Examination of technology-enhanced statistical problem-solving tasks designed by pre-service teachers

Nadide Yılmaz
Karamanoglu Mehmetbey University, Turkey

In this study, technology-enhanced statistical problem-solving tasks designed by pre-service teachers (PTs) were examined. The PTs designed 28 tasks. The designed tasks were analyzed within the context of the Considerations for Design and Implementation of Statistics Tasks (C-DIST) components. It was revealed that the tasks were mostly designed within the framework of the learning goal of “statistical questions-making interpretations based on the measures that serve to represent the data and the forms of representation” and that mostly real, multivariate and large data sets were used. In addition, it was observed that the context was employed in order to complete the prepared tasks and the tasks mostly included the entire investigation cycle. It was determined that the prepared tasks were mostly at Level B, followed by the tasks at Level A and Level C. In light of the results obtained, inferences were made for preparing PTs to teach statistics.

Keywords: task design, technology, statistical problem solving, pre-service teachers

1 Introduction

Individuals need to learn how to read and analyze data, because data are encountered in all areas of life and it is necessary to make decisions based on data (Bargagliotti et al., 2020; Boaler & Levitt, 2019; Wild et al., 2018). This emphasis has found reflections in curricula (e.g., Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015; Ministry of National Education [MoNE] (2018) and statistics has taken its place as a learning area in the curricula. Institutions giving direction to statistics teaching and curricula emphasize that students should experience statistical problem solving (SPS) in the teaching process (Bargagliotti, et al., 2020; Franklin, et al., 2005; MoNE, 2018). Undoubtedly, the tasks in the teaching process play a key role in helping students acquire these targeted statistical skills (da Ponte, 2011; Franklin, et al., 2015; Shaughnessy, 2007).

Students often encounter statistical tasks in textbooks in the teaching process (Braswell et al., 2005). However, it has been observed that the tasks in the textbooks mostly focus on the analysis of data rather than SPS (Balci, 2023; Bargagliotti et al., 2020; Jones & Jacobbe, 2014; Jones et al., 2015) and take smaller datasets into their centre (Weiland, 2019). However, in real life, students need to work with large data
sets (Casey et al., 2021). At this point, the tasks prepared by teachers play a key role (Bakogianni, 2015; Dierdorp et al., 2011; Garfield, 1995; Shaughnessy, 2007). However, studies show that teachers have difficulties in designing statistical tasks (Bakogianni, 2015; Casey et al., 2020; 2021; Chick & Pierce, 2008; Rossman et al., 2006). In the current study, it is aimed to examine the technology-enhanced SPS task design assignments designed by PTs who will be the teachers of the future.

2 Literature Review

In this section, first, the features of high-quality SPS tasks are mentioned. Then, studies focusing on the features of statistical tasks designed by mathematics teachers/PTs are presented.

2.1 Features of high quality SPS tasks

High quality statistical tasks in teaching processes should be designed in such a way as to allow students to experience SPS (Bargagliotti et al., 2020; Burgess, 2011; Franklin et al., 2007). SPS consists of the following stages; formulating a statistical question, collecting or considering data, analyzing data and interpreting the results. (see. Figure 1).

![Figure 1. The relationship between technology and statistical problem solving (Adapted from Bargagliotti et al., 2020).]
At the stage of formulating a statistical question the problem is defined clearly and the problem is shaped in a way that pays attention to variability and real-world context (Cobb & Moore, 1997; Franklin et al., 2005; Scheaffer, 2006). Collecting or considering data includes planning for the data to be collected or considered and implementing this plan (Franklin et al., 2005). Another point to be considered is that the data sets collected or selected should reflect the nature of daily life (Bargagliotti et al., 2020; Lee, 2019). At the third stage, analyzing data, the analysis of the data takes place by choosing the appropriate methods (Franklin et al., 2005). As at every stage, context should be taken into account when deciding on the appropriate data analysis method (delMas, 2004). Representing the data with different means of representation - in other words, transnumeration - or creating different meanings by interpreting a representation according to different perspectives takes place in this process (F. Curcio, 1987; F. R. Curcio, 1989; Shaughnessy, 2007; Wild & Pfannkuch, 1999). At the last stage, interpreting the results, the results are interpreted by considering variability and context and these results are associated with the initial research question (Bargagliotti et al., 2020; Franklin et al., 2005). While reading the data can reveal the information appearing on the graph, in reading between [or within] the data, the relationships between different components or data points can be defined. If inferences or generalizations are to be made based on data, reading beyond the data is required (F. Curcio, 1987; Friel et al., 1997). Reading behind the data can be performed to reveal contextual explanations about the trend in the data (Shaughnessy, 2007; Shaughnessy et al., 1996). In addition, attention is drawn to the use of technological software in order to make the SPS more effective (Bargagliotti et al., 2020) because technological tools allow creating graphical representations and producing numerical summaries of data and simulations. This paves the way for conceptual understandings to be placed in the centre by focusing more on statistical concepts and data (Bargagliotti et al., 2020; Franklin et al., 2005; 2015). Researchers working in this field have prepared a framework that reveals how the tasks to be used and implemented in the teaching process should include the SPS (Tran & Lee, 2015).

