners all decreased when only a 'chief complaint' was provided and further data collection was required, opposed to a full clinical vignette. The authors discovered the reason behind poor performance with chief complaint scenarios was the failure to gather sufficient information during the history taking process, despite being given (when asked) more information than the vignettes provided. The authors conclude that the teaching of history taking should be integrated with reasoning tasks, so that students practice using the two in conjunction and thus are able to apply this model when in practice. If extrapolated to veterinary medicine, this theory could explain the improvement in history taking, despite it not being a focus of the simulation; i.e. by reviewing the formation of differential diagnoses during the debriefing, the ability to structure data gathering also improved. In an investigation of the structure of veterinary consultations, Everitt (12) found that the history taking process was interweaved with the physical examination – suggesting that the former is used to inform the latter and vice versa. This theory further supports the increase in history taking ability being an indicator of clinical reasoning improvement.

The SA and RA do not appear to agree on the level of development during the simulation. One possibility is that students have over-estimated their improvement, or simply gained confidence but not measurable skill. A second possibility is that case specificity affected the student's objective skill level between cases. Case specificity was first noted by Elstein et al. (30) when they observed that the diagnostic ability of a physician varied – scoring well on one case examination was not an indicator of future performance. The implication of this was that knowledge plays a role in clinical reasoning; it is not simply a generalizable skill (31). Further research has shown that actually a combination of knowledge and general problem-solving ability is needed for successful reasoning (31–33), however no studies exclude the need for domain specific knowledge. If this theory were applied to this study, a student that had greater knowledge about, for example, idiopathic epilepsy would be more likely to perform well during that case simulation, regardless whether it was their first or last consultation. If their knowledge of acute diarrhoea and weight loss causes was significantly lower, any reasoning skill development might become negligible. As the students in this study were provided with a case-list several days prior to the simulation it was expected they would research the topics, thus reducing the effect of subject-specific knowledge. However, whether or not the students did partake in revision was not measured and so it is difficult to estimate the influence of case specificity. If this study were to be repeated, providing reading material or a lecture on the topics to be addressed in the forthcoming consultations and then a test of mastery would help to reduce case specificity. There would always be, however, the effect of personal experience on knowledge and decision-making that would case some degree of bias.

One further factor may have contributed to the difference between the RA and SA score improvement. Three components of the mLCJR – Examination, Identifying abnormalities and Prioritising data – had a RA 'first consultation' median score of four; the highest possible mark. This means that it was not possible for students to improve in those areas (in a way that was recognisable on the mLCJR). It is likely that this arose due to a mismatch between student ability and simulated consultation difficulty. In future work, increasing the difficulty of the cases could reduce this effect. As it may not be possible to manipulate the physical examination task, this component might need to be removed from the mLCJR.

Differences between the research groups

Group A (early) reported a significantly larger degree of improvement than group B (late) during the SA within four categories: History taking, examination, calm confident manner and clear explanation. It can be argued that these four components of the mLCJR are covered well within veterinary curricula – particularly in the final year of the course, when students engage in workplace-based learning. The fact that group B did not improve as much in these four categories as group A suggests that teaching and repetitive practice in fifth year might improve their perceived ability in these areas to a level at which they felt proficient by the time the simulation was conducted. The remaining components - identifying abnormalities, prioritising data, making sense of data and forming a well-planned intervention – represent key mental tasks during clinical decision-making. These components showed the same increase within both group A and B, implying that there is little perceived improvement in ability during fifth year. Overall, this suggests that some components of clinical reasoning are being developed by the workplace-based learning, but essential mental processes are remaining unchanged throughout. This difference is not mirrored in the RA results, in which both groups of students performed equally.

Qualitative data results

Within the theme 'It's a different way of thinking…'. Students claimed their clinical reasoning process was different within a simulation, compared to consultations with clinicians, case-based learning or examinations. This has important consequences for how clinical reasoning should be taught, as the simulation closely resembles the day-to-day work of a veterinary surgeon and thus the way clinical reasoning will be used frequently upon graduation.

