

Research Article

Undergraduates' Criteria to Justify Claims Proposed after Laboratory Experiments

Jer-Yann Lin¹ and Ding-Ying Guo²

¹Department of Materials Science, National University of Tainan, Tainan 700, Taiwan

²Department of Education, National University of Tainan, Tainan 700, Taiwan

Correspondence should be addressed to Jer-Yann Lin, yann@mail.nutn.edu.tw

Received 8 September 2010; Revised 2 November 2010; Accepted 7 December 2010

Academic Editor: Jan Elen

Copyright © 2011 J.-Y. Lin and D.-Y. Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study investigates the criteria undergraduates adopt to justify their claims proposed after laboratory experiments. There are two categories of justifications in the current literature, *empirical consistency* and *plausibility of claims*, but observations of college juniors in the laboratory demonstrated the need for a third category, *observation reliability*. This assumption was further tested by analyzing the warrants undergraduates wrote to justify their claims formed after laboratory experiments. Three students' justification criteria were identified, that is, *empirical consistency*, *plausibility of claims*, and *observation reliability*. The most frequently used criterion is *plausibility of claims* to justify *good results*, while that is *observation reliability* to justify *bad results*. Moreover, multiple justification, which means more than one attempt being made to justify a given claim, was also found. It reveals that multiple justification, rather than single justification, is suitable for students to make scientifically acceptable claims. The implications and suggestions of this study are also discussed.

1. Introduction

Justification is an attempt to convey and convince oneself or others about one's claim being correct or valid [1, 2]. According to this definition, students' justification occurs in two situations, that is, convincing either themselves or others. Both situations occur frequently in science education. A student who attempts to judge claims concerning the validity or interpretation of experimental data observed in the laboratory is a typical example of the former situation. On the other hand, the social construction of knowledge and cooperative learning in the laboratory are examples of the latter situation. In both situations, justification takes place when students attempt to convince themselves, their peers, and teachers about their claims being correct. Studies have suggested that students need to learn how to generate good arguments with true, reliable, and multiple justifications [3, 4].

The meaning of justification is different from that of logic. Logic uses the rules of inference from the premise to the claim. However, justification is the epistemic guide of an individual to justify the adequacy of a claim [5]. Logical

consistency may be an epistemic guide for an individual to justify a particular claim from evidence. However, an individual does not always justify a given claim based on logical consistency. Personal perspectives may involve in their justification of claims. For example, Samaratungavan [6] found that children could use the standards of logical consistency to choose a theory when the theory did not violate their beliefs. If a given theory violates the children's belief, they make a claim to reject this theory. Furthermore, some students even reject anomalous data, which cannot be illustrated by students' initial concepts, without changing their concepts [7–9]. Obviously, there are some epistemic guides, other than logic, that dominate the individuals' justification of their claims. Identifying these epistemic guides is an important theme in investigating the justification of claims.

Laboratory activities have long had a central role in science learning, and many benefits accrue from engaging students in science laboratory activities [10, 11]. Students in laboratories are involved in the process of conceiving problems and scientific questions, formulating hypotheses, designing experiments, gathering and analyzing data,

TABLE 1: Possible criteria undergraduates adopt to justify their claims proposed after laboratory experiment.

Criteria	Grounds ^a	Standards ^b
Empirical consistency	Empirical evidence ^c	Evidence should logically cohere with claims
Plausibility of claims	Prior knowledge, beliefs, or values	Claims should coherently relate to personal prior knowledge, beliefs, or values
Observation reliability	Observation process	Observation process should be reliable

^aGrounds: the events or concepts used to justify a claim

^bStandards: standards to judge the adequacy of a claim

^cEmpirical evidence: the data or information obtained from experiments, visual observation, or personal experiences.

and drawing conclusions about scientific phenomena [10]. Consequently, laboratory activities have the potential to enhance students' conceptual understanding and knowledge construction. However, the relationship between laboratory experiences and student learning is complicated. Several researches indicated that effective learning could not be achieved when students just involve in technical manipulation without any cognitive activities, such as predict-explain-observe learning process [11–13], presentation of results [14], or reflection and modification of ideas [15].

The need of cognitive activities in laboratory learning demonstrates that the justification of claims might play an important role for students to learn effectively in the laboratory. Students have to justify their assertions formed during these cognitive activities. Predicting or presenting experimental results, for example, students need to convince themselves or their peers that their claim is valid or believable. The ways students justify their claim, referred to as students' justification criteria, influence students' judgment upon experimental phenomena as well as their learning in scientific concepts. For example, Chinn and Brewer [8] indicated that students might reject the correct data that were unexplainable by students' concepts. Thus the criterion students adopt to justify claims seems to be a vital factor that influences their learning in the laboratory. Better understanding of students' justification criteria is important for instructors to enhance students' learning in the laboratory.

2. Literature Review

There are two categories of justification criteria in the current literature. Hogan and Maglienti [16] reported students adopted two justification criteria, *empirical consistency* and *plausibility of claims* to justify conclusions from a given set of evidence. Both justification criteria are defined in Table 1, and illustrated in the following paragraphs.

Adopting *empirical consistency*, subjects justify their claim by indicating that the claim or its supporting data were consistent with empirical evidence obtained from scientific experiments or daily life observation. *Empirical*

consistency is a predominate criterion scientists adopt to justify a given claim based on the coherence between claims and the corresponding evidence. However, several studies revealed that students also adopt empirical consistency to justify their claims. Hogan and Maglienti [16] indicated that middle school students sometimes applied the criterion of empirical consistency to justify knowledge assertions from a set of given data. By the method of questionnaire survey, Zou [17] investigated how undergraduates convinced their peers that a given scientific knowledge is correct. Results of Zou's study showed that seventy percent of the participating students justified the scientific knowledge based on empirical consistency. Samarapungavan [6] showed children some pictures as scientific evidence and found that children used the criterion of logical consistency with these evidence to choose a theory when the theory did not violate their beliefs. BouJaoude [18] demonstrated experiments to students and indicated that students' assertions on whether the weight of a candle after burning is lost or not were based on their visual observation. Galili and Bar [19] demonstrated burning-candle experiments to students and reported that 67% of the non-conservers they interviewed logically connected their assertion to the empirical evidence they observed. Using individual interviews, researchers also reported students' explanations concerning the organism's role in ecosystem [20] or the earth's gravity [21] were quite consistent with the empirical evidence students observed in daily life. In the context of socioscientific decision making, consistency with personal experiences of students plays an important role in forming an assertion on the socioscientific issues of genetic engineering [22] and distinct environmental challenges [23].