2.2 Features of the statistical tasks designed by mathematics teachers/PTs

Tasks plays a decisive role in the conduct of statistics teaching as targeted (e.g., Dierdorp et al., 2011; Shaughnessy, 2007). Attention is drawn to the need for PTs to prepare high-quality statistical tasks in mathematics teacher education (da Ponte,
Differences between mathematics and statistics, the role of context in statistics and the differences in the process of interpreting data make the preparation and implementation of statistical tasks even more critical (Bakogianni, 2015; Rossman et al., 2006). However, it has been revealed that there are limited studies on how teachers/PTs prepare statistical tasks, and in these limited studies, it has also been revealed that teachers/PTs have difficulties in designing statistical tasks (e.g., the role of context) (Bakogianni, 2015; Casey et al., 2020; Chick & Pierce, 2008; Rossman et al., 2006). Chick and Pierce (2008) gave data sets to PTs and asked them to create statistical questions and a hypothetical lesson plan for sixth graders using these data sets. It was observed that the majority of the PTs (81%) created questions asking to simply read or interpret the information shown by a table or graph, while less than half (41%) of the questions they prepared were aimed at identifying information that is not immediately visible from the data or making inferences from the data. More than half of the lesson plans prepared by the PTs directed the students to the data set. However, few of the lesson plans (23%) included the continuous and effective use of the data set to reveal statistical concepts, while the other lesson plans did not include the data set in a meaningful way. Casey et al. (2020) examined the tasks prepared by PTs and concluded that most of the tasks prepared by PTs contained large, multivariate, real datasets and allowed making associations with the context. In addition, the attempts of PTs to structure the tasks in a way that allows for a SPS have attracted attention. Casey et al. (2021) focused on the strengths of the tasks designed by PTs and the aspects that need improvement. They stated that the strengths of the tasks designed by PTs included the use of large and multivariate data sets, constant connection with the context and students’ involvement in many parts of the SPS and their use of multiple data presentations. The aspects of the tasks that need improvement were expressed as adopting a mathematical approach instead of a statistical approach, focusing on ambiguous questions or numerical calculations as well as issues related to statistical content. Bakogianni (2015) focused on the stages of mathematics teachers preparing, implementing and reflecting on statistical inquiry tasks. Statistical context has proven to be not only an elusive learning goal, but also a significant teaching challenge. It has been determined that teachers’ familiarity with the content and teaching of statistics, students’ prior statistical knowledge, classroom reality problems and the stochastic context of statistical problems affect the preparation and implementation of the tasks and reflection on them. Collaboration and interaction among teachers provided the opportunity for teachers to gain a deeper understanding.
of statistical concepts and procedures, to identify specific learning objectives related to statistical content, and to be aware of learning difficulties associated with them. The results show that PTs have various difficulties in designing statistical tasks.

3 Rationale of the study

Targeted statistics education “depends to a large extent on the teachers who will bring them to life in the classroom” (Franklin et al., 2015 p.1). It is necessary for teachers to have many skills, such as being able to plan, conduct and evaluate the teaching of statistical concepts in the classroom environment (Franklin et al., 2007; 2015, Groth, 2007; 2013). One of the important factors determining the effectiveness of students’ learning is the tasks used (Carver et al., 2016). It can be said that teachers generally tend to use the tasks from various sources (e.g., textbooks, online) (Casey et al., 2020; Shapiro et al., 2019). However, it has been revealed that some tasks in these resources focus on the calculation of concepts and analyzing the data and include small data sets, instead of dealing with SPS in its entirety (Balcı, 2023; Jones & Jacobbe, 2014; Jones et al., 2015; Weiland, 2019). Another emphasis is that statistical data are intertwined with technology and that the tasks used and the SPS should be integrated with technology when feasible (Bargagliotti et al., 2020). Many studies that draw attention to the fact that technology is one of the important components that affect the quality of statistics education also support this (Garfield, 1995; Lee et al, 2014; Neto, 2017; Suhermi & Widjajanti, 2020; Tishkovskaya & Lancaster, 2012). Seen from this perspective, it becomes important for teachers to have the necessary knowledge and skills on how to create high-quality tasks (Casey et al., 2020; 2021). Researchers pointing out that the reasons for the difficulties experienced by teachers / PTs should be examined in depth state that the lack of knowledge of PTs about SPS has the potential to affect the structure of the tasks they prepare (Bakogianni, 2015; Casey et al., 2020; 2021; Chick & Pierce, 2008). PTs’ design statistical question (Burgess, 2007; Leavy & Frischemeier, 2022), data collection (Hannigan et al., 2013; Lovett & Lee, 2018), data representation or interpretation (Casey & Wasserman, 2015; Hannigan et al., 2013) may also affect the prepared tasks. In other words, if PTs have difficulties in carrying out the SPS, it is likely that this will affect the statistical tasks they prepare (Casey et al., 2020; 2021). Seen from this perspective, the examination of the statistical tasks prepared by PTs has the potential to provide mathematics educators with important information about PTs (Casey et al., 2020; 2021; Chick & Beswick, 2018). With the current study, it is thought that it will be revealed which points PTs can easily deal
with while preparing statistical tasks and at which points they need improvement. On the basis of the results obtained, mathematics educators can better organize the content of undergraduate courses and more effectively conduct these courses (e.g., teaching statistics and probability, statistics). In addition, another point that should be taken into consideration is that limited research has been conducted in this field. The existing research has largely focused on the development of the content knowledge of PTs (Peck et al., 2013; Perkowski & Perkowski, 2007). In this context, it is thought that the results obtained will contribute to the literature. Thus, an answer to the following research question was sought: How are the technology-enhanced SPS task design assignments designed by PTs?

