SURVEY GUIDELINES AND BEST PRACTICES

Contents

I. CONDUCTING SURVEYS AT BCIT

Surveys are a useful tool for collecting information in an organized and methodical way about characteristics of interest from some or all units of population, and compiling the collected information into a useful summary form. At BCIT, surveys are used to assess quality of education (teaching and learning) and student experiences, plan for program and service reviews and enhancements, and determine the needs of our community and the best ways to meet those needs. With the increased emphasis on planning, accountability, and accreditation, there has been significant growth in the number of surveys conducted at the Institution to a point that there is a concern that surveys become burdensome resulting in a general decrease in survey response rates. In addition to the internally conducted surveys, BCIT also receives many requests from the government agencies to participate in surveys.

II. THE NEED FOR GUIDELINES

There is no doubt that surveys help the leadership teams hear from their community and accordingly make informed decisions. Thus to increase the quality and effectiveness of surveys, there should be a strong coordination when it comes to conducting surveys at the Institute. This process, unfortunately, can be negatively affected by several inconsistencies in coordination, including:

- Duplication of effort: same or similar questions are asked over and over in multiple surveys.
- Poorly designed surveys: reliable and useful survey results depend on the quality and efficiency of the survey. Poorly designed surveys mostly result in lengthy survey instruments, questions which are difficult to read and understand, and results that are not reliable and useful.
- Improper data collection practices: poorly planned surveys (e.g. inappropriate timing), the number and timing of reminders; inappropriate incentives that do not match the target group being surveyed (e.g. a BCIT bookstore gift card with a limited value and only for one textbook offered to employees, or a free recreational class for distant education students).
- Inadequate sharing of results: this will cause duplication of effort and resources as well as concerns by the respondent groups who would think their voice has not been heard, thus their time was wasted (losing trust).
- Over-sampling of specific populations: when students (and employees) are surveyed multiple times over a short period of time, this will lead to survey fatigue resulting in incomplete surveys or unanswered questions, which will yield low response rates.

The following guidelines outline the process by which surveys can be administered to students, faculty, and/or staff at BCIT. They are issued as a way to help:

- Maximize response rates providing better value through preventing redundancy and avoiding survey fatigue.
- Increase the quality, reliability, and usability of the data collected through well-designed surveys aligning with the survey objectives to ensure that survey results can be used for evaluation, assessment, and planning.
- Avoid inefficient use of Institute resources.
- Coordinate the timing of surveys.
- Help coordinate the population of surveys forming representative samples.
- Ensure compliance with relevant laws and regulations such as Freedom of Information and Protection of Privacy Act (FIPPA) FOIPOP and Canada's Anti-Spam Legislation (CASL) to protect the privacy and confidentiality of the respondents.
- Select an appropriate survey tool that meets Canadian Privacy legislation (server must reside in Canada; see the list of recommended survey tools of which servers reside in Canada.
- Ensure effective publication and use of the results across the Institute.

III. SURVEY CYCLE AT A GLANCE

BCIT Institutional Research Office (IRO) has developed this guide to help you plan and create a web-based survey that will provide you with useful and reliable results. The guide contains tips drawn from our survey experience and will help you design an effective survey, prepare your respondent list, write convincing email invitations and make the most of the resources available to you. In the following sections, you will find detailed information about a complete survey cycle.

1. IS IT REALLY NECESSARY TO CONDUCT A SURVEY?

We conduct a survey because we perceive a need for the information it will provide. Ask yourself critically if you *really* need a survey to get the information you desire. Perhaps the review of what others have done before you will serve your needs. Ask questions such as: Will your survey likely yield the same or similar results that others have found? If so, is it worth the cost (i.e. designers' analysts' and respondents' time, printing and data entry expenses, etc.)? If all you need is factual data, are you sure it is not available from another source such as student records or exam papers?

Also, is a survey the appropriate data collection method to elicit information you need? If you need answers to sensitive questions for which you can't expect honest and valid answers (such as questions on racism, drug use, sexual activity, or domestic violence) you may like to prefer either making inferences from published literature or hire a professional pollster with proven experience with sensitive surveys.

If you think a survey is the appropriate data collection method for your purpose, then:

- Make sure that there is not already available data for your purpose/research.
- Ensure that your project in which students (or faculty and staff) will participate is consistent with the Institute's mission and priorities.
- Gauge the overall impact: What will the impact of your survey be? Will your research positively or negatively impact the community? Does the survey overburden respondents?

BCIT IRO regularly administers student surveys to its enrolled students at various stages of their education at BCIT (e.g. Entry Student Survey, Part-time Students Engagement Survey, Full-time Students Engagement Survey) as well as ad-hoc surveys requested by the Institute or specific departments. The institutional survey and relevant results are available on the IRO website for your reference (you will need your BCIT employee credentials to access these reports).

Please note that depending on your purpose of collecting data from your target population, you may consider other methods of collecting information such as focus groups and personal interviews (face-to-face or telephone). These methods can be preferred over surveys if you are dealing with a time-constraint, or if the topic is not known enough and requiring deeper research to develop a reliable and valid instrument, or if you need more in-depth information/feedback from your target population.

2. DEFINING YOUR SURVEY OBJECTIVES

Before you begin your survey, determine the key research objective(s) to be answered by the survey, and develop a statement of the specific objective(s) and have it approved by the stakeholders. This is generally one of the most difficult and time consuming parts of doing survey research. The statement should be as clear as possible. It is a good idea to consult widely within a unit as to what are the important questions to ask on a topic, as this may vary depending on the stakeholder group. A good example would read, for example, "BCIT would like to obtain a better understanding of students undertaking part-time studies, with a specific emphasis on their demographic characteristics and to measure PTS students' perceptions of several aspects of their learning experience at BCIT, ranging from their experience with the registration process, student support services, security services, facilities, overall education, and other aspects related to their learning environment."

From there, develop a plan for analyzing the data and reporting the findings. What will you do with the data after you collect it? With whom you will share the results? In what format the results will be shared – as electronic or printed report, or visual presentations, or all of them depending on the audience/stakeholders who make use of the survey results?

3. SETTING A REASONABLE TIMELINE

A good survey is a major research project that involves considerable time and resources, so it is important that you spend some time planning. A survey squeezed within a tight deadline may not yield useful data that would address the overall objective of the project. Similarly, a survey that delivers the needed information too late may be a tremendous waste of time and effort. The followings should be considered when setting a reasonable timeline for a survey project:

- Decide if it will it divert staff resources away from other projects and needs?
- Develop a reasonable survey project timeline by considering the resources to be needed to conduct the survey from start to finish. It is not uncommon that developing a quality survey instrument, i.e. concise, clear, not overly burdensome on the respondents, and aligning with your objectives thus would yield valid and reliable data, can take up to six months. Activities could include:
    - literature research (this may take longer if the survey will be a prototype of a particular future survey series and explore a topic about which relatively little is known),
    - meetings with the stakeholders to determine the objectives and research questions and to collect feedback during the instrument design stage,
    - pilot-testing of the survey (testing entire survey from start to finish on a small scale) to collect feedback for potential revisions before launching the survey to the whole target group,
    - administrative errands such as preparing the data file in light of the selected survey frame; printing the survey instruments if required; in-person delivery of the survey instruments; and data entry (**Survey (sample) frame:** it is the tool to gain access to the population; means of identifying and contacting the units of survey population -e.g. a list of students at a post-secondary institute, a list of all vehicle that entered a specific parking lot during a specific time period, a list of neighbourhoods in a city; **survey units:** Items from a population selected for inclusion in a sample).
- Surveys should usually not be conducted close to the end of an academic term, during the exam periods, peak workloads, or holidays. Schools and service areas are highly advised to coordinate their survey timelines according to the timelines of the institutional survey

conducted by the IRO to avoid various surveys targeting the same student groups. Please check the Institutional Surveys Calendar to mitigate survey fatigue, collection of duplicate information, reduce cost, and maximize efficiency and resources.

4. DEFINE YOUR RESEARCH QUESTIONS

Information-collection is essential, but it is not your first step. Your first step is to define your research questions in light of project objectives. This is a critical part of the survey research process. Why are the right research questions so important?

- Clearly specified research questions are the only basis for making sensible planning decisions: Think about what you want to know. Ask yourself if you make Choice A, will you be able to answer your research questions? What if you make Choice B? Considering your research questions, and understanding the consequences of your decision, can help you achieve a successful outcome.
- Research questions identify the target population from which you will draw a sample: Should you examine both full-time and part-time students? Entry level students only, or first years as well? Do you need to include both Trades and Technology students? You should decide whom to survey after considering exactly whom you want to make decisions about. If you do not identify the people you are most interested in before collecting data, you risk excluding important respondents from your research.
- Research questions determine the appropriate level of aggregation: should you measure, for instance, effectiveness of program advising, at the level of major, program, school, student type (domestic vs international), program type (full-time vs. part-time), or institution? Research questions can be framed at different levels of aggregation. If you do not think about the issue of aggregation before data collection, you risk not having enough data at a crucial level of aggregation to answer your research questions (**aggregation** means joining or adding smaller data units together into larger totals; so **aggregate data** means information averaged or summed over a large population. The most common reasons for aggregation are to:
    - obtain more information about particular units based on specific variables such as age, gender, education, and income
    - avoid any privacy implications through anonymizing the data so that it becomes impossible to identify an individual).
- Research questions identify the outcome variables: What do you mean by, for instance, the effectiveness of program advising? Are you interested in student perceptions of, for instance, the number of program-advisor contact hours? You can determine appropriate outcomes only after considering exactly what you want to know. If you do not outline the

outcome variables before data collection, it is likely that the data collection will be ineffective on the key outcomes of the project.

