1. **Research Questions**: Identify the research questions. Sometimes, authors do not list them explicitly, but you can infer their objectives, questions, or gaps from the Introduction section.
    - VR training environment for lower extremity balance rehabilitation with real-time guidance and feedback.
2. **Hypotheses**: Determine what hypotheses the authors are testing.
    - H1: For coaching styles, the participant's balance performance will be improved with the positive-reinforcement style of real-time verbal feedback from the coach compared to autonomous-supportive style
    - H2: The participant's subjective experience will be better with the positive-reinforcement style of real-time verbal feedback from the coach compared to autonomous-supportive style.
    - H3: Regarding viewpoint, the participant's balance performance and experience will be improved with the third-person compared to the first-person viewpoint
    - H4: The participant's subjective experience will be better with the third-person compared to the first-person viewpoint
3. **Sample Size**: Understand how they determined the sample size to draw conclusions for their study.
    - Statistical G*Power analysis
    - Analysis was preformed for 2 groups and 4 measurements
    - A total sample size of 16
4. **Study Design**: Examine how they structured their user study. What are the independent and dependent variables?
    - Independent:
        - coaching style - this variable is the real time audio feedback from the virtual coach
            - Positive reinforcement: audio of suggestions to fix posture and balance as well as praise
            - autonomous supportive style: No audio feedback, only silent guidance
        - Users view point of the trainee avatar: altering the users perspective in the VR environment
            - First person: user was seeing through the avatar eyes
            - Third person: the user can see the avatars whole body
    - Dependent:
        - User performance
            - Steadiness: quantified vertical hand height and height variation

        - Foot height: measured the height of the user's non standing foot off the ground for the pose(higher foot = better)
        - Number of mistakes: number of times the non standing foot touches the ground (losing balance)
    - User experience
        - Mid questionnaire measure(1 out of 5 scale questions)
            - Social presence: the user's sense of presence with the virtual coach using questions
            - Embodiment: users' sense of avatar embodiment using questions
            - Enjoyment: measured how much the user enjoyed the condition
            - Difficulty: measured the perceived difficulty of the condition
            - Perception of coach: users' perception of the virtual coach
            - Recommendations: willingness to recommend the system to another
        - Post questionnaire (1 out of 10 scale questions)
            - System usability
            - Task load
                - Mental demand
                - Physical demand
                - Temporal demand
                - Performance
                - Effort
                - frustration
5. **Statistical Methods**: Identify the statistical methods used for analysis.
    - Two way analysis of variance (ANOVA) for the dependent variable
    - Pairwise t-test → when there was a statistically significant difference found in the ANOVA (dependent variables)
    - Pairwise t-test for the mid questionnaire
    - descriptive statistics: mean, standard deviation, etc…
    - Statistical G*Power analysis: used to determin the sample size
    - System usability scale (SUS) scoring: questionnaire
    - NASA TLX Scoring: unweighted (raw) NASA TLX questionnaire was used
6. **Descriptive Analysis**: Did they include descriptive statistics? How were they presented?
    - Participants background and characteristics
        - Ages: between 19-40, mean 24.4 $\pm$ 5.7 years

    - Physical measurements: mean height (65.6 ± 3.7 ft) and weight (148.0 ± 28.7 lbs)
    - Demographics: gender (7 males(44%) 9 females (56%)), education level, headedness, and VR experience shown using counts and percentages
    - Ethnicity and injury history: participants ethnicity and history of lower-body injury, frequent exercise (activities with balance)
  - Users performance: the descriptive results for steadiness(m), foot height(m), and number of mistakes
    - Autonomous support style mean value (std) [se]
    - 1st person pov mean value (std) [se]
    - 3rd person pov mean value (std) [se]
    - Positive reinforcement style mean value (std) [se]
    - 1st person pov under positive reinforcement mean value (std) [se]
    - 3rd person pov under positive reinforcement mean value (std) [se]
  - Questionnaire results: descriptive statistics used to report both mid and post questionnaires mean (M) and standard deviation (SD)
7. **Group Comparisons**: For statistical analysis, how many groups were compared? If more than two groups were involved, which statistical method was appropriate?
  - 2 X 2 within subject design (all 16 participants experienced every condition)
    - Coaching style (2 levels: positive reinforcement vs. autonomous supportive style)
    - Viewpoint (2 levels: 1st person pov vs. 3rd person pov)
  - 1. First-Person Viewpoint with Autonomous-Supportive Style
  - 2. First-Person Viewpoint with Positive-Reinforcement Style
  - 3. Third-Person Viewpoint with Autonomous-Supportive Style
  - 4. Third-Person Viewpoint with Positive-Reinforcement Style
  - Statistical method appropriate: two-way analysis of variance (ANOVA) for dependent variables
    - Determine the main effects of the two independent variables (Coaching Style and Viewpoint).
    - Assess the interaction effect between Coaching Style and Viewpoint
8. **Results Reporting**: How did they report their findings in the Results section?
  - 3 main categories: user performance (objective measures), questionnaire results (subjective measures) , and qualitative participant feedback
9. **Visualizations**: Did they use visualizations to support their interpretations? Do you find the plots more helpful than the descriptive results in tables?

- Yes, they did use visualizations to support their interpretations
- I find them more helpful because the numbers can kind of blur together but with the visualizations, you can immediately see the significance and similarities

10. **Qualitative Analysis**: In addition to quantitative analysis, did they perform any qualitative analysis?
    - Yes they performed qualitative analysis at the end with optional feedback from the participants
11. **Hypothesis Discussion**: Did they discuss how their results relate to their hypotheses? How many hypotheses were rejected, and how many failed to be rejected?
    - Yes they discuss how their results relate to their hypotheses
    - H1: partially failed to be rejected
    - H2: failed to be rejected
    - H3: rejected
    - H4: rejected
12. **Limitations and Future Work**: Did they acknowledge any limitations in their study? Did they propose future work to address those limitations?
    - Yes, they acknowledge their limitations and proposed future work to address those limitations
        - Sensing and data accuracy
        - Sample size and generalizability
        - Methodological biases