4 Method

4.1 Research design and participants

Since the purpose of the current study was to examine the structure of the statistical task design assignments prepared by PTs within the scope of a teacher education program, the case study design was employed. Case study allows obtaining and examining in-depth information about the case of interest in line with the research problem (Merriam, 2009; Putney, 2010). In the current study, it was aimed to examine the task design assignments prepared by PTs. To this end, the unit of analysis of the study was determined as 28 task design assignments prepared by 56 PTs participating in the study in groups of two.

4.2 Context of the study and data collection

This study was carried out in the Department of Mathematics at a state university in a city located in the Central Anatolian region of Turkey. PTs who graduate from this department can work as a mathematics teacher at the middle school level (11-14 years old) of public or private institutions. The program is a four-year program and the language of instruction is Turkish. In the first two years of the four-year program, PTs mainly take content knowledge courses (e.g., Analysis, Algebra) and in the last two years, they mainly take pedagogical content knowledge courses (e.g., Teaching practice, Teaching numbers). This study was conducted within the scope of the “Probability and Statistics Teaching” course, which is a compulsory course to be taken in the sixth term of the mathematics teaching program. The course mainly focused on the
teaching of probability and statistics concepts and aimed to improve PTs’ knowledge
and skills about these concepts and how this knowledge could be reflected in the classroom environment. The course also placed the SPS into the centre of the subjects taught (Bargagliotti et al., 2020). In addition, approaches to teaching statistics and how to implement an effective statistics teaching (Ben-Zvi, 2011; Cobb & McClain, 2004) were discussed with the PTs. In the following weeks, each stage of SPS (formulating a statistical question, collecting or considering data, analyzing data and interpreting the results) was handled. In addition, the technological software that supports the teaching of statistics and how this software could support the teaching process were discussed. Common Online Data Analysis Platform (CODAP) (http://codap.concord.org), web-based educational software, was preferred because it is free and accessible, and students were informed about how to use this software. Furthermore, sample tasks were examined (Concord Consortium, 2019). In the 10th week of the lesson, the PTs were asked to design a statistical task design assignment using CODAP in such a way as to develop students’ statistical thinking as a group. This task design assignment was requested to consist of two main parts (Figure 2).

1) Introductory Information

Brief Summary of Task:

Learning goal of the task you have prepared:

Target audience:

Time required:

Materials:

CODAP Link: (Add a link to your data and analyses)

2) Component of a Statistics Task (While structuring your task, suppose that students are solving it)

Posing Statistical questions (e.g., What is your question? What did you pay attention to while posing?)

Data Collection (e.g., How did you obtain the data?, From whom did you collect the data?)

Data Analysis (e.g., How did you analyze the data you collected/used? What representations, measurements (e.g., central tendency, dispersion) did you use?

Interpreting the results (e.g., What do the results tell us?)

Figure 2. Task design assignment
These parts are defined as (1) task summary, target audience, required time, materials, learning goal, CODAP link and (2) Component of a Statistics Task (data, context, SPS). The PTs were left free to use whichever dataset they wanted to use while designing their tasks. It was stated that they could create the data themselves if they wished, or they could use ready-made data sets. The PTs were also said that they could benefit from various websites (e.g., Turkish Statistical Institute (TUIK), Sample CODAP Datasets). In addition, it was explained to the PTs that they could get support from various sources (e.g., curriculum, academic resources) while preparing task design assignments. The PTs designed a total of 28 tasks.

4.3 Data analysis

The task design assignments developed by the PTs were analyzed in the context of C-DIST components developed by Tran and Lee (2015, p.1–2). The PTs focused on preparing the tasks, they did not engage in task implementation. Therefore, the task design assignments prepared by focusing on the components in the framework of “Considerations for Written Task” were analyzed. These components are shown in Table 1.

<table>
<thead>
<tr>
<th>Component of a Statistics Task</th>
<th>Questions to Consider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Goal</td>
<td>What learning goals does the task aim for students to accomplish? Does the task focus on answering questions that are statistical or mathematical? e.g., Does the task ask students to use computations or graphs? Are these in support of analyzing data to make a decision? or is the use of an algorithm or creation of a graph the focus?</td>
</tr>
<tr>
<td>Data</td>
<td>Does the task call for the use of data (either to collect or use already collected data to answer)? Does the data appear to come from a real source?</td>
</tr>
<tr>
<td>Context</td>
<td>Is context a salient part when solving the problem? Is the context likely to be of interest to the students engaging in the task?</td>
</tr>
<tr>
<td>SPS</td>
<td>Does the task address only one phase of a SPS, some phases, or all phases of the cycle? Consider the appropriate phases below as applicable to the intent of the task:</td>
</tr>
<tr>
<td>Pose</td>
<td>Is the question already posed (by teachers, or curriculum developers) or do students have opportunities to pose statistical questions based on their interest? What type of variability does the task attend to?</td>
</tr>
<tr>
<td>Collect</td>
<td>Does the task offer opportunities for students to plan to collect data: sampling, sample size, attribute, and measurement? Do students conduct the data collection? Does the task provide a context so that students are aware of the measurement issues and how data were collected?</td>
</tr>
</tbody>
</table>
Each component in Table 1 was detailed within itself and criteria were created. First, two main criteria were determined for the first component, the learning goal. It was questioned which learning objectives the prepared task aimed to make students accomplish. If the learning goal focused only on graphing (e.g., represent this data with a line graph) or just doing calculations (e.g., what is the mode of this data?), it was coded as a mathematical question. Learning goals where a context was used and statistical situations were required to be evaluated were coded as statistical questions. Statistical questions, on the other hand, were evaluated within the framework of two subcomponents. If the question asked for interpretation by using the measures and types of representation that serve to represent the data, that question was coded as “Interpretation based on the measures and types of representation used to represent the data-statistical question”. If the question asked for making inferences by using the measures and types of representation that serve to represent the data in the question, that question was coded as “Making inferences based on the measures and types of representation used to represent the data-statistical question”.