Students described the pressure of making decisions quickly within the simulation as something new that they have not experienced elsewhere; a way of reasoning that required different thought processes than they were used to. It is known that stress affects human decision-making – increasing the amount of risk-taking behaviour observed (34). In these circumstances, subjects use heuristics more frequently (35), possibly due to working memory overload. Studies of both veterinary surgeons and human physicians have shown that they suffer greater levels of stress than the general population, especially those recently graduated (36–38). The combination of these two factors – high stress and the impact of stress on decision-making – suggests educators need to be giving students opportunity to practice clinical reasoning under pressure. If the process of reasoning is different when time is not limited, then efforts to develop clinical reasoning in relaxed settings will not prepare students for making decisions in the real world. Simulation is known for causing stress in students – generally perceived as a negative consequence (20,39). However, this 'side-effect' of simulation-based education could be utilized for the students benefit. The timing of such an intervention would be critical – subjecting a student to decision-making under pressure before they are capable would only damage their confidence. But, for a student already competent at clinical reasoning in the classroom and clinic, simulation may provide the last key situation in which to master their skill.

Another major finding of this theme is that the simulation experience was the first time students had felt fully responsible for their own clinical decisions. Even when they are given opportunities to make decisions within WBL consultations, the students report a sense of security from the clinician present that prevents them from emotionally investing in their decision. The same problem has been reported previously in medicine, where the 'simplistic' approach to teaching clinical reasoning generates a 'sterile academic environment which avoids feelings of responsibility for any morbidity or mortality experienced by the patient as a consequence of making an inappropriate diagnosis' (40). Again, the effect of diminished responsibility is that students practice a cosseted form of clinical reasoning that is not fully representative of the skill they will need to use in practice. Thus, when they graduate, they are underprepared.

Student participants found situated decision-making another new challenge. They found incorporating owner factors particularly novel, alongside the need for multi-tasking. This probably results from the isolated nature of other clinical reasoning experiences - normally students make clinical decisions in an artificial environment where their only task is to develop an appropriate case management plan. This allows them to focus all their concentration on the decision-making process, which is not often possible in reality. On top of this, students do not always have the opportunity to complete clinical notes, prescribe and dispense drugs or calculate costs when participating in real consultations during clerkships. These form 'distractors' that interfere with clinical reasoning, however students rarely practice incorporating them into decision-making. Several studies have shown that contextual factors impact clinical decision-making (41–43), meaning teaching students to recognize and respond to these distractors is important. Again, students cited the SC simulation as an effective way to develop multi-tasking ability.

The theme 'variety of reasoning methods' developed from discussing with the students how they made decisions within the simulation. There were various methods described, including both systems one and two. This is not surprising, as Coderre et al. (44) not only showed that both system one and system two methods were used by students, but also that diagnostic accuracy was significantly higher when using the former. A later study by Ark et al. (8) found that students using dual process reasoning were most diagnostically successful. This has implication for veterinary education, as it indicated that system one reasoning should not be discouraged; in fact, students should be aware of it so they may utilise it correctly.

Limitations

To date, there is no published method of measuring clinical reasoning ability that has an acceptable construct validity. This study aimed to increase the validity and reliability of the results by using four methods of data collection to triangulate results and by attempting to evaluate the construct validity of the rubric used. However, this remains the biggest limitation of this study and the results must be interpreted accordingly.

The limitations of using self-reported data also need to be considered. These relate to the confines of introspection, and the ability to understand one's own subconscious decision-making process. Again, the use of triangulation minimizes the effect of any inaccuracy, but does not eliminate it. The fact that the focus group facilitator also facilitated the simulation may have impacted on the responses of participants. This possibly deterred students from criticising the session, however, all students were encouraged to reflect on the experience honestly and were made aware that their data would be anonymised and only used for the purpose of this research project. Finally, due to the time-span of this study, peer disclosure of the simulation structure could not be prevented using quarantine methods. The impact of this was minimised using two strategies: 1) students were asked not to discuss the cases outside of the simulation and 2) all participants were given information about the consultation topics before the simulation, thus reducing the impact of knowledge differences on the scores achieved.