- Research questions identify the key predictors. Does efficacy differ by student age, or student gender? By the advisor's years of experience? You can determine the important predictors only after thinking about all things that might be associated with your outcomes. If you do not carefully consider the predictors before data collection, you may be unsuccessful in measuring essential variables.
- Research question determine how much researcher control is needed and whether a descriptive, relational, or experimental study is most appropriate. When the effectiveness of, for instance, recently introduced online component (as the mode of delivery) is compared to traditional classroom instruction of a particular course, will you study the new system (online component) the first year it is implemented or after it has been in place for three years? Can you randomly assign students to online mode, or should you study them after they have selected the mode they prefer? If you have a great deal of control over the research setting, you can draw strong inferences. By not deciding how much control you need, you risk not having enough.
- Research questions identify background characteristics that might be related to outcome. For example when exploring the effectiveness of student counseling services, should you account for differences in student counselors' workload due to different numbers of students per counselor? Should you also consider differing goal orientation of students? Random assignment of students to counselors will eliminate most potential biases, but if you cannot use random assignment, differences in background characteristics such as counselors' counseling workload may distort your findings. Disentangling the effects of different predictors is often very difficult during analysis. It is much easier to control background influences by design.
- Research questions raise challenges for measurement and data collection. Are there published instruments that assess, for instance, student counseling services' effectiveness and that can be referenced to? Do they require individual administration? Can they be e-mailed, or must they be filled out in person? Are the measures appropriate for the students of a specific school? Or campus? Or student type? Is one measure of effectiveness sufficient, or should you use several? Data collection is costly, i.e. resources as well as respondents' and data analyst's time, so spend your resources wisely. If you do not think about measurement and data collection at the outset, you may never gather the key information you require.
- Research questions influence the number of people you must survey: Not only is research design guided by your questions but also the statistical analysis of the collected data. Different questions require different analyses which, in turn, require specific sample sizes to ensure statistically meaningful results. Once the data have been collected, it is too late

to add respondents, so you have to make this decision before data collection (For further information on defining research questions see Light, R. J., Singer, J. D., & Willett J.B. (1990); Suskie, L.A. (1996)).

In conclusion, research questions determine every aspect of research design. If your questions are not clearly defined, it will be challenging to use this research to make any crucial decisions. If you plan your research project in light of your research questions, you can ensure that your project will be able to answer those questions at the end. Thus, take the time to clearly define the research questions to be addressed in the survey. Get feedback from colleagues and relevant decision-makers. Review the literature on your topic to see how questions have been framed by other researchers, institutes, or professional associations, and what related issues or themes should be considered when exploring this topic.

5. DESIGNING THE SURVEY INSTRUMENT AND QUALITY ASSURANCE

Design the survey instrument in the online survey software if you are going to conduct a web-based survey. Survey questions should be created in light of the targeted information. Ideas for specific survey questions can come from existing instruments, colleagues, members of the target population, and your own observations and prior knowledge. You should cover your research questions adequately while keeping conciseness in mind. Do not include questions that may provide interesting but not necessarily useful results. Grammatical and typographical errors and inconsistencies should also be avoided to ensure the credibility of the survey (For further information see Sections 5.1. Standards of an Ethical and Efficient Survey: Survey Best Practices and 5.2. Survey Instrument Design Tips).

To make sure that your questions are clear, easy-to-understand and written proficiently, the instrument should be pilot tested with a relevant target group of respondents, co-workers and other stakeholders to validate its quality in terms of clarity, readability, length, coherence and cohesion, skip patterns, if any, and its appropriateness for the survey sample/target population.

5.1. Standards of an Ethical and Efficient Survey: Survey Best Practices

There are some ethical and professional standards of conducting surveys – from start to finish. Here are some key points from these standards:

- Avoid bias: Strive to conduct a survey in a way that is free of potential bias: minimize potential sources of bias, and disclose factors that may bias the results of the survey (see Sections 5.2. Survey Instrument Design Tips and 6.3. Survey Errors and Quality Assurance)

- Protect the rights of privacy of the survey respondents, and protect the confidentiality of individually identifiable information (See Section 9. Protection of Privacy and Data Confidentiality).
- Avoid harming, humiliating, or misleading respondents.
- Avoid the improper use of copyright materials.
- Take appropriate security measures before, during, and after administration of the survey.
- Informed Consent: All surveys should disclose to survey respondents why they should complete the survey, how the information will be used, how the data will be maintained (confidential, anonymous, etc.), approximate time needed to complete the survey, and whom to contact if they have any questions.
- Ability to Opt-Out: Any surveys should notify respondents that they are not required to complete the survey and can opt-out. This opt-out should remove them from future communications about the survey (i.e., reminder emails).
- Survey Length: Overall, the more concise a survey is, the more likely it is that survey respondents will complete it. So, avoid lengthy instruments and focus only on questions that will yield the information you need. Remember that the lengthy instruments could significantly affect the response rates of the surveys, and statistically not meaningful response rates most likely increase the risk of obtaining skewed results.
- Question Clarity: Survey questions should be easy-to-read/-understand/-respond to. Common problems with questions are confusing response scales, questions that are double-barreled or ask more than one question, or questions that are just unclear (for further information, see Section 5.2. Survey Instrument Design Tips)
- Beware of asking overly sensitive information: Asking for personal information can lead to a survey respondent exiting your survey, so please be aware of what and how you ask (for further information see Section 5.2. Survey Instrument Design Tips).
- Using samples is recommended: A **sample** is a subset of the target population from which inferences about the population are drawn. The method of random sampling is a good way to gather information about your population in question. By using sampling, the number of surveys sent to any one individual is lower, and hopefully, the likelihood of response higher (For further information on random sampling, see Section 6.2.2. Sampling Techniques).

5.2. Survey Instrument Design Tips

The fundamental characteristic of a good question is that it is **clearly understood.** When respondents find your question difficult to understand or answer, s/he is likely to interpret it in a way that makes it easier to answer; so they may not answer the actual question. We want our

respondents to answer our questions, not theirs. Using the following methods, you can ensure your question will be clear to your respondents so they will interpret it in the same way:

- ***Keep it short**:* Short, straightforward questions are easier to understand than compound or complex statements and therefore yield more accurate and reliable answers. They also result in a shorter instrument, thus, an increased response rate. Lengthy questions may also lead to fatigue which may cause the respondent to lose focus and interest.
- ***Make sure it asks only one question:*** Separate questions rather than having several related questions imbedded into a large one (e.g. double-barreled); otherwise a respondent may respond to the first part but ignore subsequent parts of the multi-part question. E.g. "Do you feel that using high technology is, should be, an objective **or** requirement?" To avoid this problem, check your wording for "and" and "or" and limit your questions to one adjective or adverb each.
- ***Keep it readable*****:** Keep the vocabulary simple. Go over each word in your question and make sure it is as simple as possible. Avoid the jargon of a specific field. Also avoid words which mean different things to different people.
- ***Avoid negative items:*** If you must have them, make sure they are emphasized (e.g. underlined, written in capital letters, or boldface), such as <u>NOT</u>, or **EXCEPT**. Also, beware of double negatives.
- ***Make all definitions, assumptions, and qualifiers clearly understood***: Clarify terms that could be misunderstood. A question like "Do you use the IBM mainframe operating systems?" can be interpreted differently by different people. One person could say that anyone who has ever looked at a printout is technically a user, while another could say that only those with hands-on experience are users. Similarly, rather than asking for age, ask for year of birth. Terms like "adequate," "minority," "value," "quality," "not too often," "acceptable," "now" are full of ambiguities.

  When asking for an estimate of something, specify the period of time you are interested in: "in the past year," "per day," etc. Specify the unit of measurement for all numerical responses. E.g. "How many courses do you take this term? _____courses."

  When asking for a rating or comparison, make sure you give a clear point of reference. Specify the criteria to be used. Do you want the staff rated on courtesy, cheerfulness, helpfulness, or knowledge? As mentioned above, do not include more than one criteria. Make sure your questions have no hidden assumptions written into them (e.g. "What percent of your budget do you spend on outdoor activities?"). Make sure the respondent has all the information necessary to answer the question. Be careful of presuming, for instance, that the respondent is at least high school graduate.

  When it comes to definitions, don't phrase them in such a way that your respondent feels stupid.
- ***Avoid making significant memory demands:*** It is very less likely that you can get accurate answers if you ask alumni, for instance, how they felt about the institution as first year students. Your respondents have had many experiences since then which will tint their

memories of those years. Asking "How many times did you visit the library last week" will yield more valid data than asking "How many times did you visit the library last year?" (asking the frequency instead of exact numbers would also be reasonable in the first question).

- ***Avoid asking for very precise responses:*** Questions asking the respondent to provide very precise information, such as the number of visits made to the library or number of calls made to home within a specific period of time, or exact annual salary, requires significant cognitive processing and will not be answered correctly. So, instead of asking for an exact amount, use a multiple choice or rating scale format (see Section 5.3. Formats for Survey Questions).

  Respondents will also provide less accurate answers to a question with many response categories (for instance, asking the respondents to identify which one of ten listed factors was most important in deciding to attend the Institute), then a question allowing the respondent choose, for example, the three most important categories; however the best way is to avoid long lists of response categories in either way.

- ***Avoid asking for broad generalizations about attitudes or opinions:*** It is more difficult for people to make broad generalizations about their feelings or opinions than to describe concrete actions. Opinion questions are also harder to analyze and interpret than behavioral "vignette" questions (contextualized questions based on a scenario design), since respondents' views comes from different frames of reference. For instance, student opinions of your project management program, may be based in part on the quality of their or their friends' high-school project assignments, or on stories they have heard about project management at other schools. By making as many questions as possible concrete or "behavioral", you will ensure your survey instrument is less ambiguous and more reliable. For instance, instead of asking "Do you consider yourself an active social media user?" ask, "How many hours per day do you usually spend on social media?" or "How many hours did you spend on social media yesterday."

- ***Make the questions easy and fast to answer:*** The respondent should easily understand how to decide on an answer and how to record it. A simple answer format requiring only a few checks is better than a complicated scoring key (see Section 5.3. Formats for Survey Questions).

- ***Keep It Interesting:*** Avoid making your questions monotonous. The six question categories below would quickly frustrate almost any respondent:

How important is your need to find information of the following types?