The prepared tasks were analysed in terms of the characteristics of the data and the need for data in the context of the data component. If the data used were directly collected from a real source, they were coded as “Data come from real source-primary data”. If the data were obtained from a real source (e.g., websites, OECD), they were coded as “Data come from real source-secondary data”. If the data were fabricated by PTs, they were coded as “Data come from hypothetical”. Another point examined was whether data were needed to complete the prepared task design assignments. In this context, the “Use of data to complete the task” code was created.

In terms of the context, what the context addressed contains was analyzed. In this connection, based on the contexts provided in the task design assignments prepared by PTs, components such as “Contexts related to the students themselves”, “science”, “health”, “social” and “education” were created and the task design assignments were examined within this framework. In order to determine whether the contexts in the task design assignments were contexts that could capture the attention of students,
contexts found in the textbooks and reference books commonly used by students were taken into account. In addition, the contextual information contained in the contexts was also examined in order to reveal in more detail how the tasks are related to the data’s context. Contextual title, contextual attribute names, multivariate nature of datasets were discussed in this regard. In addition to this, task design assignments were analyzed according to including/not including information about the source of the data. In other words, it was analyzed whether the source of the data was included in task design assignments. Each task design assignment prepared by the PTs was evaluated according to these criteria. While evaluating the tasks in terms of SPS components (pose, collect, analyze, interpret), the framework created by Bargagliotti et al. (2020, p.16–19) was adopted. This framework described each component at Level A, B and C. The analysis also included the examination of the shared CODAP link associated with the task design assignments. In this way, the types of data representations involved in the link and whether information on how to use CODAP or the data source was included were determined. After the data were coded, another researcher was asked to code the data independently. After the two codings, the researchers came together and analyzed the task design assignments in the context of the components and the points of disagreement were discussed until a consensus was reached.

5 Findings

5.1 Learning goal

The task design assignments prepared by the PTs were evaluated in terms of learning goals and Table 2 was created.

<table>
<thead>
<tr>
<th>Learning goal</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematical calculation/graph construction-mathematical question</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Interpretation based on the measures and types of representation used to</td>
<td>23</td>
<td>82%</td>
</tr>
<tr>
<td>represent the data-statistical question</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making inferences based on the measures and types of representation used to</td>
<td>4</td>
<td>14%</td>
</tr>
<tr>
<td>represent the data-statistical question</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It was observed that the prepared task design assignments mainly included the learning goal of “statistical questions- Interpretation based on the measures and types of representation used to represent the data” (82%). For example, the learning goal in
one of the tasks was expressed as “determining whether the use of the left or right hand is related to gender”. Only in one task, the learning goal of “mathematical questions-mathematical calculation/graph construction” was at the forefront. The learning goal of this task was determined as “calculating the average lifespan of mammals”. In four of the tasks prepared by the PTs, the learning goal was “Making inferences based on the measures and types of representation used to represent the data-statistical question”. For example, the learning goal in one of these tasks was stated as “estimating the foot length according to the height of individuals”.

5.2 Data

The task design assignments prepared by the PTs were evaluated in terms of learning goals and Table 3 was created.

Table 3. Features of the data in the tasks

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data come from real source-primary data</td>
<td>13</td>
<td>46%</td>
</tr>
<tr>
<td>Data come from real source-secondary data</td>
<td>14</td>
<td>50%</td>
</tr>
<tr>
<td>Data come from hypothetical</td>
<td>1</td>
<td>4%</td>
</tr>
</tbody>
</table>

In half of the task design assignments prepared by the PTs, data collected from real secondary sources were preferred. These data sets were obtained from TUIK data, sample data sets in CODAP, and various internet sites. In 46% of the tasks, the data consisted of primary sources collected by the PTs themselves. In only one task, it was determined that the PTs created the data themselves, that is, they obtained hypothetical data. In addition, it was observed that all the designed tasks required the use of data. Then, the number of data used in the tasks was analyzed. In the tasks, while the minimum number of cases was 27 the maximum number of cases was 900. The mean of the number of cases was 274, and the median was 218. Based on these results, it can be said that they preferred to use large data sets in most of the tasks. It was also noted that both categorical and quantitative (numerical) variables were included in task design assignments.
5.3 Context

The contexts of the task design assignments prepared by the PTs was evaluated in terms of their being interesting/suitable for the level of the students and the consideration of the context in the process of solving the problem and Table 4 and Table 5 was created.