Adopting *Plausibility of claims*, students justify claims based on whether claims are coherent with a given theory or belief. When students adopted this criterion, they justified conclusions by their belief about how the system works: such that the conclusions would be plausible relative to their own ideas or values [16]. Several studies reported in the literature also revealed that students used their preferred theory to justify their claims. Based on the literature review, Kunda [24] proposed that an individual typically seeks confirmation for the conclusion formed from evidence, compatible with one's prior belief. Such an argument, referred to as *prior belief effect* [25], was also suggested by several researchers [26, 27]. In peer discussions of bioethical issues, students' core beliefs appeared throughout the student's chain of reasoning [28]. Sadler [29] explored how college students' conceptions of evolutionary theory affect their reasoning in the context of genetic engineering issues. Many biology majors they interviewed adopted evolutionary perspectives to justify their views. Applying content knowledge to generate arguments also indicated that some students justify their assertion by conveying the consistency of their assertion with a particular scientific theory [3, 4].

3. Assumptions and Purposes of This Study

This study focuses on students' justifications in the laboratories. The situation of students' justifications after authentic

laboratory experiments is different from that of studies based on questionnaire survey [16], individual interview [6, 18–23, 29], peer discussion [28], or students' writing [3, 4, 17]. The data are produced by students themselves in the lab. In the previous studies, however, the data/evidence were from researchers' delivery [6, 16], researchers' demonstrations [18, 19], or students' thinking [3, 4, 20–23, 28, 29]. Students may doubt the validity of data they produce owing to personal manipulative error, but students seldom argue the validity of data/evidence from researchers' delivery or demonstration, or from authorized sources, such as textbooks or journal articles. An additional criterion, *observation reliability* (defined in Table 1), is possible in the laboratory by indicating the observation reliability of the process producing a claim or its supporting evidence to justify a claim. It seems possible that three criteria, *empirical consistency*, *plausibility of claims*, and *observation reliability*, may be students' criteria to justify claims proposed after laboratory experiment. Table 1 shows the comparison of three justification criteria based on the *grounds* and *standards* of justifications. The purpose of this study is to investigate undergraduates' justification criteria after laboratory experiments. The research questions are as follow: what types of justification criteria are there in undergraduates' justifications? What characteristics are there in undergraduates' justifications of a given claim by more than one attempt?

4. Methods and Procedure

This study consisted of 161 participants who were undergraduates in their junior year. The participants were education majors, seventy-four were social science education majors, and eighty-seven were primary education majors. They participated in an experimental course for credit, which was a part of the education curriculum for preservice elementary school teachers. Students in their sophomore year had completed a semester course of general chemistry in which they learned the scientific theories essential for each experimental unit in this course. The principles of each experiment were given in the handouts instructor gave, and students had to complete their reading before they entered the laboratory. In the laboratory, instructors illustrated briefly the experimental principles as well as the laboratory safety of each experiment before students' experimental activities.

There were seventeen units of experiments in this course, and three experiment units were selected for the analysis in this study, that is, (a) melting point and crystallization, (b) simple distillation, and (c) electrochemistry. Seven prompts were used as a template to assist student to learn throughout the experimental activities as well as to complete their report writing, that is, *observations*, *investigation question*, *methods*, *claims*, *evidence*, *negotiations*, and *awareness of change* [30, 31]. To explore each experiment unit in the laboratory, students carried out two phases of lab activities within two hours. In the first phase, students did experiments according to the experimental procedures described in the handouts instructor provided [32–34]. Instructor encouraged students

to find out problems or questions associated with the experimental phenomenon they *observed* when they carried out the experiment. The problems or questions students found out became the *investigation questions* they explored in the subsequent second phase of lab activity. In the second phase, students were prompted to design adequate *methods* to explore the investigation questions they generated in the previous phase. After evidence collection, students made claims and discussed with peers (*negotiation*) the relationship between *claims* and *evidence*. Each group of students had to explore at least one investigation question in the second phase of experiment. To obtain students' possible justifications of claims upon their investigation questions, instructor did not confine students' investigation methods to a given researching approach, that is, either experimental induction or theoretical deduction was acceptable.

Experiments were carried out by groups of students. Students could choose their group members to form a group, and a group ordinarily contained three students. If one or two individual(s) left after three-member group formation of a class, there was a group in this class containing four or two students, respectively. All experiments in this study were carried out by each group of students, containing 2–4 members, except that an investigation question in experiment (c) was explored cooperatively by two groups of students. In this instance, a group of students doubted the instrument they used was broken out in the first phase experiment and proposed that intergroup cooperation was required to explore the investigation question they generated. Thus, there were six students cooperatively doing this investigation question in this case.

Lab works were performed by groups of students; however, students had to complete their lab reports individually. In their reports, students described the phenomena *observed*, the *investigation question* generated, the methods designed, the *claims* proposed, and the corresponding *evidence* demonstrated as well as the *reflection* of their ideas. The instructor encouraged students to give sufficient reasons to justify their conclusions in their reports. To collect students' justifications of any possible claims made after experiments, the instructor did not limit students' judgment in conclusion making. Students can decide to accept, doubt, or reject their experimental results even though the results were factually incorrect according to current scientific knowledge. In students' reports, conclusions of their investigation questions as well as the justification of claims were sorted out for the analysis in this study. The justification students wrote were analyzed to understand how they justified their claims after laboratory experiments.

4.1. Coding Procedures. Although, possible justification criteria were proposed in the previous section, the codes were created inductively. Students' written reports were collected as protocols in this study, and each report was given a report code indicating the experiment unit of report and the background of the student writing this report. For example, the report code of "08-6-a-114-b" indicates this is a report of exp. unit 08 written by a student who was in class 6,

TABLE 2: Number of scientific claims proposed and the corresponding justifications by undergraduates.