	Very important	Important	Neutral	Not very important	Not at all important
Simple facts for course purposes	○	○	○	○	○
Simple facts for research purposes	○	○	○	○	○
Simple facts for general purposes	○	○	○	○	○
Complex data for course purposes	○	○	○	○	○
Complex data for research purposes	○	○	○	○	○
Complex data for general purposes	○	○	○	○	○

It is not uncommon to come across with survey instruments using the same kind of question format (e.g. likert scale) throughout, which is one of the most important reasons of monotonous or boring instruments; so try to use a variety of formats, but not too many; otherwise your respondents might get confused.

- ***Avoid biased, leading, and loaded questions:*** A biased, loaded, or leading question is one that suggests or leads the respondent to a certain answer; i.e., the way the question is phrased has an influence on the response. So, this type of questions can easily distort the responses and affect your survey results. While no professional survey designer and ethical researcher writes these kinds of questions on purpose, it is easy to write them unintentionally.

  Survey questions should ensure that all possible choices are provided and clear to the respondent; otherwise it can lead to only one response alternative as in the following example.

  > *Do you think most manufacturing companies that lay off workers during slack periods could arrange things to avoid layoffs, giving steady work right through the year?*

  - Yes
  - No
  - No option

  When asked in this way, this question yielded the following results:

  63% Yes, companies could avoid layoffs

  22% No, companies could not avoid layoffs

  15% No opinion

  In this question, the only alternative that respondents are offered is whether they think companies can arrange things to avoid layoffs. Usually, when only one alternative is available to respondents, they tend to agree with that alternative. Here is a different wording for the same question:

  > *Do you think most manufacturing companies that lay off workers during slack periods could arrange things to avoid layoffs, and provide employees with steady work throughout the year, or do you think that layoffs are unavoidable?*

  - Yes, companies could avoid layoffs
  - No, layoffs are unavoidable
  - No opinion

This time the question yielded the following results:

35% Yes, companies could avoid layoffs (compared to 63% in the previous question)

41% No, layoffs are unavoidable (compared to 22% in the previous question)

24% No opinion (compared to 15% in the previous question)

As seen, this question presents two clear alternatives: *companies could avoid layoffs* AND *layoffs are unavoidable, and* the results are more evenly distributed between the response categories than for the previous question.

Here is an example for a loaded question (questions including non-neutral terms, such as "harsher," "power" and "forcing"):

> *Should harsher penalties be enacted for crimes committed against the elderly?*

To avoid biased, leading, or loaded questions, try to imagine you are writing questions from the opposite point of view. If you are trying, for instance, to collect data to support the need for expanded counseling services, imagine you're trying to cut back on them.

In theory, providing all possible alternative responses and using neutral terms in the question allow respondents to reflect on the answer before responding and thus provide a more reliable answer.

- ***Ask about both the pros and cons of an issue:*** Don't ask the respondent to make criticism without giving him/her a chance to praise as well. And don't ask him/her if s/he is in favour of a new program or extending the hours that an office is open (who wouldn't be?) without noting the costs and liabilities. Here is an example from BCIT's Part-time Studies Student Survey (2017) where students were asked if they would be interested in participating in the mandatory Universal Transit Pass. As can be seen, the question includes all information the students needed to know before answering the question.

  > *Would you be interested in participating in the mandatory Universal Transit Pass (U-Pass BC) if it were offered for an additional fee of $41.00/month? (Please note: the majority of BCIT courses are four months in length)*

  - Yes
  - No

  (The neutral response category, i.e. "no opinion" or 'I'm not sure" was not included on purpose).

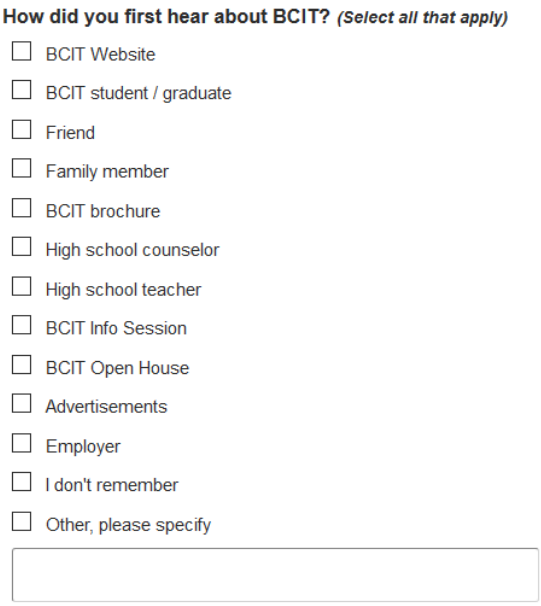
- ***Avoid questions that people are uncomfortable answering honestly:*** People tend to be comfortable with the status quo at the cost of deceiving themselves or others about their inclinations in sensitive areas, and to be reluctant to admit any inferiority or wrongdoing. Many will not answer honestly if you ask them, for instance, whether they use substances, go to church, have bullied or been bullied, or vote. Similarly, many will hesitate to report accurately their salary, age, job, grade point average, whether they failed a particular course, and whether they were rejected from other schools to which they had applied; or

many will indicate they (strongly) agree with the statement, I believe all people should be respected as equals," even if they demonstrate prejudice or intolerance towards certain groups.

When phrasing questions whose responses have potential level of social desirability, find ways to make them feel comfortable with their answers so that they respond more truthfully. For example, when asking about age or salary, prefer broad categories rather than a specific figure.

Try not to phrase questions in a psychologically intimidating way. Instead of directly asking if a respondent does something good or bad, ask *how often* he or she does it (instead of "Do you cheat on tests?" ask "How often do you cheat on tests" – and do not forget to include all response categories including "never."

- **If appropriate, let people admit they don't know or can't remember:** Forcing an opinion from people who don't have one or a memory from people who can't remember one is practically the same as forcing them to lie. Here is an example:



- Use multiple assessments with a variety of methods, formats, and topics: The only way you can be sure your questions are not biased or leading is to cross-check your responses against other data. This process is called "triangulation."

(For further information of Survey Design see Suskie (1996); Light, R. J., Singer, J. D., & Willett J.B. (1990); Statistics Canada, *Survey Methods and Practices (2010)*)

5.3. Formats for Survey Questions

Once you have defined your research questions, the first step in developing your survey instrument is to construct the questions that will get you the information you want.

This section describes some common formats for survey questions:

5.3.1. Closed-ended Survey Questions

5.3.1.1. Two-choice (Dichotomous) Questions

This type of questions have only two response categories (choices/answers) and are the simplest version of the closed questions. They are often constructed as a "Yes/No question" and split the respondents into two distinct categories (1). They are also used as screening questions to (2) determine whether respondents meet the criteria defining your target group that would make them eligible to participate in the survey (if this is the purpose, it is asked in the very beginning of the survey), or (3) prevent respondents from being asked a series of questions that do not apply to them (this is referred to as skip patterns). Here are some examples:

Have you every voted in any election (federal, provincial, municipal)?

- *Yes*
- *No*

Are you a full-time student?

- Yes
- No

(If your survey is intended for full-time students only, those who choose "No" will be terminated from the survey by directing them to the submit page of your survey (here you can provide them with information why they did not need to complete the survey; do not forget to add a thank you note here).

Do you use a mobile device, such as a cell phone, smart phone, tablet, or something similar?

- *Yes*
- *No (Go to Question 22)*

5.3.1.2. Multiple Choice and Checklist Questions

Another type of closed-ended questions is multiple-choice or fixed choice question that consists of a question followed by a list of response categories (choices/answers) that the respondent are asked to choose ***one*** from among. Checklist questions are a variation of multiple-choice questions in which respondents may choose ***one or more*** responses. Instructions, such as "select only one," "choose only one," "select all that apply," "choose all that apply," "select minimum 2 choices," "choose maximum 3 items," should be included to let the respondent know whether the question is multiple choice or checklist. Note that an 'Other (specify)' category is normally provided at the end of the choice list to ensure that the list is exhaustive. Here are examples for both types:

Which BCIT Library do you usually visit in person? (Select only one)

- *Burnaby Campus Library*
- *Marine Campus Library*
- *Aerospace Technology Campus Library*
- *I only use online services*

How do you prefer to learn about information regarding BCIT? (Select all that apply)

- ☐ *E-mail*
- ☐ *Face-to-face meetings*
- ☐ *BCIT Update Newsletter*
- ☐ *The Loop*
- ☐ *BCIT website*
- ☐ *Other (please specify)* _____

Multiple-choice questions are widely used, because they are quick and easy for respondents to answer, and particularly good for collecting factual information. The main difficulty in drafting multiple-choice questions is ensuring that you have *all* possible response categories (choices/answers). Although "Other (specify)" choice is to cover the rest of the choices not listed, you should not use it to salvage a poorly written question. The best way to make sure your list is exhaustive is either (1) pretesting the question as an open-ended question before launching the actual survey and using the most frequent answers as your responses, or (2) updating your choice list for the next survey cycle in light of the responses provided under "Other (specify)" which yielded the highest percentages.

The second difficulty in drafting multiple-choice questions is making sure you have mutually exclusive (clearly distinct and separate from one another) and exhaustive (comprehensive, all-

inclusive) response categories. For example, they are not mutually exclusive in the following example.

How old are you in years?

- *20-30*
- *30-40*
- *40-50*
- *50 or more*

As seen, the response categories overlap (the answer of a respondent who is 30, 40, or 50 years old would fall in two response categories). Data analysis would be problematic for this question as it is impossible to figure out which category such a respondent would choose. In this example, response categories are not exhaustive either. If persons under 20 are part of the target population, how would they answer this question?

An appropriate choice for response categories would be:

How old are you in years?

- *Under 20*
- *20-29*
- *30-39*
- *40-49*
- *50 or more*

One of the most important disadvantages of checklist questions is that you get no sense of the relative importance or frequency of each item checked. If one student checks three items in the list of response categories (e.g. reasons they decided to go to BCIT), we would have no idea of the relative importance of each of those reasons. They may all be equally important, or one reason may be far more important than the other two.