Table 4. Features of the context in the tasks

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contexts related to the students themselves (height, favourite subjects, number of siblings..)</td>
<td>15</td>
<td>54%</td>
</tr>
<tr>
<td>Science (atomic radius, mammals...)</td>
<td>5</td>
<td>18%</td>
</tr>
<tr>
<td>Health (Covid 19, heart attack..)</td>
<td>4</td>
<td>14%</td>
</tr>
<tr>
<td>Social (seasonal workers, population, tobacco use..)</td>
<td>3</td>
<td>11%</td>
</tr>
<tr>
<td>Education (budget in education...)</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>

It was determined that all the task design assignments prepared by the PTs were interesting and could attract the attention of students. About half (54%) of the tasks prepared by the PTs were found to include contexts related to the students themselves. This was followed by science (18%) and health (14%) contexts. Social context (11%) was preferred in one of every ten tasks prepared. The least preferred context was found to be education in the task design assignments prepared by PTs (3%).

Another point analyzed was the need for context in order to solve the statistical question. In this connection, Table 5 was created.

Table 5. Using the context to solve the question in the tasks

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Including contextual information (e.g. contextual title, contextual attribute names)</td>
<td>28</td>
<td>100%</td>
</tr>
</tbody>
</table>

When evaluated in terms of the sub-components showing that context was used to solve the statistical question, it was remarkable that all of the task design assignments included contextual information. An example of this is the title “the number of seasonal workers in the provinces of...“ in a graph created for a statistical question focusing on how the number of seasonal workers changes across different provinces. In a table containing the data in another task, the 10-letter figure refers to the sum of the letters in the name and surname of one of the students. When evaluated in terms of multivariate nature of datasets, it was revealed that there are approximately 4 (mean
4.1) attributes per case. The obtained results allow the interpretation that multivariate data sets were used in the tasks. When the designed tasks were evaluated in terms of the source of data, Table 6 was presented.

Table 6. Using the context to solve the question in the tasks

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Including information about the source of the data</td>
<td>24</td>
<td>86%</td>
</tr>
<tr>
<td>Not including information about the source of the data</td>
<td>4</td>
<td>14%</td>
</tr>
</tbody>
</table>

Table 6 was revealed that explanations were made for the data obtained from the primary sources (for example, we wanted to measure the foot length and height of the university students who wanted to be at the university for a week), and in the data sets obtained from the secondary sources, the source was included, except for 4 tasks. On the other hand, in the task constructed from the hypothetical data, it was stated that they created these data themselves, since they could not be accessed from any source.

5.4 SPS

The task design assignments prepared by the PTs were evaluated in terms of SPS and Table 7 was created.

Table 7. Using SPS in the tasks

<table>
<thead>
<tr>
<th>SPS</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Including one or more stages of SPS</td>
<td>3</td>
<td>11%</td>
</tr>
<tr>
<td>Including the whole SPS</td>
<td>25</td>
<td>89%</td>
</tr>
<tr>
<td>SPS levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level A</td>
<td>8</td>
<td>28%</td>
</tr>
<tr>
<td>Level B</td>
<td>16</td>
<td>58%</td>
</tr>
<tr>
<td>Level C</td>
<td>4</td>
<td>14%</td>
</tr>
</tbody>
</table>

While it was observed that the majority of the tasks (89%) prepared by the PTs included the whole SPS, 11% did not include the stage of interpreting the results. The majority of the tasks prepared by the PTs were found to be at Level B (58%), followed by Level A (28%). Only four of the tasks prepared by the PTs were found to be at Level C.

Level A tasks

Although statistical questions related to the tasks prepared at this level included small
groups (e.g., PTs in a classroom, students in a classroom, some cities), it was observed
that these groups were well defined. Some sample questions in these tasks are “What
kind of sports do the students in our class like?, What is the distribution of the colours
of the sweaters worn by the PTs?, Did the Black plague or Covid-19 cause more deaths
in the cities of Moscow, Venice, London, Beijing, Paris and Warsaw? What is the av-
erage lifespan of mammals?”. In addition, it was noted that the contexts used were
chosen in a way that would attract the attention of students.

It was observed that the data sources used to answer these questions in the tasks
were obtained from both primary and secondary sources. For example, the PTs col-
lected data from their classrooms to answer the research question about what kind of
sports the students in our class like, while they used the CODAP sample dataset to
answer the question about the average lifespan of mammals. They noted down the
data they collected/used in excel files or papers. It was observed that the PTs used
both categorical (e.g., sports) and numerical (e.g., weight) variables. In addition, it
was seen that they were aware of how the variable (e.g., the sum of the number of
letters in the names) in the tasks they prepared was distributed, that is, how many
times a certain result occurred.

It was observed that the PTs used different representations such as tables, bar
charts, picture graphs, and dotplots while analyzing the data they collected in order
to answer the research question they prepared, as well as taking into account the
measures of central tendency (e.g., arithmetic mean).

It was seen that the PTs interpreted the data they analyzed by taking into account
the group they dealt with. For example, “In the city of Moscow, the number of deaths
caused by Covid 19 is higher than the number of deaths caused by Black Plague. How-
ever, the number of deaths caused by Black Plague in Beijing city is higher than the
number of deaths caused by Covid 19”. The mean was calculated in a task but it was
observed that no comment was made on what this mean meant. Below are presented
two sample tasks for this level.