Experiments*	Claims	Justifications
A	185	333
B	179	234
C	183	252
Total	547	818

* A: melting point and crystallization; B: simple distillation; C: electrochemistry.

participated in group A, had a student ID of 114, and is a *female*. To obtain the criteria undergraduates adopted to justify their claims, the protocols were coded, involving four steps: (1) sorting out the claims and the corresponding justifications, (2) identifying the grounds of justification, (3) identifying the standards of justification, and (4) naming each criterion of justification.

Sorting out Claims and the Corresponding Justifications. The conclusions written in students' experiment reports were identified. Each attempt to justify a given claim was identified as a justification of this claim. If a student justified a given claim based on more than one attempt, all the attempts were identified as the justifications of this claim. The results of extraction showed that 547 claims and the corresponding 818 justifications were sorted out as the protocols in this study (Table 2).

Identifying the Ground of a Justification. The question, "what is the ground used to justify a claim?" was used to test each claim in the protocols, and the answer was the code of ground for the justification. If students described that their claim was correct because it obeys the scientifically acceptable theory, the code of ground was "*scientifically acceptable theory*." Eight types of codes for the grounds of justifications were found in this analysis: *experimental manipulation*, *experimental apparatus*, *experimental method*, *literature value*, *daily life experience*, *scientifically acceptable theory*, *naïve theory*, and *theoretically calculated value*.

Identifying the Standard of a Justification. The question, "what is the standard used to judge the validity of a given conclusion?" was used to test each claim. If students justified a claim based on such a claim being consistent (or inconsistent) with a given scientific theory, the code of standard was *consistency*. If students justified a claim by describing that the manipulative process to produce the claim was careful (or error), the codes for the standard of these claims were *careful* and *error*, respectively. Three types of codes, *consistency*, *careful*, and *error* were found for the standard of justifications.

Naming each Criterion of Justification. *Ground* and *Standard*, identified in previous coding procedure, for a given justification were combined to form a set of codes, *Ground-Standard*.

For example, a student's justification was identified in previous coding procedures to contain a *ground code*, *scientifically acceptable theory*, and a *standard code*, *consistency*. The set of codes for this justification was *scientifically acceptable theory consistency*. The results showed that there were nine sets of codes: *experimental manipulation-careful*, *experimental manipulation-error*, *experimental method-error*, *experimental apparatus-error*, *literature value-consistency*, *daily life experience-consistency*, *scientifically acceptable theory-consistency*, *naïve theory-consistency*, *theoretically calculated value-consistency*. The 1st–4th, 5th–6th, and 7th–9th sets of codes were named as *observation reliability*, *empirical consistency*, and *plausibility of claims*, respectively. Both authors coded all protocols, with 90% agreement on the protocols coded by both coders. The coders had had different views for classifying the instances of justifications and for naming of the categories such as which code, *quality* or *reliability*, was suitable to label the standard for justifying claims by the experimental manipulation. Both coders examined and discussed the situations for experimental manipulation. After discussion, both coders achieved agreement.

5. Results and Discussion

The justifications of claims proposed by students after laboratory experiments were analyzed, and the results are shown in Table 2. Some students in this study made more than one attempt to justify a given claim. Each attempt was identified as a justification. For example, a student asserted that the melting point observed for the sample of benzoic acid was 118–120°C because the experiment was not only carefully measured but the datum was also close to the literature value. The student in this instance made two attempts, which were data being carefully measured and data being close to the literature value, to justify his/her claim. Both attempts were identified as the justifications of the claim. Thus, the number of justifications identified was higher than that of claims proposed.

5.1. Students' Justification Criteria. All the 818 justifications could be classified into three categories of criteria (Table 3), that is, *empirical consistency*, *plausibility of claims*, and *observation reliability*.

5.1.1. Empirical Consistency. Undergraduates also justified their claim by indicating that the claim or its supporting data were consistent with empirical evidence obtained from scientific experiments or daily life observation. Unlike the criterion of *observation reliability* that focuses on the quality of a given experimental procedure, the criterion of *empirical consistency* focuses on the consistency of a claim with the corresponding empirical evidence. Empirical evidence is the result of experimental measurement or daily life observation, rather than direct deduction based on theory. Two types of empirical evidence were found by undergraduates in the laboratory to justify their claim, that is, literature reports and personal experiences. Both consistency and inconsistency

TABLE 3: Categories and subcategories of undergraduates' justifications of scientific claims proposed in the laboratory.

Category	Subcategory	f*
(1) Empirical consistency		
	(1.1) Literature value	82
	(1.2) Personal experience	31
(2) Plausibility of claims		
	(2.1) Scientifically acceptable theory	81
	(2.2) Naïve theory	11
	(2.3) Theoretically calculated value	55
(3) Observation reliability		
	(3.1) Careful manipulation	52
	(3.2) Personal error	397
	(3.3) Apparatus error	86
	(3.4) Method error	23

*f: frequency

with empirical evidence were used by undergraduates to justify their claims.

Literature Report. Literature reports are the experimental evidence reported in the literature, such as handbooks, textbooks, or journal article, and contain the results of most undergraduate laboratory experiments. Some students in laboratory justified their claim by the consistency of their claims with the corresponding literature values.

Claim. The experimental value is correct.

Justification. This piece of datum is almost identical to the literature value (122.4°C) reported in the literature

Experience Datum. Observed melting point of sample (benzoic acid): 119–121°C (086a114b)

By comparing the literature value and the experimental value, the student asserted his/her experimental value as correct and justified his/her claim by the consistency of both values.

Personal Experience. The personal experiences of students are the students' experimental or daily life experiences observed before laboratory experiments. Unlike the literature value subcategory that is empirically produced and reported in literature, personal experiences are the observation of students themselves.

Claim. We obtained high quality ethanol by distillation.

Justification. I have smelled a bottle of ethyl alcohol before. The odor of the distillate was identical to that of ethyl alcohol. (106h112a)

In this instance, the odor of the distillate smelled by the student was identical to that of ethanol he/she experienced.

The distillate was justified to be ethyl alcohol by the personal experience of the student.

