Indexing Esport Performance

Benjamin T. Sharpe,1,2 Nicolas Besombres,3 Matthew R. Welsh,2 and Phil D.J. Birch2

1Institute of Psychology, Business, and Human Sciences, University of Chichester, Chichester, United Kingdom; 2Institute of Sport, Nursing and Allied Health, University of Chichester, Chichester, United Kingdom; 3Institut des Sciences du Sport-Santé de Paris, Université Paris Cité, Paris, France

The assessment of an athlete’s performance can play a key role in determining their current state, their readiness to compete, the impact of an experimental manipulation, and/or the influence of an intervention. At present, there is limited empirical evidence stating the indicators that encapsulate individual performance within any esport. To identify the variables that are historically associated with indicating Counter-Strike: Global Offensive performance, a literature review was conducted. Identified variables were accumulated and presented to three technical expert panels composed of world-class esport athletes, researchers, and practitioners. We utilized a modified Delphi method to provide direction concerning the examination of performance in esports.

The expert panelists presented numerous opinions on what encapsulates performance, considerations for best practices, and concerns associated with the semantics of performance. This study presents the opinions of various domain-specific experts and encourages the use of more explicit terminology when discussing performance measurement. It was the intention of the project to generate an open discussion rather than draw a unified conclusion on best practices.

Keywords: counter-strike: global offensive, CS:GO, Delphi method, semantics

From a hidden gem for late 90s video game enthusiasts (Jonasson & Thiborg, 2010) to arguably one of the fastest growing sectors in sport in recent decades (Nagel & Sugishita, 2016), esports has seen a dramatic acceleration in growth over the last 10 years (Himmelstein et al., 2017) and has been estimated to be worth US$24.9 billion (Ahn et al., 2020). Esports can be defined as the competitive activity of playing specific video games that provide professional and/or personal development to the player (for full definition see Pedraza-Ramirez et al., 2020). Esports is quickly growing into a global phenomenon and a major area for empirical study (Hallmann & Giel, 2018; Lee & Schoenstedt, 2011). It is surprising that many perceive the concept of esports being an officially recognized sport as subversive (Jonasson & Thiborg, 2010), with differences only appearing to be in the competitive environment (real world vs. cyberspace; Chae & Kang, 2011; Pizzo et al., 2018). The expansion of research has provided the field with early insight into the impact of esports performance on individual psychophysiology (Leis & Lautenbach, 2020), the influence of cognition on performance (Pedraza-Ramirez et al., 2020), and the testing of our current theoretical understandings (Pluss et al., 2019) as well as providing direction for practitioners to generate future interventions (García-Lanzo et al., 2020; Poulos et al., 2020; Smith et al., 2022; Trotter et al., 2020). Nonetheless, a fundamental question that has been overlooked is what are the indicators of individual performance?

Such a question is not unique to the field of esports. In fact, researchers questioning performance markers have challenged the way we select traditional sporting athletes (Bhattacharjee & Saikia, 2014), manage sectors of industry (Bittici et al., 2012), evaluate health care quality (Werner & Asch, 2007), and assess individual mental states (Lane, 2012). Within the field of esports, authors often discuss the integration of perceptual–motor, perceptual–cognitive, and experiential skills to perform optimally (Pluss et al., 2020), and a substantial proportion of literature has used a plethora of indicators to encapsulate what it is to be an athlete (e.g., Behnke et al., 2022; Moen et al., 2021; Momí et al., 2018; Roebuck et al., 2018). For example, authors have assessed individuals on metrics from standalone accuracy (Neri et al., 2021), to game score (Liu et al., 2021), to perceived performance (Hopp & Fisher, 2017). Although noted by prior authors (Leis et al., 2021; Pedraza-Ramirez et al., 2020; Pluss et al., 2022), the most empirically justified means to indicate performance is limited.