In the first task, the PTs created the following research question to be answered
on the basis of their own classroom “A sports tournament is planned to be held at the
end of the year for PTs. For this purpose, it should be determined which sport is liked
by the PTs. What kind of sports do the PTs in our class like?”. To collect data for this
purpose, they created the following survey question “What is your favourite sport?”. The questionnaire prepared to collect data is given below in Figure 3.
Here, it is seen that the data in the data collection tool are categorical data. It can be said that since the survey question asked to choose between four options, it became easier to organize and analyze the data. However, a limitation can be pointed out here. The fact that sports branches are limited to four options will make it difficult for PTs who do not like one of these sports and like another sport (e.g., athletics) to choose. Below in Figure 4 are given the raw data collected by the PTs.

![Figure 3. Data collection tool in the task](image)

**Figure 3. Data collection tool in the task**

Here, it is seen that the data in the data collection tool are categorical data. It can be said that since the survey question asked to choose between four options, it became easier to organize and analyze the data. However, a limitation can be pointed out here. The fact that sports branches are limited to four options will make it difficult for PTs who do not like one of these sports and like another sport (e.g., athletics) to choose. Below in Figure 4 are given the raw data collected by the PTs.

![Figure 4. Raw data collected for the task](image)

**Figure 4. Raw data collected for the task**
Based on these collected raw data, the PTs created horizontal and vertical dotplots to represent the data and they also represented the data with a table.

![Diagram](image)

Figure 5. Types of representations used in the task

The PTs interpreted the data they analysed in the tasks they prepared as follows; “The numbers of the sports branches that the PTs like are very close to each other, the PTs preferred tennis the most, we can say that when one of the PTs is chosen randomly, his/her probability of liking volleyball is less than his/her probability of liking tennis.” These interpretations can be considered as an indicator of what they are doing is reading the data and reading between the data.

Another sample task prepared by the PTs at Level A was about the lifespan of mammals. In the task they prepared, the PTs asked the question “What is the average lifespan of mammals?”. The data set used was a sample data set in CODAP and retrieved from [https://codap.concord.org/app/static/dg/tr/cert/index.html](https://codap.concord.org/app/static/dg/tr/cert/index.html).
The PTs calculated the average lifespan of mammals here. They found the average as 24.85. However, they did not make any interpretations on this result. Here, the PTs were expected to make evaluations about which mammal lifespan is closer to the average and which mammal lifespan is farther from the average because such interpretations are also evaluations of variability across the data set obtained.

Level B tasks

In the tasks prepared at this level, it was observed that statistical questions were formed for comparison (e.g., do the types of music that students like differ between classes?) and association (e.g., is using right or left hand related to gender?) between variables based on a larger sample. It was also seen that they included questions that required examining the change of a variable over time (for example, how the average life expectancy of women in Turkey changed by years). In addition, PTs designed questions that would include two categorical (do students who like bananas tend to like/dislike strawberries?), two numerical (Is the height of students related to their jumping height?) variables and one categorical and one numerical variable (Is using...
a computer program effective on students’ statistical exam grades?). In addition to collecting first-hand data, they created research questions using data sets obtained from online sources and websites (e.g., TUIK). It was noted that the collected/obtained data were recorded in the Excel program. They also made arrangements to perform random assignments to control certain traits. For example, a research question “Does going to the support course affect the number of words that students read in 1 minute?” was asked. In this context, students studying in ....primary school 1/A, 1/B, 1/C and 1/D classes were determined. Although the students in these classes were not randomly selected, the students were randomly assigned while conducting the relevant experiment.

It was revealed that various representations (e.g., bar graph, scatterplots, two-way graph, dot plot, mosaic plot, a time series plot) were used to analyze the collected data, measurements (e.g., mean absolute deviation (MAD), measurements of central tendency) were used to describe the distribution and measurements (e.g., correlation coefficient) were used to elicit the relationship between two variables.

It was seen that while interpreting the analyzed data, they used expressions to look at reading beyond the data as well as reading the data and reading between the data. It was also observed that they made comments to compare the results for different conditions in an experiment (e.g. how using/not using a computer program affects statistical grades). In addition, it was observed that they stated that although the selected samples were larger than the samples at Level A, they still might not represent the population (for example, although we evaluate the average life expectancy of women in Turkey, we cannot say that the average life expectancy of women all over the world is like this). A sample task is presented below.

In the sample task, the PTs focused on the periodic table of the elements. Based on the various properties of the elements in the periodic table, they prepared research questions to reveal the relationship between these properties. For example, they posed the following research question; “Is there a relationship between the melting points of the elements and their boiling points?” To this end, they obtained the data from a secondary data source. The data were obtained from https://codap.concord.org/app/static/dg/tr/cert/index.html. An example of the data they handled is given below in Figure 7.
Figure 7. The data set obtained for the task

They created scatterplots to analyze the data they obtained (See Figure 8).

Based on this graph, it was noted that the relationship between the melting point and boiling point of the elements was close to linear, that is, they made interpretations that the melting point of the relevant element would increase as the boiling point of the related element increased.
Level C tasks

It was noted that the statistical questions asked by the PTs in the activities they prepared at Level C included two or more variables and focused on causality and prediction. For example, the following questions can be given as examples to the questions at this level; “What is the relationship between the time of having a heart attack and its being fatal?, Is there a relationship between the height of individuals and the length of their feet? Can height be estimated from the foot length?, How has the working population in Turkey changed over the years?”. They preferred to use primary or secondary data sets to answer these questions. They were able to determine the appropriate method (e.g., survey research, observational studies and experiments) according to the research question. They became aware of the role of random selection when selecting samples and the fact that the random assignment in experimental assignments influenced cause-and-effect interpretations.