Here are some suggestions for writing multiple-choice questions:

- The question should ask a complete question, even it is phrased as an incomplete statement. The respondent shouldn't have to read the responses to figure out what you are asking, as they would with, for instance, "The library is:".
- Limit the number of possible response categories to five or six (if possible). It's hard for most respondents to keep more than that many in mind.
- Your response categories should all be mutually exclusive.
- ***If*** there is a dominant response category that you are not interested in, include it in the question itself (e.g., "Aside from housing, what…?").
- Order the response categories logically-numerically if the answers are numbers, for example. This will help your respondents who know their answer find it quickly (if your

list of response categories is a list of numbers, most people will pick from the middle of the list).

- "Stack" your response categories *vertically.*
- If it is a checklist question, consider randomizing the order of response categories. This will help you ensure that the order of response categories (choices/answers) is not influencing your survey results, so randomization eliminates order biases (it has been claimed that whereas the categories/choices in the upper part of the list are more likely to be picked, rated high, etc., those in the lower part are more likely to be skimmed especially if the list is so long that it can cause the respondent to get fatigued.

5.3.1.3. Rating Questions

Rating questions are employed to ask respondents to rate their answer. ***Likert scales*** are most widely used attitude or opinion scale type in the social sciences. They are comparatively easy to construct, can deal with attitudes of more than one dimension, and tend to have high reliabilities. Likert scales are more practically used to measure attitudes and opinions rather than factual information. Below are examples of Likert scale rating questions:

*How satisfied are you with classroom facilities and equipment at BCIT? **

- Very satisfied
- Satisfied
- Dissatisfied
- Very dissatisfied

* (Middle scale point is optional: "Neutral" or "Neither Satisfied nor Dissatisfied")

Please indicate the extent to which you agree or disagree with each of the following statements:
The instructor...

	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
Treated students fairly and impartially	○	○	○	○	○
Demonstrated knowledge of course materials	○	○	○	○	○
Was prepared for class	○	○	○	○	○
Was available outside of class	○	○	○	○	○
Stimulated interest int he course	○	○	○	○	○

How likely is it that you will recommend BCIT to your family, friends, and/or co-workers?

10 Very Likely	9	8	7	6	5	4	3	2	1 Very Unlikely
○	○	○	○	○	○	○	○	○	○

Many people are familiar with Likert scales and therefore find them easy to complete. Like other rating scales (graphic rating scale, semantic differential scale, ecosystem rating scale, side-by-side matrix), they are efficient as a lot of information can be provided quickly and compactly. They also allow comparisons among answers within the scale. However, researchers and survey designers should be careful when designing rating questions for their survey instrument. You may need to consider the followings when formulating rating questions:

- Do not construct your entire survey instrument of Likert scale questions to facilitate your data analysis. Likert scale may not always be appropriate or applicable.
- Decide on the number of response categories that the rating scale will have. It could have just two (e.g., satisfied, dissatisfied) or as many as 10, ranging from 1 (Very Unlikely) to 10 (Very Likely).
- When using a Likert scale on agreement, consider the "yeasayer/naysayer effect": Some people with generally positive feelings and opinions toward a subject will tend to check all the "Strongly Agree" responses without reading the individual questions. Similarly, some people with generally negative feelings and opinions will tend to check all the "Strongly Disagree" responses with reading the questions. This is called yeasayer/naysayer effect" (Suskie: 1996): To prevent it, keep at least some of your statements negative or presenting an opposing view.

  Should you force an opinion/response at all or offer a "mid-point," acting as a neutral option, by creating a scale with an odd number of response categories (examples of mid-point category are: "Neutral,", "Undecided," or "Neither Satisfied/Agreed nor Dissatisfied/Disagreed). If you decide to leave out this choice and force an opinion on either the negative or positive spectrum of the even scale, you may be forcing a lie from someone unfamiliar with the issue. This will result in biased results as truly neutral people will have to choose a category that indeed does not reflect their opinion. On the other hand, if you include it, some people will tend to choose it to avoid expressing their opinion on which they would take the time to contemplate. You would also not be able to figure out, unless you ask a follow-up question, whether an "Undecided" respondent, for instance, is really "Undecided" or "Uninformed." This will also bias your results. Each approach has its strengths and weaknesses. Suppose you are surveying your part-time students on whether their "course instructor spoke clearly." It could be argued that the instructor either spoke clearly or didn't, and each respondent should be able to either agree or disagree with that statement; however if the same group of students are asked if "the President is doing a good job," it could be argued that some students might be truly neutral on this issue. Similarly, if your survey results are going to be used towards making a business decision about a particular current practice (e.g. continuing to opt out the credit cards as an acceptable method of payment during course registration or starting to accept them as an alternative method of payment with an extra cost), then even number of scales could work better to make solid business decisions. Let's see how an odd number of scale can result in significant bias which, in turn, would affect your business decision.

Please indicate to what extent you agree or disagree with the following statement:

"I would prefer using my credit card for course registration if it were offered for an additional cost"

	N	%	mean
Strongly Agree	130	16%	
Agree	220	27%	
Neutral	320	39%	
Disagree	100	12%	
Strongly Disagree	50	6%	
Total	820	100%	3.3

If we calculate the average score (mean) of the level of agreement regarding the ability of using credit cards for course registration if it were offered for an additional cost, we would get 3.3 (**neutral**, 1=Strongly Disagree, 5= Strongly Agree), but if we define the "Neutral" category as missing (eliminating the category from the data) and use the 4 category rating list, our average score would be 2.9 (**agree**, 1=Strongly Disagree, 4= Strongly Agree). So, eliminating "neutral" answers moves the data in favour of the preference for credit card usage (if we assume that those who chose "Neutral" as their answer are truly neutral in their opinion, then omitting neutral responses will bias the data; however, the opposite, i.e., keeping the "Neutral," can also cause bias as some of the respondents will use it as a way out of expressing their true opinion. Please note that eliminating the "Neutral" category from your data would reduce the accuracy and reliability of your data as you intervene with your respondents' responses. The best solution is to decide on the number of response categories before you launch your survey in light of the purpose of your survey).

- Consider if you need to include a category on your rating scale such as "Don't know," "No opinion," or "Not Applicable." This depends on the question being asked: e.g., when asking about a particular service that the respondent may have never used, it is necessary to include a "Not Applicable" category.
- Note that "Don't Know" is very different from a "Neutral" opinion. A "Neutral" column should be placed in the centre of your Likert scale and included in your data analysis. A "Don't Know" response, however, quantifies those refusing to answer the question, thus should be treated as missing data in your data analysis, unless you ask your respondents if they do or don't know a specific thing, or if their knowledge or lack of it is merely what you wanted to know.
- Your decisions on (1) the number of response categories on your rating scale (2) the use of a neutral alternative, and (3) the use of "Not Applicable" should depend on your survey objectives and the questions to be asked.

5.3.1.4. Ranking Questions

In a ranking question, the respondent is asked to rank either all or a subset of response categories. For example:

> *Please rank your 3 most favourite colours. Rank your most favourite colour "1," the second "2," and third "3."*
>
> ☐ *Blue*
>
> ☐ *Red*
>
> ☐ *Orange*
>
> ☐ *Green*
>
> ☐ *Yellow*
>
> ☐ *Black*
>
> ☐ *White*
>
> ☐ *Violet*
>
> *Here is a list of source of financing your tuition fees. Please rank each category to indicate how important it is to you by placing the number "1" beside the source you think is most important, a "2" beside the source you think is second most important, and so on.*
>
> ☐ *Full- and/or part-time employment*
>
> ☐ *Personal savings*
>
> ☐ *Parental assistance*
>
> ☐ *Spousal assistance*
>
> ☐ *Student loans*
>
> ☐ *Grants*
>
> ☐ *Scholarships*
>
> ☐ *Social Security benefits*

As mentioned in the previous section, if you randomize the order of your rank categories, you will eliminate potential order biases.

Ranking questions have some serious weaknesses:

- They can be tedious to complete, especially if the items to be ranked are many in number or different from each other. Often a respondent may have to reread the list over and over when s/he is ranking it.
- Rankings often incorrectly assume that respondents feel differently about every category (choice) in the list; this is often not true. A respondent may think or feel two or more items of the same importance or preference. In the above example, a respondent's actual ranking for both "Personal savings" and "Parental assistance" as 1.
- Rankings provide very limited, and mostly incomplete information, making the results difficult to interpret. Consider one respondent's response to the following question:

  Please rank the reasons you decided to attend BCIT, giving a "1" to the most important reason and a "7" to the least important reason:

  - ☐ *Reputation (respondent assign **1** to this choice)*
  - ☐ *Hands-on/practical training (4)*
  - ☐ *Program length (3)*
  - ☐ *Program/course only offered at BCIT (5)*
  - ☐ *Convenience of the location (6)*
  - ☐ *Lower cost/tuition fees (**2**)*
  - ☐ *Other (please specify): _____ (7 or left blanked)*

The problem is that you will never know if this respondent felt that "Reputation" was far and away the most important reason or if 'Reputation" and "Lower cost/tuition fees" were actually equally important. Even more significantly, you will never know if "Reputation" was the most important reason or just the best of the choices given. Similarly, you will never know if "Convenience of the location" was not a factor in the decision to attend BCIT, or just the least important of 7 very important factors (or 6 if s/he doesn't have any "Other" reason to indicate).

- The sizes of the rank intervals are unknown and unlikely to be equal, i.e. the interval between 1 and 2 cannot be assumed to be the same as that between 2 and 3. This makes data analysis difficult. For example, if four items (categories) are listed to be ranked, the respondent will rank them 1, 2, 3, and 4, but it is quite likely that s/he considers the first three to be very close and the last item to be a distant fourth. Such information cannot be collected from the simple ranking.
- It is not uncommon that respondents are not be able to rank all items (categories) on the list. It may be reasonable to expect them to only rank a few (e.g. four or less among seven items).