Conclusion

In summary, this study has shown that standardised client simulation can be used to increase student confidence in clinical reasoning ability. There is also some evidence that simulation objectively improves some aspects of clinical reasoning, including differential diagnosis formation. This study also highlights the differences between the decision-making students practice during their time in education, and the decision-making they will use once working in practice. High fidelity simulation is indicated as one successful way to align the curriculum content to the career needs

References

1. Gray C, Blaxter A, Johnston P, Latham C, May S, Phillips C, et al. Communication Education in Veterinary Education in the United Kingdom and Ireland: The NUVACS Project Coupled to Progressive Individual School Endeavors. J Vet Med Educ. 2006 Mar;33(1):85–92.

2. Kneebone R, Baillie S. Contextualized simulation and procedural skills: a view from medical education. J Vet Med Educ. 2008;35(4):595–8.

3. Byron J, Johnson S, Allen C, Brilmyer C, Griffiths R. Development and pilot of Case Manager: a virtual-patient experience for veterinary students. J Vet Med Educ. 2014;41(3):225–32.

4. Latham C, Morris A. Effects of formal training in communication skills on the ability of veterinary students to communicate with clients. Vet Rec. 2007 Feb 10;160(6):181–6.

5. Vinten CEK, Cobb KA, Freeman SL, Mossop LH. An investigation into the clinical reasoning development of veterinary students. J Vet Med Educ. 2016;43(4).

6. Croskerry P, Petrie D, Reilly J, Tait G. Deciding about fast and slow decisions. Acad Med. 2014;89(2):197–200.

7. Canfield P, Whitehead M, Johnson R, OBrien C, Malik R. Case-based clinical reasoning in feline medicine: 1: Intuitive and analytical systems. J Feline Med Surg. 2016;18(1):35–45.

8. Ark T, Brooks L, Eva K. Giving learners the best of both worlds: Do clinical teachers need to guard against teaching pattern recognition to novices? Acad Med. 2006;81(4):405–9.

9. Ark T, Brooks L, Eva KW. The benefits of flexibility: The pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. Med Educ. 2007;41(3):281–7.

10. Schmidt H, Norman G, Boshuizen H. A cognitive perspective on medical expertise: theory and implication. Acad Med. 1990;65(10):611–21.

11. Schmidt H, Rikers R. How expertise develops in medicine: knowledge encapsulation and illness script formation. Med Educ. 2007;41(12):1133–9.

12. Everitt S. Clinical decision making in veterinary practice. University of Nottingham; 2011.

13. Vandeweerd J, Vandeweerd S, Gustin C, Keesemaecker G, Cambier C, Clegg P, et al. Understanding Veterinary Practitioners' Decision-Making Process: Implications for Veterinary Medical Education. J Vet Med Educ. 2012;39(2):142–51.

14. May S. Clinical Reasoning and Case-Based Decision Making: The Fundamental Challenge to Veterinary Educators. J Vet Med Educ. 2013;40(3):200–9.

15. Cobb K, Brown G, Hammond R, Mossop L. Alumni-based evaluation of a novel veterinary curriculum: are Nottingham graduates prepared for clinical practice? Vet

Rec Open. 2015;2(2):1–9.

16. Gilling M, Parkinson T. The transition from veterinary student to practitioner: A "make or break" period. J Vet Med Educ. 2009;36(2):209–15.

17. Cant R, Cooper S. Simulation-based learning in nurse education: systematic review. J Adv Nurs. 2010;66(1):3–15.

18. Steadman R, Coates W, Huang Y, Matevosian R, Larmon B, McCullough L, et al. Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. Crit Care Med. 2006;34(1):151–7.

19. Powell-Laney S, Keen C, Hall K. The Use of Human Patient Simulators to Enhance Clinical Decision-making of Nursing Students. Educ Heal. 2012;25(1):11–5.

20. Lasater K. High-fidelity simulation and the development of clinical judgment: Students' experiences. J Nurs Educ. 2007;46(6):269–76.

21. Gibbs J, Trotta D, Overbeck A. Human patient simulation versus case study: Which teaching strategy is more effective in teaching nursing care for the hypoglycemic patient? Teach Learn Nurs. 2014;9(2):59–63.

22. Yuan H, Williams B, Man C. Nursing students' clinical judgment in high-fidelity simulation based learning: A quasi-experimental study. J Nurs Educ Pract. 2014;4(5):7–15.