It was seen that they used high-level statistical concepts (e.g., population proportion (p), Pearson’s correlation coefficient (r), Quadrant Count Ration (QCR)) and high-level analysis methods (e.g., Chi-squared tests) when analyzing the data. In addition to making interpretation on what the estimation of a variable’s property means, they also used expressions to understand how the variables affect each other. Here, it can be said that PTs could make advanced interpretations (reading between, beyond and behind the data). Below is given a sample task.

In the sample task, the PTs formulated a statistical question; “Is there a relationship between the height of the individuals and the length of their feet? Can height be estimated from foot length?” In order to answer this question, PTs preferred to collect data themselves and prepared a questionnaire given below in Figure 9.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Female ☐</th>
<th>Male ☐</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height :</td>
<td>(cm)</td>
<td></td>
</tr>
<tr>
<td>Length of the feet :</td>
<td>(cm)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Data collection tool for the task
Here, the PTs were aware that they could not reach all students in every faculty. Thus, they asked various eager students to fill out the questionnaire. They were also able to define and record the obtained variables.

As an example, regarding the problem of estimating the height from the foot length, the PTs measured and recorded the height and foot length of 100 university students randomly selected on the campus.

![Raw data collected for the task](image)

They transferred the collected data to the CODAP program and calculated Quadrant Count Ratio (QCR) by drawing a scatterplot. They also included calculating the Pearson correlation coefficient (r), which takes into account the distance of the data from the mean lines. The results obtained showed a strong positive linear correlation between height and foot length. Based on this correlation, they wrote the equation for
the least-squares line, with the help of technology to estimate the foot length from the height.

Foot length = 0.1639*(height) - 2.5

They evaluated whether the model prepared here was suitable and had a good fit and showed the positive and negative deviations in the data by drawing the fit line.

For example, the foot length of a student whose height is 185 cm

Foot length = 0.1639*(185) - 2.5 = 27.82

By analyzing the data, they argued that there was a linear correlation between foot length and height and that foot length could be used to predict height and they supported this argument with examples.

6 Discussion and conclusion

The current study aimed to examine SPS tasks prepared by the PTs using the CODAP dynamic statistics software tool. One of the motivations of the study was that many other studies (Casey et al., 2020; 2021; Langrall et al., 2017; Shaughnessy, 2007) drew attention to the difficulties experienced by PTs in this regard. In this connection, the PTs held various discussions on how to prepare SPS tasks within the scope of a course,
and made examinations about the preparation of appropriate tasks and the selection of appropriate technological software. Then, the PTs prepared task design assignments. The prepared tasks were analyzed within the context the C-DIST framework.

When the task design assignments are evaluated in terms of learning goals, it can be said that the pre-service teachers tended to prepare statistical questions. From among the learning goals, interpretation based on the measures and types of representation used to represent the data was largely adopted. Making inferences based on the measures and types of representation used to represent the data was a less preferred learning goal by PTs in their task design assignments. Only one task focused on calculating the arithmetic mean, but this measure was carried out by taking the mathematical calculation into the centre. These findings can allow the interpretation that the PTs determined the goals of statistical tasks with a statistical approach. However, in a designed task, it was focused on calculating the arithmetic mean by adopting a mathematical approach. Casey et al. (2021) stated that contrary to the findings of the current study, the PTs mostly structured the tasks they prepared with a mathematical approach. When Chick and Pierce (2008) examined the lesson plans of the PTs, they observed that the majority of them made plans with the aim of simply reading or interpreting the information shown by a table or graph, and less than half of the questions they prepared were aimed at identifying information that was not immediately visible from the data or making inferences from the data.

When the data used in the tasks were examined, it was seen that in half of the task design assignments, the data came from the real source and the secondary data source was used, while in 46% of the tasks, the data came from the real source and were collected by the PTs themselves. The PTs’ use of mostly real but secondary data can be evaluated under several headings. The first of these may be to provide information about various websites in the course so that PTs can access the data sets. The PTs, who gained knowledge about how to access the data sets, may have preferred secondary data sets for the tasks they prepared. It was observed that a hypothetical data set was preferred in only one task. The fact that the PTs could not reach the appropriate data set may have caused them to prefer the hypothetical data set. When the tasks were evaluated in terms of the number of cases, it can be said that the PTs preferred to use large data sets. In addition, it was observed that the data were multivariate and different attributes were taken into consideration. The use of large, multivariate and real data sets for effective statistics teaching is a point that has been emphasized in many studies (Bargagliotti et al., 2020; Casey et al., 2020; 2021; Franklin et al., 2015). In
addition to studies that have reached similar results (Casey et al., 2020; 2021), there are studies that draw attention to different results (Chick & Pierce, 2008; Weiland, 2019). For example, Weiland (2019) revealed that small, univariate and bivariate tasks containing imaginary data are included in high school textbooks. In addition, another finding is that the task design assignments created by PTs were structured in such a way as to need data in order to be completed successfully. Parallel to this result, Chick and Pierce (2008) examined the lesson plans prepared by PTs in their study and stated that more than half of the lesson plans directed their students to the data set.

