Student Checklist (1A)

This form is required for ALL projects.

1. a. Student/Team Leader: Ankit Sayed Grade: 12

   Email: ankitsayed3@gmail.com Phone: 5167273499

   b. Team Member: ______ c. Team Member: ______

2. Title of Project:

   Analyzing long-term improvement in adolescent cross country performance

3. School: W. Tresper Clarke High School School Phone: 5168767451

   School Address: 740 Edgewood Dr, Westbury, NY 11590

4. Adult Sponsor: Erika Rotolo Phone/Email: erotolo@emufsd.us

5. Does this project need SRC/IRB/IACUC or other pre-approval? ☐ Yes ☑ No Tentative start date: ______

6. Is this a continuation/progression from a previous year? ☑ Yes ☐ No
   If Yes:
   a. Attach the previous year's ☑ Abstract **and** ☑ Research Plan/Project Summary
   b. Explain how this project is new and different from previous years on
   ☑☐ Continuation/Research Progression Form (7)

7. This year's laboratory experiment/data collection:

   8/10/19 — Actual Start Date: (mm/dd/yy)

   10/10/19 — End Date: (mm/dd/yy)

8. Where will you conduct your experimentation? (check all that apply)

   ☐ Research Institution ☐ School ☐ Field ☑ Home ☐ Other: ______

9. List name and address of all non-home and non-school work site(s):

   Name: ______ ______

   Address: ______ ______

   ______ ______

   Phone/email ______ ______

10. **Complete a Research Plan/Project Summary following the Research Plan/Project Summary instructions and attach to this form.**

11. **An abstract is required for all projects after experimentation.**

Research Plan/Project Summary Instructions

A complete Research Plan/Project Summary is required for ALL projects and must accompany Student Checklist (1A).

1. All projects must have a Research Plan/Project Summary
   a. Written prior to experimentation following the instructions below to detail the rationale, research question(s), methodology, and risk assessment of the proposed research.
   b. If changes are made during the research, such changes can be added to the original research plan as an addendum, recognizing that some changes may require returning to the IRB or SRC for appropriate review and approvals. If no additional approvals are required, this addendum serves as a project summary to explain research that was conducted.
   c. If no changes are made from the original research plan, no project summary is required.

2. Some studies, such as an engineering design or mathematics projects, will be less detailed in the initial project plan and will change through the course of research. If such changes occur, a project summary that explains what was done is required and can be appended to the original research plan.

3. The Research Plan/Project Summary should include the following:
   a. **RATIONALE:** Include a brief synopsis of the background that supports your research problem and explain why this research is important and if applicable, explain any societal impact of your research.
   b. **RESEARCH QUESTION(S), HYPOTHESIS(ES), ENGINEERING GOAL(S), EXPECTED OUTCOMES:** How is this based on the rationale described above?
   c. Describe the following in detail:
      - **Procedures:** Detail all procedures and experimental design including methods for data collection. Describe only your project. Do not include work done by mentor or others.
      - **Risk and Safety:** Identify any potential risks and safety precautions needed.
      - **Data Analysis:** Describe the procedures you will use to analyze the data/results.
   d. **BIBLIOGRAPHY:** List major references (e.g. science journal articles, books, internet sites) from your literature review. If you plan to use vertebrate animals, one of these references must be an animal care reference.

Items 1–4 below are subject-specific guidelines for additional items to be included in your research plan/project summary as applicable.

1. **Human participants research:**
   a. Participants: Describe age range, gender, racial/ethnic composition of participants. Identify vulnerable populations (minors, pregnant women, prisoners, mentally disabled or economically disadvantaged).
   b. Recruitment: Where will you find your participants? How will they be invited to participate?
   c. Methods: What will participants be asked to do? Will you use any surveys, questionnaires or tests? If yes and not your own, how did you obtain? Did it require permissions? If so, explain. What is the frequency and length of time involved for each subject?
   d. Risk Assessment: What are the risks or potential discomforts (physical, psychological, time involved, social, legal, etc.) to participants? How will you minimize risks? List any benefits to society or participants.
   e. Protection of Privacy: Will identifiable information (e.g., names, telephone numbers, birth dates, email addresses) be collected? Will data be confidential/anonymous? If anonymous, describe how the data will be collected. If not anonymous, what procedures are in place for safeguarding confidentiality? Where will data be stored? Who will have access to the data? What will you do with the data after the study?
   f. Informed Consent Process: Describe how you will inform participants about the purpose of the study, what they will be asked to do, that their participation is voluntary and they have the right to stop at any time.

2. **Vertebrate animal research:**
   a. Discuss potential ALTERNATIVES to vertebrate animal use and present justification for use of vertebrates.
   b. Explain potential impact or contribution of this research.
   c. Detail all procedures to be used, including methods used to minimize potential discomfort, distress, pain and ~~injury~~ to the animals and detailed chemical concentrations and drug dosages.
   d. Detail animal numbers, species, strain, sex, age, source, etc., include justification of the numbers planned.
   e. Describe housing and oversight of daily care
   f. Discuss disposition of the animals at the termination of the study.