5.1.2. Plausibility of Claims. Undergraduates also justified their claims by convincing that their assertions were consistent with a particular scientific theory. Unlike the criterion of *empirical consistency* that focuses on the comparison with empirical evidence, the criterion of *plausibility of claims* focuses on the comparison with scientific theory. Three types of scientific theories, which were scientifically acceptable theory, theoretically calculated value, and naïve theory, were adopted by undergraduates to justify their assertions.

Scientifically Acceptable Theory. Undergraduates justified their assertions by indicating that the assertion was consistent with a particular scientifically acceptable theory. Undergraduates also justified their assertion by indicating that the experimental data were inconsistent with a scientifically acceptable theory. For example, a student justified the claim that the volume ratio of the distillate and residue should be constant by the scientifically acceptable molecular model of evaporation and condensation.

Claim. The volume ratio of the distillate and the residue in the flask is a constant for a given mixture separated by distillation.

Justification. The principle of the distillation could be illustrated by the evaporation and condensation processes. In a given temperature, a particular number of molecules in the liquid evaporate into vapor. On the other hand, a particular number of molecules in the vapor condense into the liquid. The component with a lower boiling point becomes the distillate, and the component with a higher boiling point would become the residue in the flask in a given distillation temperature. Thus, the volume ratio is constant for a given mixture (106h107a).

The equilibrium between condensation and evaporation during the distillation process is a scientifically acceptable theory. Although molecules are microscopic and not observable in the undergraduate's laboratory, some undergraduates used such a theory to justify their claims.

Naïve Theory. Not all scientific theory students used in this study belonged to scientifically acceptable theory, however. The personal naïve theory, scientifically unacceptable, was also found to justify assertions.

Claim. The melting point of the sample (benzoic acid): 117–118°C

Justification. The particle size of the sample used in our experiment was larger. The lower melting point of this sample, compared to the literature value, results from its smaller surface area to absorb heat (085f024b).

If the samples are not nanoparticles, the theory that the sample with larger particle size has lower melting point based

on smaller surface area is scientifically unacceptable. All the samples used in this experiment were not nanoparticles although the samples had nonuniform particle size. Thus, the melting point of sample was not affected by its particle size. In this instance, the scientific theory this student adopted to justify his/her assertion was in fact a personal naïve theory.

Theoretically Calculated Value. The scientifically calculated value is the value obtained by mathematic calculation based on scientifically acceptable theory. A student, for instance, adopted this criterion to justify his/her claim.

Claim. The observed gas volume ratio for electrolysis of water was 1 : 2.

Justification. The Experience entry 2 was a failed experiment because it is significantly different from the calculated value according to the reaction stoichiometry. On the other hand, the Exp. entry 1 is consistent with the calculated value.

Experience data. Gas volume obtained from the electrolysis of water: Exp. entry 1, 1 : 2; Exp. entry 2, 2 : 3 (142a078b).

The volume ratio of the gases obtained from the electrolytic reaction of water could be calculated based on the reaction stoichiometry. This calculated value was used by student to justify the assertion of his/her observed datum.

5.1.3. Observation Reliability. Students attempted to justify their claim by indicating the reliability of observation to produce a claim or its supporting evidence. The criterion of observation reliability was found most frequently among three categories (Table 3). Students evaluated various types of manipulations of observation, for example, the manipulative methods, apparatus, or personal techniques. Both good and poor qualities of observation were adopted by students to justify their claims. There are four subcategories identified in the category of *observation reliability*, that is, *careful manipulation*, *personal error*, *apparatus error*, and *method error*.

Careful Manipulation. Students justified a particular claim by directly indicating that this claims or its supporting evidence were produced by careful manipulation. Observation reliability, such as careful measurement, and critical control of experimental conditions became the warrants of students to justify their claims. For example, a student in the laboratory asserted correctness of his/her experiment by conveying that he/she critically controlled the experiments.

Claim. The melting point of the sample is 106–110°C.

Justification. The result was correct because the experiments were carried out carefully. The sample tube was heated slowly (082k088b).

To obtain an accurate experimental value of the melting point for a sample, it is important to heat the sample tube slowly. The student in this instance asserted that the experimental result was correct by conveying that he/she carefully controlled the heating temperature of the sample tube.

Personal Error. Unlike *careful manipulation*, students may assert that a particular experimental result was wrong by indicating that personal manipulative errors occurred in a given procedure of experiments. For example, a student asserted that his/her recovery of crystal was too low because of crystal loss resulting from personal manipulative error.

Claim. The recovery of crystal in this experiment was too low.

Justification. Some crystals spilled from the funnel when the suspended solution was filtrated. Without such an operation error, the recovery should be higher.

Experience Datum. Recovery of crystal = 60% (082f064b).

In this instance, the loss of crystal is one of the factors that reduce recovery of crystal. Conceiving such a manipulation error, the student ascribed the poor experimental result to such a manipulation error. Personal error was most frequently proposed by students to justify their claims (Table 3). In the laboratory, undergraduates typically had insufficient confidence to perceive their personal manipulation as valid because they had no experience with such manipulations before the experiment.

Apparatus Error. The quality of experimental apparatus was another subcategory for undergraduates to justify their claims.

Claim. The datum for the volume ratio of gases collected from the electrolysis of water is wrong.

Justification. The volume scales of our tubes blanked out. It became difficult to read the exact volume of the gas collected in our tube.

Exp. Datum. Gas volume, H₂:O₂ = 2.5:1 (145e035b).

Students collected oxygen and hydrogen gases produced by the electrolysis of water separately in two test tubes. The students observed that the volume scales drawn in the test tubes blanked out, and justified his/her claim by correlating this observation with the corresponding datum.

Method Error. The experimental method was a subcategory to argue the reliability of the data producing process. Unlike the personal error or apparatus error, the method error means that experimental procedure is invalid. For example, a student reflecting on his/her experimental procedure found that errors existed in the manipulative procedures described in the textbook. The student asserted that such an experiment was wrong owing to method error.

Claim. Our experiment is wrong.

Justification. Because the method of oven dry method was not used in this experiment, water still remained in the product. This resulted in an experimental error.

Experience Datum. The recovery of crystals was 60% (082f089b).