23. Issenberg B, McGaghie W, Petrusa E, Lee Gordon D, Scalese R. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach. 2005;27(1):10–28.

24. Rudolph J, Simon R, Rivard P, Dufresne R, Raemer D. Debriefing with good judgment: combining rigorous feedback with genuine inquiry. Anesthesiol Clin. 2007 Jun;25(2):361–76.

25. Lasater K. Clinical judgment development: using simulation to create an assessment rubric. J Nurs Educ. 2007;46(11):496–503.

26. Bland J, Altman D. Cronbach's alpha. Vol. 314, British medical journal (Clinical research ed.). 1997. p. 572.

27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77–101.

28. Royal K, Hecker K. Understanding Reliability: A Review for Veterinary Educators. J Vet Med Educ. 2015 Nov 11;e(e):1–4.

29. Nendaz M, Raetzo M, Junod A, Vu N. Teaching Diagnostic Skills: Clinical Vignettes or Chief Complaints? Adv Heal Sci Educ. 2000;5(1):3–10.

30. Elstein A, Shulman L, Sprafka S. Medical problem solving: An analysis of clinical reasoning. Vol 2. Cambridge, MA: Harvard University Press; 1978.

31. Wimmers P, Fung C. The impact of case specificity and generalisable skills on clinical performance: a correlated traits−correlated methods approach. Med Educ. 2008

Jun;42(6):580–8.

32. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? Med Educ. 2006 Jun 23;40(7):618–23.

33. Dory V, Gagnon R, Charlin B. Is case-specificity content-specificity? An analysis of data from extended-matching questions. Adv Heal Sci Educ. 2010;15(1):55–63.

34. Starcke K, Brand M. Decision making under stress: A selective review. Neurosci Biobehav Rev. 2012 Apr;36(4):1228–48.

35. Keinan G. Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. J Pers Soc Psychol. 1987;52(3):639–44.

36. Bartram D, Yadegarfar G, Baldwin D. Psychosocial working conditions and work-related stressors among UK veterinary surgeons. Occup Med (Chic Ill). 2009;59(5):334–41.

37. Gardner D, Hini D. Work-related stress in the veterinary profession in New Zealand. N Z Vet J. 2006;54(3):119–24.

38. Wall TD, Bolden R, Borrill S, Carter A, Golya D, Hardy G, et al. Minor psychiatric disorder in NHS Trust staff::Occupational and gender differences. Br J Psychiatry. 1997;171(December):519–23.

39. DeCarlo D, Collingridge D, Grant C, Ventre K. Factors Influencing Nurses' Attitudes Toward Simulation-Based Education. Simul Healthc J Soc Simul Healthc. 2008;3(2):90–6.

40. Patel R, Sandars J, Carr S. Clinical diagnostic decision-making in real life contexts: A trans-theoretical approach for teaching: AMEE Guide No. 95. Med Teach. 2014;37(3):211–77.

41. Durning S, Artino A, Boulet J, Dorrance K, van der Vleuten C, Schuwirth L. The impact of selected contextual factors on experts' clinical reasoning performance (does context impact clinical reasoning performance in experts?). Adv Heal Sci Educ. 2012;17(1):65–79.

42. Sibbald M, Panisko D, Cavalcanti R. Role of clinical context in residents' physical examination diagnostic accuracy. Med Educ. 2011;45(4):415–21.

43. Durning S, Artino A, Pangaro L, van der Vleuten C, Schuwirth L. Context and clinical reasoning: Understanding the perspective of the expert's voice. Med Educ. 2011;45(9):927–38.

44. Coderre S, Mandin H, Harasym P, Fick G. Diagnostic reasoning strategies and diagnostic success. Med Educ. 2003 Aug;37(8):695–703.