In light of these limitations, we would hesitate to recommend using rankings. Using rating scales, which are discussed earlier, will yield more useful and reliable data.

(For further information of the formats for survey questions, see Suskie (1996); Statistics Canada, *Survey Methods and Practices (2010)*)

5.3.2. Open-ended Survey Questions

Open-ended or free-response questions leave a blank space where the respondents compose their own answers as in the following examples:

What was the *one best* part of this program?

_____

How can the Student Support Services be improved to better meet your needs?

_____

Open-ended questions are useful and yield invaluable information, because they:

- Do not provide a limited set of predefined answers, so answers and comments provided are mostly genuine.
- Are unprompted, so they reflect respondents' thinking process, and frame of reference.
- Allow the respondents to answer the question in detail and clarify the points they want to make.
- Can, and often do, bring out unforeseen findings. This will help the survey designer revise the question and/or response categories for the next survey cycle.

There are a few occasions when open-ended questions may be best:

- When many response categories, say more than seven, are possible, or when a multiple-choice question might overlook some important responses.
- When a multiple-choice question might bias responses by leading respondents to a particular direction.
- When your question is interesting enough that people will want to answer it. People enjoy being asked, "How can we improve…" or "How would you improve…", for example.

Despite many advantages, open-ended questions are not very popular with respondents as they make them contemplate more and lengthen the time required to complete the survey. It is common that many respondents leave them blank.

Open-ended questions are also not very popular with researchers as the answers (comments) provided are generally difficult to read (for paper-based surveys), categorize, process, and

interpret. Also, responses may be irrelevant or unnecessarily detailed, which makes the analysis harder.

While open-ended questions can contribute immensely to the research when used sparingly, the researcher/survey designer should avoid excessive use of this group of questions, or consider other data collection methods such as individual interviews or focus group discussions.

5.4. How Can You Maximize Your Response Rate?

The following factors probably have the highest effect on your response rate (Suskie, 1996):

- ***The topic of your survey:*** A survey asking for simple facts will probably get a higher return rate than one asking for opinions on a sensitive issue.
- ***The target population:*** You will probably get a higher response rate if your respondents are interested in your project or if they think their participation would add a value to your research or to the decision making mechanism.
- ***How considerate you are of your respondents:*** You should recognize your respondent's effort and time and that they are doing you a favour by completing your survey. If you show your appreciation by doing all you can to minimize your respondents' trouble and make their job as easy as possible, they will be much more likely to respond to your survey and give you valid information. Offering monetary incentives is another way of showing your appreciation (see Section 7. Considering Incentives).
- ***How professional and important the project appears:*** If you appear professional and the research appears important, you will make your respondents' contributions seem much more worthwhile.

There is not much you can do to change the first two factors: most probably you cannot change the topic you have been asked to research or the group you have been asked to survey. The only way to handle respondents' concerns about the sensitive information is to (1) ensure that their responses will be treated as confidential data and their privacy be protected (See Section 9. Protection of Privacy and Data Confidentiality), and to (2) provide a contact information they can reach out for their questions or concerns.

Your can do, however, many things about showing how considerate and professional you are: To ensure that you are considerate of your respondents, ask yourself the following:

- *How long is your survey instrument?* – the shorter it is, the more considerate you are of your respondents' time.
- *How clear are the questions?* – respondents should not spend extra time trying to figure out what you really mean.

- *How clear are the directions and layout?* Respondents should be able to move easily from one question to the next – consider skip patterns to prevent respondents from being asked a series of questions that do not apply to them.
- *Was the survey instrument pilot-tested?* – If it wasn't you can't be sure your questions and directions are clear enough.
- *Which questions are first? or last?* – The first question should be interesting to your respondents. The long, complicated, and boring questions and questions asking sensitive information should be at the end where your respondents will see that the survey instrument won't take much longer and have built a trust needed to answer the sensitive content questions.
- *When is the survey invitation including the survey access link e-mailed or the survey delivered in person (for paper*-based surveys*)?* – Will your survey arrive at a time when your respondents are busy with other matters, such as exam weeks or statutory holidays, or school breaks?
- *Do you offer a summary of findings?* – Your respondents would like to see the impact of their efforts.

To find out how professional you appear to your respondents, ask yourself the following:

- *Does your survey instrument ask interesting, important-sounding questions?*
- *Is it carefully laid-out and grammatically flawless?*
- *Does it have a title?*
- *Can the survey invitation convince your respondents that it's worth taking the time to respond?*
- *Is your survey invitation well-articulated and comprehensive enough? (See Section 8. Drafting the Survey Invitation and Reminders)*
- *Does your survey invitation include a privacy statement? (See Section 8. Drafting the Survey Invitation and Reminders)*
- *Is it well-reproduced on quality paper? (for paper-based surveys)*
- *Does it render properly on multiple devices, i.e. desktop computer, mobile phone, tablet, etc. (for online surveys).*

6. THE POPULATION AND THE BASICS OF SAMPLING

6.1. What is the Population to Be Covered?

- **A population** is defined as a collection of units and their characteristic (including a geographic and time reference); e.g. the student population of BCIT is defined as all students studying at BCIT regardless of their student grouping (Full-time, Part-time), program type (Trades & Technical Studies, Apprentice, Technology), or other academic demographics.
- **The target population** is the set of units about which information is wanted and estimates are required, i.e. the units your survey is intended for, the units which you *want* to cover and your results to *apply* to. Depending on the nature and the purpose of the survey, these units will likely be individuals, households, schools, businesses, etc. It can be, for instance, *all* cars that use a certain road in a specific time period, *all* bears that were permanently marked and later recovered dead by hunters or researchers, *all* currently enrolled students, *all* full-time students, full-time students in a specific program type (e.g. Trades students, or Technology students), students in a specific school, full-time students living on campus, or only entry-level full-time students; *all* part-time students, part-time students taking courses in a specific campus (e.g. Burnaby, Downtown, and Marine), or only those who take distance/online classes; graduate students; faculty and staff, or a combination of two or more of these groups.
- **Survey Population** consists of the units which have some chance of actually being included in the survey, thus it is the population that is actually covered by the survey and the results *do apply* to. It may not be the same as the target population, though the two populations, ideally, should match closely in coverage (characteristics the population was specified according to: the type of units that comprise the population and their defining characteristics (who or what?), a time reference (when, which period?), a geography reference (where?). For this reason, the survey populations should be clearly defined in the survey documentation.

(All definitions in this section are taken from the *Survey Methods and Practices (2010)*)

6.2. Sampling

It is not always necessary to survey an entire population in order to have reliable, valid and generalizable results to make inferences. It has been observed at BCIT that in recent years there has been a significant decline in response rates of student surveys due to survey fatigue. Using a sample instead of the entire student population would help reduce the survey fatigue, which will in turn increase the survey response rate. However, if the target population is small, such as the first year full-time students of a specific school or all enrolled students of a specific program, it may be more suitable to survey the entire target population. If a sample survey is preferred, then the target population should be well-defined, otherwise the sample selected to be surveyed may contain people outside the intended population, or exclude those who should have been included. In that case, the survey results are likely to be biased.

6.2.1. Sample Selection Approaches

Sampling approaches fall into two categories:

1. Probability Sampling: Probability sampling, carefully designed and carried out, has greater validity and credibility than nonprobability sampling. Often, the costs and time required to conduct a probability sampling are greater than for a nonprobability sample. While being more complex, costly, and time-consuming, only probability sampling allows for reliable estimates, so inferences (generalization) can be made about the target population with a known degree of confidence, because units from the population are *randomly selected* and each unit has some known chance of being selected into the sample.

2. Non-probability: Non-probability sampling is used for empirical studies (e.g. qualitative research); results are usually non-attributable to the population; and no measure of sampling error is possible, so you would not know how good or bad your estimates (results) might be. In order to make inferences about the population from a non-probability sample, the data analyst must assume that the sample is representative of the population.

Choosing between these two types of approaches is a matter of weighing the requirements for validity and credibility against a realistic assessment of the requirements for timeliness, cost-effectiveness, and simplicity (for further information on sampling techniques *Survey Methods and Practices (2010);* Henry, G.H. (1990)).

6.2.2. Sampling Techniques

A randomly selected sample will work in the same way as a census survey while minimizing related costs – including the costs of printing and data entry (for paper-based surveys), survey fatigue (See Section 6.2.1. Sample Selection Approaches for "probability" and "non-probability sampling" before you decide for your sample design). **Randomly** means that every member of the target population has an equal chance of being included in the sample. In other words, the process of randomization is done to minimize the process bias/error, but it is still possible that all the randomized participants are not representative of the target population.

In general, a good sample reflects the relevant characteristics of the population and represents the target population, but if the population itself is not homogenous and the sub-groups significantly vary in size, precision can be increased through **stratified sampling**, which is a sampling method with which the population is divided into homogenous, mutually exclusive groups called strata (and each group is called stratum), and then independent samples are selected from each stratum. A population can be stratified by any variables that are available for all units on the survey frame prior to the administration of the survey (e.g. size or demographic characteristic, location, etc.). Let's assume you wanted to find out the e-journal usage of BCIT students. Knowing that Trades students behave quite differently from the technology students in terms of the e-journal usage, you may want to separate them into different groups/strata; or when surveying the part-time studies students of BCIT, you would consider the size of the schools offering part-time courses (e.g. the School of Business versus the School of Transportation), and thus stratify your population by school. In simple random sampling, there is no assurance that a sufficient number of part-time students from the School of Transportation would actually be included in the sample.