The solid crystal obtained from the aqueous solution could be dried by different methods, such as oven dry, vacuum suction, or drying agent. The student depicted a deficiency, without oven dry, to justify his/her claim of the crystal recovery from aqueous solution being too high.

Several studies reported that certain scientific reasoning processes of students are similar to those of scientists [6, 7, 35–37]. In scientific journal articles, scientists frequently adopt *observation reliability* to justify their claims [38, 39]. Poor observation reliability becomes a criterion to assert that a given datum is invalid. For instance, McCrone [39] asserted that the data of the titanium content in pigment particle produced by particle-induced X-ray emission (PIXE) are invalid because the instrument, PIXE, is not a good choice for detecting subnanogram samples. The finding that students participated in this study adopt justification criterion of *observation reliability* seems reasonable.

The finding that the criterion, *observation reliability*, should add to the justification criteria previously reported [16] may ascribe to the learning environment of laboratory. The data were given by researchers in previous study [16], while those were produced by students in the laboratory. Students seldom argue the data validity when data source was authorities, but students often doubt the data validity in the laboratory [40]. Similar to scientists' facing anomaly in the laboratory, students in the laboratory frequently encounter new experimental phenomenon they havenot observed. Some students probably correlate such a phenomenon to manipulative processes, and this correlation is similar to scientists' correlating invalid data to improper instruments used [39].

The finding of three students' justification criteria seems rational according to epistemic theories of justification. The criterion of *plausibility of claims* is similar to the epistemic theory of *internalistic foundationism*, in which the beliefs in question are judged in terms of whether they relate well with the foundation or not [5]. The criterion of *observation reliability* is similar to the theory of *reliabilism*, in which a belief is justified if the belief-forming process to produce this belief is reliable [41]. On the other hand, the criterion of *empirical consistency* is similar to *empiricism*, that a belief is justified only if this belief is causally related to the fact making the belief true [42, 43]. Three justification criteria represent three distinct dimensions of epistemic guides a student adopt to be justified in believing a claim.

Two types of claims, *good results* and *bad results*, were identified in this study. The claims in which students asserted a given experimental result or concept is valid or correct are defined as *good results*. On the other hand, the claims in which students asserted a given experimental result or concept is invalid or false are defined as *bad results*. Among the 818 justifications, most justifications (79%, 646 instances) were adopted to justify *bad results*, such as invalid data or false concepts, they concluded after exploring their investigation questions. Only 174 justifications (21%) were adopted to justify *good results*, such as valid data or rational concepts. As described in previous section, investigation questions were generated by students in the first phase of lab activities, and most

investigation questions students explored in this study were the interpretation or judgment of anomalous phenomenon. Either scientists or students usually concluded anomaly to be a *bad result* [8]. Several reports indicated that students usually disbelieve and discount anomalous data [8, 40].

As shown in Figure 1, the behaviors students adopted as justification criteria to justify *good results* and *bad results* they claimed are quite different. The most frequently used criterion is *plausibility of claim* to justify *good results*, while that is *observation reliability* to justify *bad results*. Near a half (48%) of the justifications students adopted to justify *good results* was *plausibility of claims*; that is, results were good because of consistence with *scientific theory*, personal *naïve theory*, or *theoretical value* (ST, NT, and TV in Figure 1(a)). On the other hand, 78% of the justifications to justify *bad results* was *observation reliability*; that is, results were bad owing to *personal error*, *apparatus error*, or *method error* (PR, AR, and MR in Figure 1(b)).

In some aspects, the context of justifying data or concepts as *good results* in lab is similar to that of justifying a given conclusion based on evidence provided by Hogan and Maglienti [16]. Both contexts, students need to illustrate how good the data, concepts, or conclusions are. Both studies show that students frequently adopt *plausibility of claims* to justify how claims are good. However, justifications in lab allow students repeatedly doing experiments to explore a certain anomaly student doubt, while, repeating such experiments was not provided in the previous study [16]. Repeating experiments and subsequently confirming manipulative reliability in laboratories, students have confidence to justify a given claim as a good one. As a result, *empirical consistency* and *observation reliability* are also vital in justifying *good results* in this study (Figure 1(a)). However, the former criterion was seldom used, and the latter was not adopted in the previous study [16].

Unlike justifying *good results*, *observation reliability* is the criteria students adopted the most to justify *bad results*; both *plausibility of claims* and *empirical consistency* are seldom used. Among the subcategories of observation reliability, *personal error* seems a useful criterion to justify anomaly students encountered as a *bad result*. This finding probably results from that students who are novices at experimental manipulation in the lab. Consequently, students have insufficient confidence to assert anomaly as a *good result* based on their manipulation. Chinn and Malhotra [44] found that students can change their belief based on their observation, but they have difficulties to make accurate and precious observation. The finding of this study also implies that personal manipulation is vital in students' learning in the lab.

It was also identified that some *bad results* students adopted personal errors to justify were actually *good results*. In experiment (c), for example, students had to observe the temperature at distillation head every minute, and interpret the temperature change during distillation period. Some students actually obtained *good results*, but they made conclusion of *bad results* and adopted *personal error* to justify their claims.

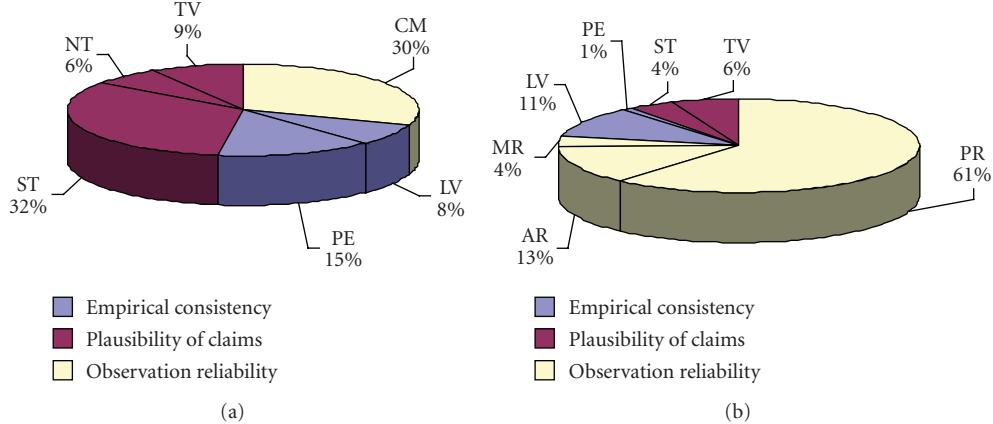


FIGURE 1: The percentage of each justification criterion students adopted to justify (a) *good results* or (b) *bad results*. AR: apparatus error; CM: careful manipulation; LV: literature value; MR: method error; NT: naive theory; PE: personal experience; PR: personal error; ST: scientifically acceptable theory; TV: theoretically calculated value.