3. **Potentially hazardous biological agents research:**
   a. Give source of the organism and describe BSL assessment process and BSL determination.
   b. Detail safety precautions and discuss methods of disposal.

4. **Hazardous chemicals, activities & devices:**
   - Describe Risk Assessment process, supervision, safety precautions and methods of disposal.
   - Material Safety Data Sheets are not necessary to submit with paperwork.

Research Plan

Ankit Sayed

W.T. Clarke High School
740 Edgewood Drive, Westbury, NY, 11590

Analyzing Long-Term Improvement in Adolescent Cross Country Performance

Behavioral and Social Sciences

a. Rationale:

Sports are ripe for experimentation. They offer psychologists controlled conditions, a set of rules for participants to abide by, and set goals for success. Within sports, there are defined performance metrics, such as race times for running or points scored in basketball. These have a high degree of observability compared to other study methods (Balafoutas et. al 2019). Scientists can study certain mental processes within the context of sports, such as decision-making. Decision-making, in sports, is the sum of the actions that an athlete does during an activity, such as pacing changes during a long distance race in a cross country meet.

Predicting future outcomes, and subsequently over and underestimating success is a valuable skill in sports. In running, estimation of success is shown in pacing, where a long-distance runner will run at a certain speed while gauging their energy level. Overconfidence can be seen with the characteristic sharp increase in pace followed by sharp decrease, displaying that the runner overestimated their ability to increase speed and was forced to slow down. Therefore, slowdown during a race can be used to determine overconfidence levels (Krawczyk and Wilamowski 2018). Longer race distances tend to have higher slowdowns, i.e. more overconfidence. If studied by gender, males have a higher tendency to overestimate their performances, at least in marathons, though other distances have been studied to similar effect. Age also plays a part, with young male runners having the highest slowdown out of any age group, especially in the marathon (Figure 1).

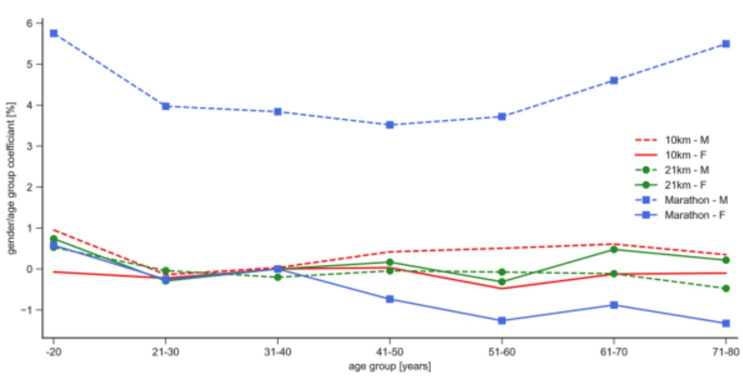

Figure 1:
Race slowdown (RSD) based on age and gender
(Krawczyk and Wilamowski 2018)

Inexperience may be another concern for younger runners slowing down, as they are less likely to know strategy than their older counterparts. Younger runners tend to have high levels of variability in competition, especially when combined with a grueling schedule of three or more competitive races in a month (Hopkins and Hewson 2000). Less time spent running can, of course, spend less time developing the runner's physical abilities, including their oxygen intake and anaerobic power output, both of which are crucial factors in determining running performance (Bulbulian et al. 1986).

Overconfidence may also be due to a heightened emotional state among young runners before they compete. The competitive aspects of running can cause both positive feelings (consonance) and negative feelings (dissonance), both of which reflect as stress. Competitive sports demand certain emotional constructs, such as a fear of the opponent, anger at the performance of others, and a competitive spirit (Lapa et. al 2013). Therefore, the ideal distance runner will have the emotional intelligence and stress tolerance to handle these concerns, although the average teenage runner may have a hard time reaching this ideal in actuality. Many young, competitive runners still have hostile emotional profiles during competition, especially for males (Vasilica et. al 2013).

Statistical analyses are a major part of many sports research projects as well, and a major tool is the ANOVA simulation. ANOVA, or at least the standard one-way ANOVA test, is an analysis of variance that tests the variance between samples against the variance within samples (Chandrakantha 2014). This will output an f-statistic and p-value. The F-statistic shows the variance ratio that is compared to the f-critical values derived from the degrees of freedom, while the p-value shows the statistical significance of the result (Miller 2010). A Monte Carlo simulation is another tool that allows for analysis of the range of possible values when there are unknown factors affecting the data by simulating the data collected many times to create a spread of possible values.