A stratified sampling example from BCIT: Let's say we want to explore the Library usage patterns of our Trades and Technology students. Our enrolment data shows that 80% of our students are from Technology programs and 20% from Trades programs. We need, for instance, at least 100 respondents for our Trades segment/stratum. This means that we would need four times as many Technology students as they represent four times as many students at BCIT, for a total of 500 respondents. This is called **proportionate stratified sampling.** If a particular strata is very important to the survey project but occurs in too small a percentage to allow for meaningful analysis, then **disproportionate sampling** is undertaken where you oversample your smaller stratum and then apply **weighting to** your collected data. The latter is called **post-stratification weight** as it is computed after the all data has been collected. Computing the post-stratification weight is necessary to eliminate the bias introduced into any estimate to be obtained from our sample data because statistical computation will give greater weight to those individuals we intentionally oversampled (Please note that we sometimes unintentionally oversample some stratum/groups and undersample others. In that case, we still correct the biases with the post-stratification weighting procedure).

6.3. Survey Errors and Quality Assurance

There are two kinds of surveys: census and sample survey. A **census** collects information from all units of the populations; a **sample survey** collects information from only a fraction of units of the population. Regardless of the kind of the survey and the variety of statistical techniques, the information collected is used to make claims, or inferences, about the population in question as a whole.

Accurate and precise survey results are ensured by minimizing or eliminating survey errors caused by sampling error and non-sampling error; if these errors are not anticipated and dealt with in the planning, design, and development phases of the survey and if necessary quality control techniques are not used during the implementation phase, these errors can render survey results useless.

6.3.1. Sampling Error

Sampling error is the error that occurs naturally when sampling. In fact, it is the *difference (*not error) between a sample and the population from which it is drawn. It arises from estimating a population characteristic by measuring only a portion of the population (sample) rather than the entire population, because samples are generally not identical to the populations from which they come. Therefore, all sample surveys are subject to sampling error.

The magnitude of the sampling error can be controlled by the sample size (it decreases as the sample size increases) or method of sampling (e.g. if the population is not homogenous and the sub-groups are very different in size, stratified sampling will work better than a random sampling) (for further information on sampling, see Section 6.2. Sampling). We can, hence, say that a census has no sampling error since all members of the population are surveyed, thus results from a census are expected to be more accurate than results from a sample survey. That being said, all surveys are subject to non-sampling errors, i.e. all errors that are unrelated to sampling.

6.3.2. Non-sampling Error

Non-sampling error arises during the course of virtually all survey activities and results in the data values not accurately reflecting the “true” value for the population. While sampling errors are associated with the process of selecting a sample and are only present in sample surveys, non-sampling errors are the errors that arise during the course of all survey activities, including both sample surveys and censuses.

Non-sampling errors can be classified into two groups:

- **Random errors** are the unpredictable errors resulting from estimation whose effects are generally nullified if a large enough sample is used; otherwise they usually lead to an increased variability in the characteristic of interest. In other words, the greater the

difference between the population units, the larger the sample size required for reliable results (**Estimation** is the means by which the values for the target population are obtained so that conclusions can be drawn about that population based on information gathered from only a sample of the population. An estimate may be a total, mean, ratio, percentage, etc.).

- **Systematic errors** are the errors that tend to go in the same direction and accumulate over the entire sample, so they result in bias in the final results. For example, you recorded the temperature at noon every day for a month in your backyard. If your thermometer was wrongly calibrated so that it was always 3 degrees high, the faulty thermometer would produce a systematic error in your measurements, in which case all temperatures would be systematically overestimated. Unlike sampling error or random errors, this bias is not reduced by increasing the size of the sample or changing the process of the sample selection, so they are the main source of concern for survey's data quality.

Non-sampling errors are caused primarily by coverage, measurement, non-response, and processing:

- Coverage Error: These errors consist of omissions (undercoverage), erroneous inclusions, duplications (overcoverage) and misclassifications of units in the survey frame). They have an impact on every estimate obtained by the survey, so they are one of the most important types of error.
- Measurement (Response) Error: It is the difference between the recorded response to a question and the "true" value and may occur due to:
  - Respondent errors (e.g. wrong recollections, tendencies to provide 'socially desirable' answers, etc.)
  - Poor instrument design (e.g. technical jargon, or confusing/poorly worded questions resulting in misunderstandings)
  - Interview bias (e.g. inadequate interviewer training resulting in unprofessional interviewing attitude, such as being too friendly or too distant)
  - Survey process (e.g. using proxy responses (taking answers from someone other than the respondent)
- Non-response Error: There are two types of non-response errors:
  - Partial (item) non-response errors: This type of error occurs when respondent provides incomplete information mostly due to the lack of understanding of the question(s). This is why designing and testing of the survey instruments are crucial steps to reduce bias caused by this type of error.

- Complete (total) non-response errors: This type of errors occur when results fail to include the responses of certain units (respondents). Reasons of this type of error include unavailability of the respondent (for paper/mailed surveys: respondent may have moved from the address in the survey frame or been temporarily absent, and for online surveys: respondent may be unable (bounced back email invitations or wrong e-mail addresses in the survey frame) or refuse to participate in the survey. If several survey units (respondents) do not respond to a survey and if non-response is not corrected properly, this can result in biased survey estimates because it is not uncommon that non-respondents often have different characteristics from respondents, and a significantly high rate of non-response can lead to invalid survey results due to bias. To reduce this type of error, care should be taken in describing the purpose clearly and efficiently, preparing up-to-date survey frames, and conducting sample rather than census surveys to avoid contacting a particular unit (respondent) repetitively for the same or different surveys, which otherwise leads to respondent's refusals to participate due to survey fatigue.

- Processing: Processing transforms survey responses obtained during collection into a form that is suitable for tabulation and data analysis (e.g. from an online survey software tool, e.g. Verint, FluidSurveys, and LimeSurvey to a data analysis tool, e.g. MS Excel, SPSS, SAS, and R). It is a mixture of automated and manual activities, which can be very time-consuming and resource-intensive, and is likely a potential source of error. Processing errors can occur while data is being coded (assigned a numerical value to facilitate data capture and processing in general), captured (transformed into a machine-readable format), edited (checked to identify missing, invalid or inconsistent entries that point to data records that are potentially in error), or imputed (identified and assigned replacement values to resolve problems of missing, invalid or inconsistent data). Similar to others, this type of errors can be random or systematic in nature, expand the variance of the survey's estimates and introduce bias. This group of errors are often monitored and controlled using quality control techniques (For quality control techniques, see *Survey Methods and Practices (2010)*).

Non-sampling errors are likely to lead biased survey results, they, unfortunately, are usually extremely difficult and sometimes impossible to measure. They can only be controlled by anticipating problems before they occur and take necessary steps to prevent or minimise them during the planning and designing phases of a survey.

Since errors can be expensive and difficult to resolve, emphasis should be placed on error prevention in the early stages of the survey (ideally at the planning and design phases). The

survey must also be designed and implemented to ensure that the data is relevant to users and produced in a timely manner and that users are made aware the data is available in a medium they can access.

Examples of quality assurance include:

- Intensive planning
- Conducting a pilot survey
- Training everyone involved in the entire survey process (data capture operators, coders, survey designer, data analyst, etc.)
- Improving the sampling frame
- Improving the survey instrument design
- Modifying the data collection methodology if necessary (e.g. moving from paper-based to web-based collection, or vice versa)
- Better follow-up routines
- Clearer processing procedures

(For further information on survey errors and quality assurance, see *Survey Methods and Practices (2010)*).

7. CONSIDERING INCENTIVES

There is a substantial body of research exploring the effectiveness of incentives in surveys and arguing that monetary incentives increase the survey response rates (See a recent HEQCO report for the extent of the impact of incentives on post-secondary student participation in online surveys).

If there will be a prize draw as incentive, the following or a similar paragraph should precede the personal contact information part in the survey (usually at the end of the instrument).

*To participate in the draw for a chance to win **one of four $50 or two $100 gift cards to Metropolis at Metrotown**, please enter your name and email address below. Prizes are not transferable and must be accepted as awarded with no substitutions in cash or otherwise. Participation in the prize draw is voluntary and you have a choice whether or not to disclose any information. We will use the contact information only to notify the winners and award their prize, and this information will be discarded once the prizes are awarded. You may withdraw your consent to be contacted at anytime by sending an e-mail to [Contact person's first and last name, title/position, e-mail] with a message asking us to remove your name from the draw.*

See Section 9: Protection of Privacy and Data Confidentiality and 9.3.1. Personal Information Provided by Respondents for a Prize Draw" for further information on the security of the data containing direct or identifying personal information.

8. DRAFTING THE SURVEY INVITATION AND REMINDERS

Survey invitation is the first thing your target group of respondents see, so it has an important effect on the survey response rate. It is one of the powerful tools with which you can convince your potential participants to take the survey.

8.1. Survey Invitation

Survey invitations or correspondence should include a statement that clearly explains (they don't have to be in the same order):

- Title of the survey (optional, you can either mention it in the subject bar or within the invitation text)
- Who is the target group to be surveyed (e.g. *first-year full-time students, returning part-time students, staff/faculty of BCIT*)
- The purpose of the survey (describe who is conducting the survey and how the information collected through the survey will be used)
- That participation in the survey is voluntary
- That the respondent can skip questions he or she would prefer not to answer
- Whether responses provided will be treated as anonymous and/or confidential data
- How information from the survey will be reported and used
- The length of time the survey might take to complete
- Statement of incentives
- Timeline for the survey (how long the survey will be available)
- Contact information in case the respondent has questions or concerns
- Option to opt out from all Institute surveys – using the "unsubscribe link"
- Confidentiality statement and if possible links to the relevant privacy policies should be available either in the survey invitation or in the very beginning of the survey instrument (see Section 9. Protection of Privacy and Data Confidentiality).

Here is an invitation sample for your reference:

Dear student:

BCIT is committed to providing you with an outstanding learning experience. As a current part-time student, you have the unique opportunity to give your feedback on the course registration process and the student support services provided at BCIT, as well as on the classrooms, buildings, and grounds. By completing this survey, you will help inform BCIT on its commitment towards improving your learning experience while you are studying here.

Your participation in this survey is completely voluntary and you may skip any questions that you do not wish to answer. All of your responses will remain confidential and be reported only in aggregate form.

The survey will take 10-15 minutes to complete, and upon completion of the survey you will have an opportunity to enter a draw to win [describe the prize].