Claim. At the start of distillation, the temperature recorded is error.

Justification. At the start of distillation, solution was boiling, but the temperature observed at distillation head was still room temperature. We checked our apparatus and found we had an error in setting the cooling-water inlet. After correcting this error, we found that the temperature increased (106b123b).

The purpose of this activity is to observe and learn several phenomena during distillation period, such as refluxing and condensation ring. At the start of distillation, the mixture was brought just to boiling and refluxing, but the distillation head would be indeed at room temperature. After the condensation ring traveled up and arrived at the distillation head, the temperature of distillation head would gradually increase. The phenomenon, boiling solution with lower temperature at distillation head, observed at the start of distillation is a *good result*; however, it is anomalous for some students. They claimed the anomaly is a *bad result* and associated some personal manipulative error with this result. They tried to do something to correct *personal errors* they thought in order to make the temperature increase, such as changing cooling-water inlet or checking vapor leaking. Probably because the condensation ring had already arrived at the distillation head after they completed their error correction, the temperature indeed increased. Such an observed fact made students misrelate the *personal error* they thought to their claim of a *bad result*.

5.2. Multiple Justification. Some students justified a given claim by more than one warrant, as shown in Table 4. The multiple justifications could further be classified according to the similarity of the attempts in a given multiple justification, and *homotype* and *heterotype* were obtained. The homotype multiple-justification consists of the justifications with identical criterion; for example, a student justifies a claim by *personal error* and *apparatus error*, both belonging to the criterion of *observation reliability*. On the other hand,

TABLE 4: Comparison between multiple justification and single justification.

Category	N _c ^a	N _{sci} ^b	N _{sci} /N _c (%)
Multiple-justification ^c	213	113	53
Homotype ^d	76	39	51
Heterotype ^e	136	74	54
Single-justification ^f	334	159	48

^aN_c: the number of the total claims justified by a certain category of justification

^bN_{sci}: the number of scientifically acceptable claims

^cMultiple justification: more than one attempt is made to justify a given claim

^dHomotype: a multiple justification with identical criterion of justifications

^eHeterotype: a multiple justification with different criteria of justifications

^fSingle justification: one attempt is made to justify a given claim.

the heterotype consists of different criteria of justifications to justify a given claim; for example, a student justifies a claim by *careful manipulation* and *literature report*, which belong to the criteria of *observation reliability* and *empirical consistency*, respectively.

This study reveals that multiple justification is suitable for students to form scientifically acceptable claims. Table 4 shows the number ratio (N_{sci}/N_c), which is the ratio of the scientifically acceptable claims to the total claims proposed, of multiple justification as well as single justification. No significant difference in number ratio was found between both types of justifications. However, further qualitative analysis revealed that in some instances students could not achieve scientifically acceptable claims if students just adopted single justification, instead of multiple justification. For example, a student who misunderstood the mass conservation law doubted a valid datum he obtained from his initial experiment. After confirming experiment and reconsidering his justification ground, the student made a scientifically acceptable claim, finally.

Claim. The datum obtained is correct.

Justification. In the initial experiment, the mass of crystal obtained was lower than that of the substance in mixture we wanted to purify. We doubted the result because no mass should be lost during this physical change according to mass conservation law. We repeated experiment carefully and obtained identical datum. After group discussion, we found that part of the substance remained in solution after crystallization; we misunderstood the mass conservation law.

Experience Datum. The recovery of crystals: 78% (083k005a).

The student in this instance adopted two types of justifications, *plausibility of claim* as well as *observation reliability*. At the beginning of his investigation, he had a poor justification ground owing to misunderstanding of the mass conservation law. If the student made a claim just based on this poor justification ground without further investigation, he would not form a scientifically acceptable claim. Students are novices at either applying scientific theory or manipulating experiments. Multiple justification makes students reconsider their thinking and/or observation process, and it is suitable for them to form scientifically acceptable claims.

The finding that multiple justification is suitable to make scientifically acceptable claims seems reasonable because different justification criteria are associated with distinct types of students' nonscientific thinking. In our recent study [45], students' nonscientific thinking occurs in three dimensions of justification, that is, *poor grounds*, *inadequate standards*, and *misrelating grounds to claims*. It should be noted that different justification criteria have different *grounds* and *standards*, as shown in the section of coding procedures. Therefore, students adopting different justification criteria have distinct types of nonscientific thinking in justification. For example, students adopting *plausibility of claims* may have nonscientific thinking of *naïve theory*. However, *naïve theory* is not the *ground* of the other two justification criteria (Table 3), and thus students adopting *empirical consistency* or *observation reliability* have no such type of nonscientific thinking. On the other hand, students adopting *empirical consistency* may have nonscientific thinking of misrelating *apparatus error* to claims. Because *apparatus error* is not the *ground* of the other two justification criteria (Table 3), such type of nonscientific thinking would not occur in justifications associated with the other two justification criteria.

For heterotype multiple-justifications, students' justification flaws occurred in different categories of *grounds* and/or *standards*. Even though using homotype multiple justifications, the justification flaws existed in different subcategories of *grounds* or *standards*. A conflict situation may occur as flaws exist in one justification of student's multiple justification. The resulting discordance causes students to reconsider the validity of their claims. Thus, the multiple justification becomes a useful way to produce scientifically acceptable claims.

6. Conclusions and Implications

This study identifies three students' criteria to justify their claims obtained after authentic laboratory experiments, that is, *observation reliability*, *empirical consistency*, and *plausibility of claims*. Three justification criteria have distinct *grounds* and/or *standards* of justification.