Feature	Description	Implications for the design of this study
Feedback	Providing a form of feedback to the learners regarding their performance	Detailed personal feedback was given to each student after every simulated consultation by the facilitator
Repetitive practice	Multiple opportunities for students to practice tasks - must be with the aim of improvement	Each student took part in three simulated consultations to allow them to practice decision-making multiple times and implement feedback given
Capture clinical variation	Portraying a variety of clinical cases to maximise case exposure	The three consultations the student took part in all simulated different clinical cases
Controlled environment	An environment where mistakes can be made safely and the facilitator can focus on the student, not the patient	The simulation was completely controlled – errors could be made without patient consequences
Individualised learning	Students should be active participants in a simulation experience that is individualised to each students needs	For this reason, students took part in the simulation alone and did not passively observe other students completing the simulation
Simulator validity	The simulation must have a high fidelity and be comparable to a genuine experience	The simulation was designed to be as high fidelity as possible – including the absence of peers/facilitators in the consultation area

Table 1: Components identified by Issenberg et al. (2005) to promote effective learning during simulation that were incorporated into the simulated consultation design.

Case	Signalment and history	Most likely diagnosis	Appropriate treatment plan example	Owner considerations
Acute diarrhoea	• 5-year-old dog • Watery diarrhoea lasting two days • No other relevant history • Clinical exam normal	Dietary indiscretion	Advise the owner to feed a bland diet (e.g. chicken breast) and administer digestive support paste (e.g. Protexin Pro-kolin) twice daily according to weight	Usually seen by the senior vet, who always prescribes antibiotics for diarrhoea.
Seizure	• 3-year-old dog • First seizure yesterday • No other clinical signs, change in behaviour or relevant history • Clinical exam normal	Idiopathic epilepsy	Offer the owner a blood test (biochemistry, haematology minimum) and advise monitoring at home for further seizure activity	Has no insurance and can spend a maximum of £75 during this visit
Weight loss and polydipsia	• 9-year-old dog • 6 month history of slow but progressive weight loss • No historical cause for weight loss • Observed drinking more water than usual latterly • Clinical exam normal	Diabetes mellitus/ Chronic kidney disease	Advise the owner to submit a urine sample for dipstick/specific gravity testing and recommend a blood test (biochemistry and haematology minimum)	Mother recently died from cancer so is extremely sensitive to the possibility of tumours

Table 2: Summary of the three cases developed for the standardised client simulation.