Finally, while statistics may be a powerful tool to determine the mathematical significance of results, sports, which rely heavily on human social and cultural patterns, can also benefit from a qualitative understanding of group and individual psychosocial dynamics. Qualities like team resilience, the capacity for negative stress before it affects performance is almost impossible to quantitatively define, but it can be studied through understanding of the sport at hand and the unique culture and environment of the players, such as a team's taste for humor and experience as a group beyond simple years spent playing together. These factors cannot simply be studied quantitatively, due to the difficulty in creating operational definitions for these concepts, but their effect can be observed through social observations and analysis (Morgan 2019).

b. **Procedures:**

First, running data will be obtained. The data sources are publicly available repositories of high school cross country data for teams in Nassau and Suffolk Counties. The three sources used will be the webpages Just in Time Racing, Section VIII Nassau County Athletics, and NY MileSplit. These sites were recommended by a cross country coach and are updated almost daily with accurate information. Cross country teams on Long Island run 2.4 km, 4 km, or 5 km races on two courses: Bethpage State Park and Sunken Meadow State Park. For each park, the top 250 male runners in 2016 will be selected and cross-referenced for accuracy, with the data already de-identified as well. Only males will be selected to keep the focus of the project narrowed to improvement within a population (male runners), not the general group of runners. This group of runners had their personal record times for each year recorded on each course as well, until each runner has four times, each corresponding to their four-season bests.

To test the hypotheses, the average improvement of a runner will have to be determined. For each runner, the times of each subsequent year will be subtracted so the change per year could be determined. The difference between their improvement on Bethpage and Sunken Meadow respectively was calculated.

Each grade of high school boys (freshmen, sophomores, juniors, and seniors) had their improvement charted. Each runner was graphed, so the graph would show all 250 lines of improvement over high school, which could be averaged out to find the mean improvement at each point.

c. **Risk and Safety:**

There is no significant risk in this project.

d. **Data Analysis:**

Statistics simulations will be conducted using the data. First, ANOVA simulations will be constructed in Excel and conducted for both the Bethpage group and the Sunken Meadow group. The three populations for the test are the three improvements in high school cross country: freshman to sophomore year, sophomore to junior year, and junior to senior year. A standard weighted-means analysis will be conducted since all three variables were linked, as each population was the same runner. The ANOVA simulations will output an f-statistic giving the variance of the data and a p-value about the significance of the variance. Second, a Monte Carlo simulation will be constructed, also using Excel. The given scenario, 250 runners, will be replaced with random values within the bounds of the original data set. This will be randomized multiple times in a what-if data table which outputted values like the mean, average, and number of times a value is produced outside two standard deviations from the mean.

e. **Bibliography:**

Abby Miller (2003) The Development and Controlled Evaluation of Athletic Mental Preparation Strategies in High School Distance Runners, Journal of Applied Sport Psychology, 15:4, 321-334, DOI: 10.1080/714044200

Balafoutas, L., Chowdhury, S.M., Plessner, H., Applications of Sports Data to Study Decision Making, Journal of Economic Psychology (2019), doi: https://doi.org/10.1016/j.joep.2019.02.009

Bulbulian, R., Wilcox, A. R., & Darabos, B. L. (1986). Anaerobic contribution to distance running performance of trained cross-country athletes. Medicine & Science in Sports & Exercise, 18(1). doi: 10.1249/00005768-198602000-00018

Chandrakantha, L. (2014). Learning ANOVA concepts using simulation. Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education. doi: 10.1109/aseezone1.2014.6820644

Hopkins, W. G., & Hewson, D. J. (2001). Variability of competitive performance of distance runners. Medicine & Science in Sports & Exercise, 33(9), 1588–1592. doi: 10.1097/00005768-200109000-00023

Krawczyk, M., Wilamowski, M., Task difficulty and overconfidence. Evidence from distance running, Journal of Economic Psychology (2018), doi: https://doi.org/10.1016/j.joep.2018.12.002

Lapa, T. Y., Aksoy, D., Certel, Z., Özçelik, E. Ç. M. A., & Çelik, G. (2013). Evaluation of Trait Anger and Anger Expression in Taekwondo Athletes in Relation to Gender and Success. Procedia - Social and Behavioral Sciences, 93, 1976–1979. doi: 10.1016/j.sbspro.2013.10.151

Morgan, P. B., Fletcher, D., & Sarkar, M. (2019). Developing team resilience: A season-long study of psychosocial enablers and strategies in a high-level sports team. Psychology of Sport and Exercise, 45, 101543. doi: 10.1016/j.psychsport.2019.101543

Nancy Kay Butts (1982) Physiological Profiles of High School Female Cross Country Runners, Research Quarterly for Exercise and Sport, 53:1, 8-14, DOI: 10.1080/02701367.1982.10605219

Vasilica, G., Cristina, A. I., Iulian, A. D., Georgeta, M., & Radu, P. (2013). The Profile of the Emotional-affective State of the Middle and Long Distance Running Athletes during the Competition. Procedia - Social and Behavioral Sciences, 84, 1269–1274. doi: 10.1016/j.sbspro.2013.06.742