Undergraduates adopting *empirical consistency* justify their claims by indicating that claims or its supporting data were consistent with empirical evidence obtained from scientific experiments or daily life observation. Empirical evidence is the result of experimental measurement or daily life observation, rather than direct deduction based on theory. Two types of empirical evidence were found by undergraduates in the laboratory to justify their claim, that is, literature reports and personal experiences. Both consistency and inconsistency with empirical evidence were used by undergraduates to justify their claims.

Undergraduates also justify their claims by conveying that their assertions were consistent with a particular scientific theory. Unlike the criterion of *empirical consistency* that focuses on the comparison with empirical evidence, the criterion of *plausibility of claims* focuses on the comparison with scientific theory. Three types of scientific theories, which were scientifically acceptable theory, theoretically calculated value, and naïve theory, were adopted by undergraduates to justify their assertions.

Unlike the criterion of *empirical consistency* that focuses on the consistency of a claim with the corresponding empirical evidence, the criterion of *observation reliability* focuses on the quality of a given experimental procedure. Students attempted to justify their claim by indicating the reliability of the observation to form a claim or its supporting evidence. There are four subcategories identified in the category of *observation reliability*, that is, *careful manipulation*, *personal error*, *apparatus error*, and *method error*.

The behaviors students adopted justification criteria to justify *good results* and *bad results* they claimed are quite different. The most frequently used criterion is *plausibility of claim* to justify *good results*, while that is *observation reliability* to justify *bad results*. Furthermore, *personal error* is a major subcategory criterion to justify anomaly students encountered is a *bad result*. This finding probably results from that students are novices at experimental manipulation in the lab, and they have insufficient confidence to assert anomaly as a *good result* based on their manipulation. The finding of this study implies that personal manipulation is vital in students' learning in the lab.

This study implies that the multiple justification assists students perceive and improve the flaws or conflict situation in justification. Students' nonscientific thinking occurred because of flaws in three dimensions of justification, that is, *poor grounds*, *inadequate standards*, and *misrelating grounds to claims*. Different justification criteria have different *grounds* and/or *standards*, and are associated with distinct types of students' nonscientific thinking. For heterotype multiple justifications, students' justification flaws occur in different categories of *grounds* or *standards*. Even though using homotype multiple-justifications, the justification

flaws exist in different subcategories of *grounds* or *standards*. The results imply that a conflict situation may occur as flaws exist in one justification of student's multiple justification. The resulting discordance causes students to reconsider the validity of their claims. Thus, the multiple justification becomes a useful way to produce scientifically acceptable claims.

This study also reveals that the addition of the third criterion, *observation reliability*, to the framework of student's justification criteria is necessary. Science teachers should aware that nonscientific thinking may appear in any justification criterion students adopt. Encourage students to reconsider their claims by alternative justification criteria when students have flaws in a given justification criterion. After conflict situation is formed among justifications, prompt students to clarify their nonscientific thinking in justifications. For example, students misjudged a correct datum by a justification criterion of *plausibility of claim-naïve theory*, misunderstanding the Definite Proportions Law. Encourage students to reconsider the misjudged datum by an alternative justification criterion, *observation reliability*, such as discussing the accuracy, precision of measurements. A conflict situation may form between two justifications from both criteria. Then, teachers prompt students to clarify the Definite Proportions Law, and improve their justification. In authentic laboratory learning, it suggests that science teachers use three justification criteria to understand students' justifications and improve students' justifications in scientifically acceptable methods.

Acknowledgment

Funding of this research was supported by National Science Council, Taiwan, under Grant nos. NSC 95-2511-S-024-002-MY3.

References

- [1] C. N. Shealy, "A model and method for "making" a combined-integrated psychologist: equilintegration (EI) theory and the beliefs, events, and values inventory (BEVI)," *Journal of Clinical Psychology*, vol. 60, no. 10, pp. 1065–1090, 2004.
- [2] C. N. Shealy, "Justifying the justification hypothesis: scientific-humanism, equilintegration (EI) theory, and the beliefs, events, and values inventory (BEVI)," *Journal of Clinical Psychology*, vol. 61, no. 1, pp. 81–106, 2005.
- [3] A. Zohar and F. Nemet, "Fostering students' knowledge and argumentation skills through dilemmas in human genetics," *Journal of Research in Science Teaching*, vol. 39, no. 1, pp. 35–62, 2002.
- [4] V. Sampson and D. B. Clark, "Assessment of the ways students generate arguments in science education: current perspectives and recommendations for future directions," *Science Education*, vol. 92, no. 3, pp. 447–472, 2008.
- [5] M. A. Bishop and J. D. Trout, "Epistemology's search for significance," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 15, no. 2, pp. 203–216, 2003.
- [6] A. Samarapungavan, "Children's judgments in theory choice tasks: scientific rationality in childhood," *Cognition*, vol. 45, no. 1, pp. 1–32, 1992.
- [7] C. A. Chinn and W. F. Brewer, "The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction," *Review of Educational Research*, vol. 63, pp. 1–49, 1993.
- [8] C. A. Chinn and W. F. Brewer, "An empirical test of a taxonomy of responses to anomalous data in science," *Journal of Research in Science Teaching*, vol. 35, no. 6, pp. 623–654, 1998.
- [9] J. Y. Lin, "Responses to anomalous data obtained from repeatable experiments in the laboratory," *Journal of Research in Science Teaching*, vol. 44, no. 3, pp. 506–528, 2007.
- [10] A. Hofstein, O. Navon, M. Kipnis, and R. Mamlok-Naaman, "Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories," *Journal of Research in Science Teaching*, vol. 42, no. 7, pp. 791–806, 2005.
- [11] A. Hofstein and V. N. Lunetta, "The laboratory in science education: foundations for the twenty-first century," *Science Education*, vol. 88, no. 1, pp. 28–54, 2004.
- [12] T. Lord and T. Orkwiszewski, "Moving from didactic to inquiry-based instruction in a science laboratory," *American Biology Teacher*, vol. 68, no. 6, pp. 342–345, 2006.
- [13] V. Sampson and L. Gleim, "Argument-driven inquiry to promote the understanding of important concepts and practices in biology," *American Biology Teacher*, vol. 71, no. 8, pp. 465–472, 2009.
- [14] D. Taitelbaum, R. Mamlok-Naaman, M. Carmeli, and A. Hofstein, "Evidence for teachers' change while participating in a continuous professional development programme and implementing the inquiry approach in the chemistry laboratory," *International Journal of Science Education*, vol. 30, no. 5, pp. 593–617, 2008.
- [15] B. J. S. Barron, D. L. Schwartz, N. J. Vye et al., "Doing with understanding: lessons from research on problem- and project-based learning," *Journal of the Learning Sciences*, vol. 7, no. 3-4, pp. 271–311, 1998.
- [16] K. Hogan and M. Maglienti, "Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions," *Journal of Research in Science Teaching*, vol. 38, no. 6, pp. 663–687, 2001.
- [17] X. Zou, "How students justify their knowledge in the investigative science-learning environment," in *AIP Conference Proceedings*, vol. 720, pp. 105–108, 2004.
- [18] S. B. BouJaoude, "A study of the nature of students' understandings about the concept of burning," *Journal of Research in Science Teaching*, vol. 28, pp. 689–704, 1991.
- [19] I. Galili and V. Bar, "Children's operational knowledge about weight," *International Journal of Science Education*, vol. 19, no. 3, pp. 317–340, 1997.
- [20] D. H. Palmer, "Exploring the link between students' scientific and nonscientific conceptions," *Science Education*, vol. 83, no. 6, pp. 639–653, 1999.
- [21] D. H. Palmer, "Investigating the relationship between students' multiple conceptions of action and reaction in cases of static equilibrium," *Research in Science & Technological Education*, vol. 19, no. 2, pp. 193–204, 2001.
- [22] T. D. Sadler and D. L. Zeidler, "The significance of content knowledge for informal reasoning regarding socioscientific issues: applying genetics knowledge to genetic engineering issues," *Science Education*, vol. 89, no. 1, pp. 71–93, 2005.
- [23] T. D. Sadler, "Informal reasoning regarding socioscientific issues: a critical review of research," *Journal of Research in Science Teaching*, vol. 41, no. 5, pp. 513–536, 2004.
- [24] Z. Kunda, "The case for motivated reasoning," *Psychological Bulletin*, vol. 108, no. 3, pp. 480–498, 1990.