The survey will be open until midnight on [Exact Date].

If you have any questions or concerns about this survey, please contact [first and last name, title/position/Office address, e-mail].

Many thanks in advance for your participation!

[Survey access link as a hyperlink]

Please note this survey is administered by BCIT's [Department Name] using [Name of the Survey software], all data remains confidential and is stored in Canada.

If you would like to unsubscribe from this email list [Unsubscriber's link as hyperlink]

If you have any questions about BCIT's privacy practices, contact, Cynthia Kent, Associate Director, Privacy, Records Management and Copyright, BCIT Library Services, T 604.432.8508 | E cynthia_kent@bcit.ca |W bcit.ca/privacy.

8.2. Survey Reminders

Students and staff/faculty are very busy most of the time and they may need reminders to complete a survey. Most of the online survey software tools have the ability to filter respondents by response status and send reminders only to those who have not participated in the survey yet. The number of reminders depends on the duration of the survey; one to three are recommended to be sent, preferably with a one-week interval between the original invitation and the reminder(s).

9. PROTECTION OF PRIVACY AND DATA CONFIDENTIALITY

Key Concepts:

9.1. Privacy

"Privacy refers to an individual's right to be free from intrusion or interference by others. It is a fundamental right in a free and democratic society. Individuals have privacy interests in relation to their bodies, personal information, thoughts and opinions, personal communications with others, and spaces they occupy […]. An important aspect of privacy is the right to control information about oneself. The concept of consent is related to the right to privacy. Privacy is respected if an individual has an opportunity to exercise control over personal information by consenting to, or withholding consent for, collection, use and/or disclosure of information" (Panel of Research Ethics).

9.2. Confidentiality

"The duty of confidentiality refers to the obligation of an individual or organization to safeguard information entrusted to it by another. The duty of confidentiality includes obligations to protect information from unauthorized access, use, disclosure, modification, loss or theft" (Panel of Research Ethics).

9.3. Identifiable Personal Information

"Identifiable personal information" means information relating to an individual that could be used to identify or re-identify that individual through a combination of indirect identifiers (such as date of birth, place of residence, or a unique personal characteristic). It includes information about personal characteristics such as age, culture, educational background, employment history, health care, life experiences, religion, social status and other matters where an individual has a reasonable expectation of privacy" (Panel of Research Ethics).

Privacy rights of individuals are protected in federal and provincial legislation and breaches of privacy rights might have serious consequences for the reputation of the institutes and organizations. Similarly, violations of confidentiality may harm the individuals and the trust relationship between the institutes/organizations and the individuals.

The surveying party is liable for compliance with all applicable legal and regulatory requirements concerning the protection of privacy during all stages of surveying (collection, analysis, dissemination, storage, retention, and disposal of information). The surveying party may need to collect and use information that contains identifiable personal information about which a respondent has a reasonable expectation of privacy protection. In this case, the survey owners/conductors are responsible for informing all prospective survey respondents in the survey invitation or in the beginning of the survey itself if their responses will be anonymous, kept confidential or non-confidential, and aggregated (reported as group data) or non-aggregated, or if data even in aggregate form may be disclosed and pose a risk to the respondent. Anonymous data does not include any direct identifier such as name, social insurance number, personal health number, or student ID number. Once respondents are informed that their responses would be confidential, the reported data must not include any identifiable personal information, alone or by linking it to other available information that would make it possible to associate a response with any given respondent.

Sample of a Privacy Statement IR uses in the Institutional Surveys:

Privacy Statement

The collection of any personal information you provide is permitted in accordance with section 26(c) of the Freedom of Information and Protection of Privacy Act (FIPPA), which allows BCIT to collect personal information for purposes related to and necessary for its operating programs and activities. Any information that you provide by answering the survey questions will be used on an aggregated (not individually identifiable) basis for analysis and reporting, and will be provided anonymously. Please do not provide any information that would identify yourself or others. If you have any questions about BCIT's privacy practices, contact, Cynthia Kent, Associate Director, Privacy, Records Management and Copyright, BCIT Library Services, ***T*** *604.432.8508 |* ***E*** *cynthia_kent@bcit.ca |**W** bcit.ca/privacy.*

If you have any questions or concerns about this survey, please contact [Analyst's first and last name; Position/Title; Office address; e-mail address].

The party conducting the survey must protect the information entrusted to them. Storage of and access to the relevant data files containing identifiable personal information must be handled and managed with great care (e.g. completed paper survey instruments must be kept in a secure locked location/filing cabinet, while computer files of completed web-based surveys should be password-protected and encrypted. Data should never be stored or downloaded onto an unsecured computer). Institute-related data creation, collection, storage, maintenance, use, dissemination, and disposal must comply with relevant laws, regulations, policies, procedures, guidelines, and ethics (for further information see External Links).

9.3.1. Personal Information Provided by Respondents for a Prize Draw

For surveys that offer incentives to participate, respondents must provide contact information. If identifying information such as names or e-mail addresses are kept with the survey responses and confidentiality is promised to respondents, the contact information data should be safeguarded by taking all the appropriate security measures (see above) and discarded as soon as the winners are notified and awarded their prize.

For example, IR selects the prize winner(s) by random draw from among all survey respondents the first day of the survey closure, selecting a few alternates in case the first winner cannot be contacted. Once all of the prizes are awarded, all relevant contact information is removed from the data file prior to the analysis of the survey responses.

Canadian Anti-Spam Legislation (CASL) applies to surveys offering incentives. CASL applies to an electronic message if it contains any commercial element that is not exempted from the scope of the legislation. A prize draw or contest is considered "commercial activity" that is not related to core activities of BCIT. If a prize draw is included with a survey message, the entire message is considered to be a Commercial Electronic Message (CEM) subject to CASL.

All CEMs must contain the following information:

- Office/branch within BCIT that is sending the message
- Contact information for the office/branch including mailing address and telephone no. and/or email address and/or website address
- Unsubscribe option (i.e. information about how to unsubscribe from receiving future CEMs; this could be by sending an email with the request to be taken off the mailing list).

The message should include a statement indicating that personal information is being collected for the purpose of administering the draw (For information on "Incentives," see Section 7. Considering Incentives).

10. ANALYSIS OF SURVEY DATA

After exporting the survey data from the online survey software to a data analysis tool, such as Excel or SPSS, the file should be checked for any anomalies, and all collected data should be cleaned. Text responses are checked for any obscene or abusive language (e.g., negative comments or connotations about the instructors or classmates) as well as the apparent or potentially identifiable personal information. The checking/cleaning of the data is usually done by the analyst (for more information on "identifiable personal information," see Section 9.3. Identifiable Personal Information).

If the paper-based survey instruments are to be sent to a contracted data entry company all data should be cleaned beforehand, and a code sheet needs to be prepared. After the data entry file is received back, the file should be checked for any anomalies before starting the analysis.

10.1. How Well Do Your Sample Results Correspond to the Opinions of the Targeted Student Population

When we analyze the survey data we tend to assume that our sample results are similar to those we would get if the entire population (every student in our target population) was surveyed. Suppose, for instance, we surveyed 4000 randomly-selected part-time studies students and found that 65% of them are "satisfied" with the current registration process. We would like to be able to say that 65% of *all* students at BCIT are satisfied with the registration process. But can we say this? Unfortunately, no! It's very unlikely that *exactly* 65% of the entire part-time student population is satisfied with the registration process. It's possible that the overall percentage is 66%, 65%, maybe even 72%. This discrepancy between 65% and the true percentage is called **sampling (sample) error**. In fact, it is not an "error"; it occurs because even a good random sample is unlikely to match *precisely* the entire population (for further information on "sampling error," see Section 6.3. Survey Errors and Quality Assurance; also see Suskie, 1996; Henry, 1990; Linneman, 2011, Heeringa, West, and Berglund, 2010; and Statistics Canada, Survey Methods and Practices (2010)).

When reporting results, you will increase the credibility of your report if you mention the possible sampling error of your findings. For example, instead of simply saying "65% of students are satisfied with the registration process", say, "65% of students who are satisfied with the registration process *with an error margin of plus or minus 4%.*" Statistically, this means you are 95% sure that between 61% and 69% of students are satisfied with the registration process. In other words, your statistic will be within 4% points of the real population value

10.1.1. Margin of Sampling Error and Its Calculation

As mentioned in Section 6.3. Survey Errors and Quality Assurance, survey errors are grouped into "sampling" and "non-sampling" errors.

Before talking about the calculation of the margin of sampling error (margin of error for a sample proportion), we should distinguish the term **margin of sampling error** from the **margin of error**. While it is not uncommon to use these terms interchangeably, they actually refer to different statistical concepts. While the former is associated only with sample surveys and measurable, the latter is associated with both sample and census surveys/polls and harder to measure as it cannot be quantified. As mentioned in Section 6.3. Survey Errors and Quality Assurance, non-sampling errors arise during all survey activities but sampling. The sources of these errors may range from how the questions or response categories were worded to mistakes in recording and coding the data and to response errors.

There is no perfect match between the sample survey results and the "true" values that we would obtained if everyone in our population would provided responses to our questions. A margin of sampling error tells us how close we can expect the value of our survey data to be to the real (true) population value. For example, a 95% confidence interval with a 4% margin of sampling error means if we conducted the same survey 100 times, we would expect the result to be within plus or minus 4 percentage points of the true population value 95 of the time.

When we report the survey results, we need to include the margin of sampling error. The general formula for the margin of sampling error is:

$$ME = z\sqrt{\left(\frac{p(1-p)}{n}\right)}$$

Where p is the sample proportion and n is the sample size and z is the appropriate z-value for your preferred level of confidence (for example, for the 95% confidence level, the z-value is always 1.96 and for the 98% confidence level it is 2.33). [Please note: A sample proportion is the decimal version of the sample percentage. In other words, if you have a sample percentage of 7%, you must use 0.07 in the formula, not 7. To change a percentage into decimal form, simply divide by 100. After all your calculations are finished, you can change back to a percentage by multiplying your final answer by 100%.]