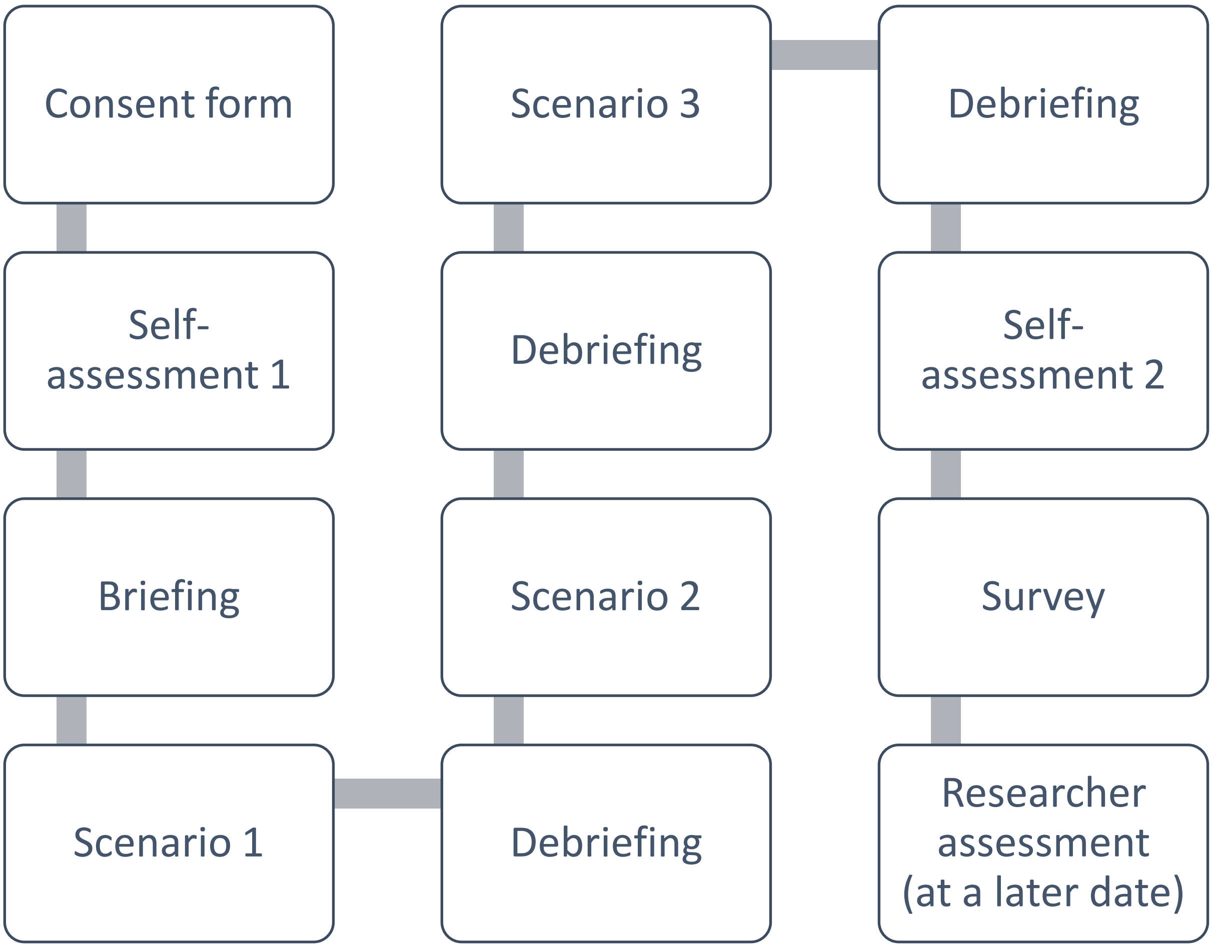

Figure 1: The overall simulation session process; repeated for each student

mLCJR Component	Median (mean) pre-sim score	Median (mean) post-sim score	Wilcoxon signed-ranks test statistic (Z score)	P-value
History taking	2.00 (2.47)	3.00 (3.13)	-4.36	<0.001*
Examination	2.00 (2.16)	3.00 (2.81)	-4.38	<0.001*
Identifying abnormalities	2.00 (2.03)	3.00 (2.75)	-4.23	<0.001*
Prioritising data	2.50 (2.53)	3.00 (2.78)	-2.14	0.033*
Making sense of data	2.00 (2.38)	3.00 (2.72)	-2.40	0.016*
Well planned intervention	2.00 (2.19)	3.00 (2.78)	-3.34	0.001*
Calm, confident manner	3.00 (2.59)	3.00 (3.03)	-3.13	0.002*
Clear explanations	3.00 (2.75)	3.00 (3.22)	-4.61	<0.001*
Total	20.50 (21.53)	25.00 (25.91)	-4.61	<0.001*

Table 3: Group A pre and post simulation self-assessment scores, with results of the Wilcoxon signed-ranks test to determine if the difference between pre/post-simulation self-assessment scores is statistically significant. *P-value shows a statistically signficant differnece (≤0.05)

mLCJR Component	Median (mean) pre-sim score	Median (mean) post-sim score	Wilcoxon signed-ranks test statistic (Z score)	P-value
History taking	3.00 (2.83)	3.00 (3.11)	-2.50	<0.012*
Examination	3.00 (2.60)	3.00 (2.77)	-1.90	0.057
Identifying abnormalities	2.00 (2.31)	3.00 (2.74)	-3.27	<0.001*
Prioritising data	3.00 (2.63)	3.00 (2.77)	-1.51	0.131
Making sense of data	3.00 (2.49)	3.00 (2.71)	-2.14	0.032*
Well planned intervention	2.00 (2.23)	3.00 (2.69)	-3.77	<0.001*
Calm, confident manner	3.00 (2.89)	3.00 (3.97)	-1.00	0.317
Clear explanations	3.00 (2.97)	3.00 (3.11)	-1.67	0.095
Total	23.00 (23.57)	26.00 (25.66)	-3.44	0.001*