- [25] K. Edwards and E. E. Smith, "A disconfirmation bias in the evaluation of arguments," *Journal of Personality and Social Psychology*, vol. 71, no. 1, pp. 5–24, 1996.
- [26] P. H. Ditto and D. F. Lopez, "Motivated skepticism: use of differential decision criteria for preferred and nonpreferred conclusions," *Journal of Personality and Social Psychology*, vol. 63, no. 4, pp. 568–584, 1992.
- [27] P. Pollard and J. S. B. T. Evans, "The effects of prior beliefs in reasoning: an associational interpretation," *British Journal of Psychology*, vol. 72, no. 1, pp. 73–81, 1981.
- [28] D. L. Zeidler, "The central role of fallacious thinking in science education," *Science Education*, vol. 81, no. 4, pp. 483–496, 1997.
- [29] T. D. Sadler, "Evolutionary theory as a guide to socioscientific decision-making," *Journal of Biological Education*, vol. 39, no. 2, pp. 68–72, 2005.
- [30] C. W. Keys, B. Hand, V. Prain, and S. Collins, "Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science," *Journal of Research in Science Teaching*, vol. 36, no. 10, pp. 1065–1084, 1999.
- [31] B. I. Grimberg and B. Hand, "Cognitive pathways: analysis of students' written texts for science understanding," *International Journal of Science Education*, vol. 31, no. 4, pp. 503–521, 2009.
- [32] C. F. Wilcox, *Experimental Organic Chemistry*, Macmillan, New York, NY, USA, 1979.
- [33] D. J. Pasto, C. R. Johnson, and M. J. Miller, *Experiments and Techniques in Organic Chemistry*, Prentice-Hall, Upper Saddle River, NJ, USA, 1992.
- [34] R. A. D. Wentworth, *Experiments in General Chemistry*, Houghton Mifflin, Boston, Mass, USA, 6th edition, 2002.
- [35] T. Ruffman, J. Perner, D. R. Olson, and M. Doherty, "Reflecting on scientific thinking: children's understanding of the hypothesis-evidence relation," *Child Development*, vol. 64, no. 6, pp. 1617–1636, 1993.
- [36] B. Sodian, D. Zaitchik, and S. Carey, "Young children's differentiation of hypothetical beliefs from evidence," *Child Development*, vol. 62, pp. 753–766, 1991.
- [37] W. F. Brewer, C. A. Chinn, and A. Samarapungavan, "Explanation in scientists and children," *Minds and Machines*, vol. 8, no. 1, pp. 119–136, 1998.
- [38] T. A. Cahill, R. N. Schwab, B. H. Kusko et al., "The vinland map, revisited: new compositional evidence on its inks and parchment," *Analytical Chemistry*, vol. 59, no. 6, pp. 829–833, 1987.
- [39] W. C. McCrone, "The vinland map," *Analytical Chemistry*, vol. 60, no. 10, pp. 1009–1018, 1988.
- [40] M. Pickering and D. L. Monts, "How students reconcile discordant data: a study of lab report discussions," *Journal of Chemical Education*, vol. 59, no. 9, pp. 794–796, 1981.
- [41] S. I. Miller and M. Fredericks, "Reliabilism 'naturalized,'" *Social Epistemology*, vol. 16, pp. 367–376, 2002.
- [42] F. Dretske, "Conclusive reasons," *Australasian Journal of Philosophy*, vol. 49, pp. 1–22, 1971.
- [43] A. I. Goldman, "A priori warrant and naturalistic epistemology," in *Philosophical Perspectives*, vol. 13, pp. 1–28, 1999.
- [44] C. A. Chinn and B. A. Malhotra, "Children's responses to anomalous scientific data: how is conceptual change impeded?" *Journal of Educational Psychology*, vol. 94, no. 2, pp. 327–343, 2002.
- [45] J.-Y. Lin and D.-Y. Guo, "Students' nonscientific thinking in justification," in submission.