It was revealed that the PTs mostly tended to use the contexts related to themselves in the designed task design assignments, followed by the contexts in the field of science. It was observed that health and social contexts were preferred by PTs in approximately one out of every ten tasks. In the task design assignments prepared by PTs, the least preferred context was found to be education. The efforts of PTs to include information about the context to solve the questions prepared in the task design assignments were also remarkable. Giving a contextual title, context-based attribute names can be given as examples to this. However, although information about the source of data was given in most of the tasks, no information was given about the source of the data in four tasks. In general, it can be said that the PTs considered importance of creating connections with the context in the tasks they prepared. While this result concurs with the results of Casey et al. (2020; 2021), it differs from the results of Bakogianni (2015). While Casey et al. (2021) stated that one of the strengths of the tasks designed by the PTs is the constant connections made with the context, Bakogianni (2015) revealed that statistical context is not only a difficult learning goal to reach but also an important teaching challenge for mathematics teachers. Although calculating statistical measures and constructing representations are important skills, it is emphasized that students should be given the opportunity to conduct research within contexts in order to gain statistical skills (Bargagliotti et al., 2020; Casey et al, 2020; Franklin et al., 2007).

When the prepared task design assignments created by PTs were examined in terms of the SPS, it was observed that most of the task design assignments included the entire SPS, and only three tasks did not include the stage of interpreting the results. When the designed task design assignments were examined in terms of their SPS levels, it was revealed that the tasks were mostly at Level B, followed by Level A. The number of the tasks prepared at Level C was found to be the lowest. It can be said
that the prepared task design assignments contained statistical questions and provided guiding information about data collection. Preparing survey questions and questionnaires on how to collect data can be given as an example to this. In different studies (Casey et al., 2020), attention was drawn to the paucity of tasks that prompted one to consider how data were collected or how the collection method might affect their interpretation. However, it is emphasized that teachers should support students to understand the data collection process (McClain & Cobb, 2001). In addition to enabling the exploration of different forms of representation, the CODAP program may have supported PTs in the process of analyzing the data by allowing them to perform calculations to visualize relationships between variables (e.g., the equation for the least-squares line). Many studies (e.g., Casey et al., 2020; Prodromou, 2015) have drawn attention to the convenience that different technological software provides in the statistics teaching.

In the process of interpreting the obtained data, interpretations were made on the basis of reading the data and reading between the data in Level A tasks, in addition to these interpretations, reading beyond the data was also performed in Level B tasks and in Level C tasks, reading behind the data was also performed. In general, reading the data, reading between the data and reading beyond the data interpretations were mainly performed in the task design assignments. In studies (Casey et al., 2020; Chick & Pierce, 2008), it is pointed out that the interpretations made by PTs are at a simpler level (reading the data, reading between the data) and that more advanced levels should be included. In this sense, recent results (Casey et al., 2021) and the findings of the current study can indicate that there are improvements in the quality of statistical tasks.

The fact that PTs mainly prepared task design assignments at Level A and B shows that they are limited in preparing task design assignments at Level C. There might be two reasons why PTs are limited in preparing tasks at Level C. Lack of knowledge of PTs may have caused them to have difficulties in preparing task design assignments at Level C. Studies (e.g. Burgess, 2007; Casey & Wasserman, 2015; Hannigan et al., 2013) draw attention to the fact that the lack of knowledge of teachers and PTs affects their task preparation skills. Another reason for the difficulties they experienced in preparing tasks at Level C may be the objectives in the curriculum in Turkey. It was explained to PTs that while preparing task design assignments, they could receive support from the curriculum as well as academic resources. Studies conducted (Batur, et al., 2021) have determined that the curriculum in Turkey mainly includes Level A
objectives, followed by Level B objectives. There are no objectives at Level C. Furthermore, Balcı (2023), who examined textbooks and curriculum in Turkey, concluded that the step of interpreting findings in the curriculum is only limited to making interpretations and decisions based on evidence obtained from data analysis and does not explicitly include making inferences.

The results showed that although the PTs had some difficulties in the preparation of task design assignments (for example, asking mathematical questions, having few tasks at Level C), in general, their efforts to consider SPS task preparation components attracted attention. It can be thought that the PTs’ discussions on how effective statistics teaching should be during the course they took, as well as their examining sample tasks, led to such a result. The fact that the CODAP software allows for the creation of multiple and various graphs and the calculation of various measurements (e.g., central tendency, dispersion) easily may have helped the PTs focus more on other components of the tasks (e.g., context, interpreting the results). Recent studies (Gorman, 2017) have shown that teachers tend to create their own teaching materials rather than using textbooks. When evaluated in this context, it is seen to be important for effective statistics teaching that PTs, who will be the teachers of the future, make progress in preparing SPS tasks.

The current study focused on the PTs’ preparation of task design assignments. What was done in the tasks was based on guessing the thoughts of the students and the tasks were not implemented in the classroom environment. This can be considered as a limitation of the study. In future studies, it can be discussed how these tasks are reflected in the implementation process. Moreover, opportunities can be provided for PTs to work on large, multivariate and real data to design higher-order tasks and to experience a SPS by using software such as CODAP (Casey et al., 2020) so that they can prepare tasks at Level C.

References


Concord Consortium (2019). *Common online data analysis platform (CODAP) [computer software]*. https://codap.concord.org/


Ministry of National Education [MoNE], (2018). Middle school mathematics curriculum (Elementary and middle schools 1,2,3,4,5,6,7 and 8 grades). [Matematik dersi öğretim programı (il-kokul ve ortaokul 1,2,3,4,5,6,7 and 8. Sınıflar]. Ankara.