Table 4: Group B pre and post simulation self-assessment scores, with results of the Wilcoxon signed-ranks test to determine if the difference between pre/post-simulation self-assessment scores is statistically significant. *P-value shows a statistically signficant differnece (≤0.05)

mLCJR Component	First consultation median (mean) score	Third consultation median (mean) score	Wilcoxon signed ranks test statistic (Z score)	P-value
History taking	2.00 (2.50)	3.00 (2.93)	-3.00	0.003*
Examination	4.00 (3.50)	4.00 (3.51)	-0.01	0.992
Identifying abnormalities	4.00 (3.75)	4.00 (3.55)	-1.57	0.116
Prioritising data	4.00 (3.60)	4.00 (3.61)	-0.12	0.906
Making sense of data	2.00 (2.75)	3.50 (3.13)	-2.16	0.031*
Well planned intervention	3.00 (2.85)	2.00 (2.84)	-0.49	0.625
Calm, confident manner	3.00 (3.13)	3.00 (3.09)	-0.41	0.684
Clear explanations	3.50 (3.38)	3.00 (3.18)	-1.90	0.058
Total	25.50 (25.47)	26.00 (25.84)	-0.50	0.619

Table 5: First/third simulated consultation scores according to the researcher-assessment, with results of the Wilcoxon signed ranks test to determine if the difference between first/third consultation researcher-assessment scores is statistically significant (P≤0.05). *P-value shows a statistically signficant difference (≤0.05)

Question	Group A n=32		Group B n=35		Mann-Whitney test statistic	P-value
	median (mean) score	Percentage agreement	median (mean) score	Percentage agreement		
The session was enjoyable	6.00 (5.53)	100.00	5.00 (5.17)	100.00	378.00	0.010*
The session was a good use of my time	6.00 (5.72)	100.00	6.00 (5.54)	100.00	463.50	0.144
I would like to participate in a session like this again	6.00 (5.69)	100.00	5.00 (5.17)	97.10	323.00	0.001*
My knowledge improved during the session	6.00 (5.53)	100.00	5.00 (4.97)	94.30	333.50	0.002*
My practical skills improved during the session	5.00 (4.72)	96.90	4.00 (4.29)	88.60	411.00	0.043*
My overall confidence in making decisions improved during the session	5.50 (5.41)	100.00	5.00 (5.03)	100.00	392.50	0.021*
My overall ability to reach a diagnosis has improved as a result of the session	5.00 (5.09)	100.00	5.00 (4.77)	100.00	420.00	0.049*
My overall ability to form a treatment plan has improved as a	5.00 (5.00)	100.00	5.00 (4.83)	100.00	491.00	0.341

result of the session						
I feel more prepared to undertake small animal consultations now	5.50 (5.41)	100.00	5.00 (5.20)	100.00	460.00	0.164
I found the session challenging	5.00 (5.03)	96.90	5.00 (4.61)	97.10	415.00	0.051
I found the session demoralising	1.00 (1.42)	0.00	2.00 (1.74)	0.00	405.00	0.030*
I found the session and scenarios unrealistic	1.00 (1.44)	6.20	2.00 (1.65)	2.90	429.00	0.060
I felt embarrassed participating in the session	1.00 (1.78)	15.60	2.00 (2.14)	20.00	435.50	0.092
The feedback sessions were informative	6.00 (5.87)	100.00	5.00 (5.31)	97.10	276.00	<0.001*
The feedback sessions were demoralising	1.00 (1.06)	0.00	1.00 (1.06)	0.00	338.00	<0.001*

Table 6: Median and mean average ratings for each survey question, percentage agreement with each questions and results of Mann-Whiney test to determine if groups A and B answered the survey differently. *P-value shows a statistically signficant difference (≤0.05)

Appendix 1

Component	Score			
	1	2	3	4
History taking	Is ineffective at taking a history. Obtains very limited information from the owner. *E.g. Only asks one or two of the mark sheet history questions.*	Asks SOME required questions, but misses a few important ones out. Seems unsure what information to ask for and may ask irrelevant questions. *e.g. Does not ask about water intake when faced with the weight loss case*	Asks MOST required questions, but occasionally does not follow up or clarify important leads. May miss one minor point, but asks all vital questions. *e.g. Does not ask about in-contact animals when faced with the D+ case*	Asks ALL relevant questions when taking a history. *e.g. Asks all questions on the mark sheet*
Examination	Examination is very limited, only one or two components are checked. *e.g. Only auscultates chest*	Performs a LIMITED clinical examination. Important aspects of the exam are missed out. *e.g. Does not perform any neurological examination when faced with the seizing case*	Performs a THOROUGH clinical examination; a few minor components are missed. *e.g. Does not check lymph nodes on any case*	Performs a COMPLETE clinical examination, does not miss any components relevant to the case. *e.g. Completes all points on the mark scheme*
Identifying abnormalities	Misses the importance of clinical findings – unjustly dismisses them. *e.g. Not appreciating significant weight loss that requires investigation in the weight loss case*	Recognises SOME abnormalities, but overlooks some important findings from the history/exam. *e.g. Not noting polydipsia when faced with the weight loss case*	Recognises MOST abnormalities that need to be considered, missing only minor aspects. *e.g. Not noting lethargy in the diarrhoea case*	Recognises ALL problems that need to be addressed. *e.g. Identifies all relevant abnormalities*