Here is an example:

Suppose that the BCIT Institutional Research Office's latest Part-time Studies student survey was conducted to 1,000 students selected by random sampling, and the results show that 520 students (52%) indicated that their learning objectives had been met, compared to 48% who don't think so. First, assume you want a 95% level of confidence, so $z^* = 1.96$. The number of students in the sample who said their learning objectives were met was found to be 520. This means that the sample proportion (p) is 520 / 1,000 = 0.52. (The sample size, n, was 1,000.) The margin of error for this survey question is calculated in the following way:

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{(0.52)(0.48)}{1{,}000}}$$

$$= (1.96)(0.0158) = 0.0310$$

According to this data, you conclude with 95% confidence that on average 52% of all part-time students think that their learning objectives were met, plus or minus 3.1%.

Please note that in order to use a *z*-value in the formula for the margin of sampling error, two conditions need to be met:

1. *np* must be at least 10 (in our case: 1,000 x 0.52 = 520)

2. n (1-p) must be at least 10 (in our case: 1,000 x 0.48= 480)

As most institutional surveys conducted at BCIT are based on hundreds or even thousands of students, meeting these two conditions will not be a problem.

10.2. Summary Measures:

Data analysis involves summarising the data in a meaningful way in tables and charts/graphs and interpreting it in a way to provide clear answers to the questions that necessitated the survey in question to be conducted. Most of the time, it includes tables and charts displaying various summary measures such as:

- **frequency distributions** (a table that shows each category of a variable and the frequencies within each of these categories)
- **measures of central tendency/location** such as the mean (arithmetic average: the sum of all the values of a variable divided by the number of values), the median (middle value of a set of data arranged in numerical order), and the mode (most frequently occurring data value in a distribution)
- **measures of spread** such as range (the difference between the largest and smallest data values of a variable) and variance (an estimate of average variability (spread) of a set of data. It is the average squared deviation from the mean, and useful for comparing the amounts of variation among various groups) for each variable.

Cross-tabulations of responses across subsets of respondents can also be useful for some users of the report (cross-tabulation is a table that crosses two variables over one another so that we can see if there is a relationship between them, e.g., cross-tabulation of students' gender and the type of student services they use) (for further information, see Statistics Canada, Survey Methods and Practices (2010), and Linneman, 2011)

10.3. Major Types of Statistical Procedure

Data analyst may just want to describe (summarize) the observed units in a given data set, which can be either a representation of the entire population or a sample of it, in clear summary form (**descriptive statistics or summary statistics**). Sometimes, that's all data users want to hear: What's the average highest education level of BCIT's part-time students; what percentage of students take online courses.

There are a bit more complicated procedures that allow us to speak beyond a small sample of people, or to say some pretty specific things about very large populations, even though we haven't surveyed everyone in these populations (**inferential statistics**). Inferential statistics can be preferred when surveying everyone in the target population is not possible. For instance, while we can survey every student on the campus to know exactly what percentage of students lives on campus (please note that this is also not ideal most of the time due to survey fatigue), it is not possible to ask every single Canadian their attitudes about the legalization of Marijuana; thus we sample instead (Linneman, 2011).

It is not uncommon that descriptive statistics raise interesting causal questions: Does where you live (on campus or off campus) affect your academic performance? Among which demographic groups has there been increasing support of marihuana legalization? All of these imply causality: something causes higher grades; something causes attitudes to change. With this set of statistical procedures (**explanatory statistics**), we don't just have variables, we have independent and dependent variable. The **dependent variable** is what we are trying to explain. Why do some students get higher grades than others? Does their residential status have an effect on their grades? That is, do grades depend on residential status? That is why, in this situation, we would call grades the dependent variable. In this situation, we are saying that residential status is not dependent on anything; therefore we call it the independent variable. The **independent variable** is the variable we are using to explain why the dependent variable varies: we think that grades vary by residential status. The dependent or independent nature of a variable relies on the situation we are exploring. While "grades" are the dependent variable in the previous situation, it could be used as the independent variable **to explain why** some students pursue certain types of careers (Linneman, 2011).

While cross-tabulation provides more detailed information and allow comparisons with other data sources, the analyst must take all security measures to mitigate privacy risks and ensure the confidentiality of respondents. IR analysts suppress/mask the data in table cells that risk revealing information about a respondent. As a best practice, they do not report result categories/cells in cross-tabulations containing five or fewer respondents. In the instance of masking results, a symbol (~) is used to draw the reader's attention to this practice.

11. DISSEMINATION OF SURVEY RESULTS

One of the main methods of disseminating survey results is a paper or digital survey report with tables and charts. The report should also include a description if the survey methodology along with measures (summary measures, measures of central tendency, and measures of spread if applicable) and sources of sampling and non-sampling error (For sampling and non-sampling error see Section 6.3. Survey Errors and Quality Assurance). Before the report is disseminated, data must be examined to ensure that it does not raise any risk of violation of respondents' privacy and confidentiality of their responses (for further information on "data suppression/masking," see Section 10. Analysis of Survey Data).

BCIT's IR publishes all its institutional reports on the IR portal (accessible to the entire BCIT community) and Tableau Server (BCIT Tableau Server is currently in the first phase of implementation and only a small number of individuals within the BCIT community has access to it; more and more users will be ultimately added). For the surveys reported to a specific department (e.g., Library Services, Housing, or Recreation Services), it's that department's discretion to publish their results. IR cannot share a specific department's survey results with others without prior permission. Remember that survey results deemed as "research" cannot be presented or published beyond BCIT without Research Ethics Board (REB) approval (for further information on REB, see the Section on the "Research Ethics Board").

Include the followings in any report of survey results:

- The ownership and undertaker: who sponsored the survey, and who conducted it.
- The survey instrument, including any preceding instruction or explanation to the respondents that might practically be expected to affect the response.
- A definition of the target population the survey is intended for.
- A description of how the respondents were selected by the researcher/analyst; including eligibility criteria and screening procedures, if applicable.
- Sample size and response rates.
- Method of data collection:
    - self-administered (paper survey instruments that are mailed and web surveys that are implemented online)
    - interview-based (face-to-face, including focus-groups, or telephone)
    - Mixed mode method: e.g. self-administering the sensitive questions and continuing with a face-to-face interview; or collecting information with a self-administered survey and following up with a focus-group, if applicable,
- Dates/period of data collection

- Information on the precision of the findings, including estimates of sampling error, and a description of any weighting of estimating procedures used, if appropriate (see Section 10.1.1. Margin of Sampling Error and Its Calculation).
- A fair, objective, and thorough presentation of the results without any censorship.
- Appropriate attributions of the work and ideas of others, if applicable.
- Promote the use multiple sources of information about persons, programs, or services.
- Provide all necessary information for the reader to avoid misinterpretations (such as information on the suppressed data, symbols and abbreviations used, etc.)

12. POST-SURVEY ASSESMENT

After completing the survey project entirely, the survey designer/analyst assesses the effectiveness of the survey by reviewing, for instance, response rates, questions with highest dropouts (to assess if there was something wrong with the question wording or choices/response categories, etc.), or outcomes of a new survey methodology, to shed light on the improvements of future projects. In 2015, for example, IR started to use proportionate survey sample: staff proportionated the survey cohort by program type, i.e., Trades and Technology and program mode (part-/full-time), and then stratified it by School. IR determined this new methodology reduced the sampling error and improved the representativeness of the sample.

IV.OTHER CONSIDERATIONS

1. RESEARCH ETHICS BOARD REVIEW MAY BE NECESSARY

If you plan to make general conclusions about the results of your survey for presentation or publication outside BCIT, your survey may be considered research under federal guidelines. If so, you should get approval from the BCIT Research Ethics Board (REB) before you begin your survey. The REB is responsible for monitoring all research at BCIT that involves human participants. Federal regulations require research with people to meet essentially four criteria:

1. Risks to participants must be minimized
2. Risks to participants must be balanced with the benefits
3. Selection of participants must be impartial
4. Informed consent must be sought where appropriate

If your study involves people, it should be reviewed and approved by the REB prior to commencing your research. To get your study reviewed, you should complete a standard REB form and provide a copy of your survey and any letters or email messages that will be sent to people you will invite to participate in your study. If you are unsure on how to proceed, please

contact the chair of the REB. For more information, please visit the BCIT Research Ethics Board (REB) homepage.

Individuals and organizations who are not affiliated with BCIT and who would like to invite members of the BCIT community to participate in a research project must first contact the REB. The REB serves as the primary point of contact for all internal and external research proposals that involve BCIT students, faculty and/or staff as investigators or participants.

2. SELECTING THE ONLINE SURVEY SOFTWARE OR SUPPLIER

To ensure confidentiality and avoid privacy breaches, the collected data should be stored in Canada, thus the online survey tool used for designing the survey instrument should be housed in Canada; otherwise, a legal document called a "data agreement" with the vendor may be needed. Survey conductors/owners should work with the Information Access and Privacy Office of BCIT to ensure that all requirements are met and detailed in the application.

V. REFERENCES

TCPS2 2014, *Tri-Council Policy Statement, Ethical Conduct for Research Involving Humans*

Statistic Canada (2010), *Survey Methods and Practices*, Minister of Industry: Ottawa, Ontario

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. (4th ed). London: Sage

Gary T. H. (1990). *Practical Sampling*. Newbury Park, CA: Sage

Heeringa, S. G., West, B. T., & Berglund, P.A. (2010). *Applied Survey Data Analysis.* Boca Raton, FL: Chapman & Hall

Light, R. J., Singer, J. D., & Willett J.B. (1990). *By Design.* Massachusetts: Harvard University Press

Linneman, T. J. (2011). *Social Statistics: the basics and beyond,* New York: Routledge.

Suskie, L.A. (1996). *Questionnaire Survey Research: What Works.* (2nd ed.). Tallahassee, FL: Florida State University

Vogt, P. (1993). *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences.* Newbury Park, CA: Sage.