Prioritising data	Does not know which findings to concentrate on, prioritises an unimportant problem over the relevant issue – may not attend to the main problem. *e.g. Focusing on lack of flea treatment at length during the weight loss case*	Attempts to focus on the main problem, but gets distracted. Alternatively, does not prioritise relevant findings as important. *e.g. Does not prioritise polydipsia as a problem when discovered in history of the weight loss case*	Generally concentrates on the most important findings, but does talk about irrelevant aspects of the exam/history BRIEFLY. *e.g. Recommending worming when faced with the acute D+ case (except as general recommendation to worm regularly)*	Just discusses and forms a treatment plan for the relevant findings. *e.g. Only discusses aspects directly related to the current problem*
Making sense of data	Struggles to interpret history and exam findings. Is unsure how to proceed. Does not determine a feasible way to proceed with the case. *e.g. Sends owner of weight loss case home with view to monitor weight over coming months*	Attempts to interpret the clinical findings, but misses an IMPORTANT differential diagnoses or includes irrelevant ones. *e.g. Does not consider toxin ingestion when facing seizing case*	Is able to interpret the history and clinical exam to form several differential diagnoses, but may miss a MINOR differential or include a differential that is very low in likelihood. *e.g. Considers worm infestation a differential for acute D+*	Is able to interpret the history and clinical exam to form a set of accurate differential diagnoses. *e.g. Clearly has considered all relevant differential diagnoses when deciding how to proceed with case*
Well planned intervention	Treatment plan is not acceptable treatment for the case. *e.g. Prescribing antibiotics when facing acute D+ case*	Treatment/investigation is not the most appropriate for the case, but some aspects are correct and will aid diagnosis/treatment. *e.g. Not conducting urinalysis on patient with*	Treatment/investigation plan is correct for the case, but there may be minor, aspects missed or incorrectly included. *e.g. Not advising Prokolin for acute D+ case*	Treatment choice ideal for case (considering animal and owner factors). *e.g. Follows treatment plan on mark sheet*

		PUPD but performing blood test		
Calm, confident manner	Is visibly stressed/anxious and lacks confidence. Relies on client to make decisions and direct consultation. *e.g. Long silences and obvious uncertainty when deciding on treatment plan*	Is tentative in the leader role; redirects some responsibility for decision making to the client. Moments of self-doubt, not 100% sure of treatment plan. *e.g. Offers treatment options but does not direct client/make recommendation – client decides how to proceed*	Is calm and confident in MOST situations. Directs the consultation but occasionally is unsure. *e.g. Changes mind about recommendations mid-consultation but otherwise confident and assumes responsibility for decision making*	Assumes responsibility; is confident with diagnosis/treatment plan. *e.g. Decides a treatment plan and relays this confidently to client*
Clear Explanation	Explanations are confusing and directions are unclear or contradictory. Owners are confused. *e.g. Owner cannot make sense of instructions given*	Explanations are mostly clear, though one element may cause confusion for the owner and need to be clarified. *E.g. Does not explain opinions clearly, owner has to ask questions to clarify*	Explains carefully to clients and gives clear directions. The pace/tone may be inappropriate or may not check for owner understanding. *e.g. Explains plan well but speaks too quickly*	Communicates at good pace; explains interventions clearly; checks for understanding. *e.g. Explains plan at appropriate speed, clearly and checks for owner comprehension*

662 Appendix 1: The modified Lasater Clinical Judgement Rubric