An approach that continues to capture the attention of researchers interested in performance is expertise. Expertise often refers to maximal adaptations to task constraints (Ericsson & Lehmann, 1996; Gruber et al., 2010), often acquired through task experience (Williams et al., 2011). To categorize participants based on their level of expertise, researchers frequently categorize participants into high versus low or expert versus novice expertise groups using a variety of indicators. Unfortunately, these chosen indicators may be arbitrary given that no formal investigation has demonstrated their association with domain-specific esports performance. As such, the current paper encourages the field to take a step backward to question what performance means in esports. Indeed, this is a question that has been proposed during the development of the most popular paradigms in expertise literature (e.g., Ericsson & Tonne, 2013; Ericsson & Ward, 2007; Ericsson & Williams, 2007). Bedard and Chi (1992) highlighted the importance of this distinction, suggesting that those categorized as experts may not demonstrate optimal performance when tested. With performance defined as the action or process of performing a task or function (Oxford Dictionaries, 2021), we wish to generate a conversation among the esports community about the processes which capture true performance within esports. If consensus can be drawn, esports researchers may be better equipped to generate groups or even develop a taxonomy for classifying expertise as seen in traditional sports (e.g., Swann et al., 2015).
One issue facing esports researchers is the vast amount of esports titles available. Given the breadth of esports types (Kowal et al., 2018), the current study will exclusively focus on the genre of first-person shooters (FPS). Counter-Strike: Global Offensive (CS:GO) is currently the top FPS video game consumed by players in the Western world. With approximately 500,000 players at any given time, with a peak of over 900,000 (https://steamcharts.com/), the esport has grown exponentially since its conception. The precursor to the game was unofficially released through a modification of the similarly themed video game Half-Life in 1999 and later went on to take its own platform of Counter-Strike in 2000. The popularity of the game grew into a major franchise with multiple iterations (e.g., Counter-Strike: Condition Zero, Counter-Strike: Source, to name a few), with the first major tournament held in 2001 (Ferguson, 2018). With over 25 million units sold, the game developers Valve Corporation and Hidden Path Entertainment have helped Counter-Strike grow into the biggest game series in esports (Llewellyn, 2018). Prize pools for major competitions (third-party tournaments supported by Valve Co.) often exceed US $1 million (Rizani & Iida, 2018).

To familiarize the reader with CS:GO, we will provide an abbreviated background. Two teams of five players compete in multiple rounds (maximum 1 min and 55 s per round) with the goal of eliminating the opposition or completing the objective. An objective is dependent on the team position (terrorists or counter-terrorists). Terrorists succeed by successfully planting and exploding a bomb, which takes 40 s, in one of the two designated areas, whereas the counter-terrorists must prevent the bomb from being planted or exploding (a bomb can be defused). Teams play the first half of the game (15 rounds) as either terrorist or counter-terrorist before switching sides and playing a further half (15 rounds). Ultimately, the outcome of a single game is best-of-30 rounds. Two aspects of gameplay that may alter the advantages held by either team include kill-death-assist count and round success, which contribute to in-game economy. Economy refers to the monetary balances held by each player, which are used to purchase items (e.g., weapons, armor, bomb-defusal kits) at the beginning of every round. As such, every beneficial act (e.g., eliminating an opponent) increases player balance, and successfully completing team objectives (e.g., bomb plants, round wins) provides a monetary reward to all members of a team. However, to provide an advantage to the weaker team, a monetary reward is provided to the team that consecutively loses multiple rounds. Economy plays a critical role within the game as players independently decide how to spend in-game money which may benefit the team strategy (e.g., defusal kits) or personal objectives (e.g., better weapons to promote more kills). Once a player has been eliminated, their purchases disappear, unless collected by the opposition or a teammate before the round ends, hence providing a high ratio of risk and reward.

The appeal of CS:GO to players stems from the simplicity and predictability of the core game mechanics, unlike many of the other esports titles presenting on the world stage. The appeal for researchers, particularly those focusing on player analytics, however, stems from the video game’s capacity to log in-game events into a data file (e.g., percentage successful shots-on-target). The challenge for researchers interested in individual performance stems from the fundamental nature of the sport being a team-based multiplayer game that is not weighted on the performance of an individual. As such, the following study conducted a literature review to identify peer-reviewed research that adopted a specific indicator(s) to encapsulate individual performance within CS:GO or discussed the topic specifically. Given the limited evidence available, the relative age of the field, and the anticipated challenges associated with examining a single genre title, we assembled a technical expert panel (TEP) to provide direction regarding ways to encapsulate CS:GO performance. To achieve expert consensus on which variable(s) of interest may be best associated with CS:GO performance, a modified Delphi method was adopted to utilize the panel’s domain-specific experiences within esports. To minimize bias among the panelists, experts first individually ranked the identified variables from the literature review based upon their perceived ability to record individual player performance (Phase 1). To generate an open conversation among panelists, a series of focus groups were held to provide an opportunity for individuals to express or debate expert opinions (Phase 2). Therefore, the aim of the present study was to harness an expert panel to glean expert opinions about what encapsulates CS:GO performance and provide specific direction for future research. We hope the study provides preliminary evidence to help guide understanding about how to best capture individual performance within an esport that is, by nature, played as part of a team.

**Method**

**Literature Review**

We targeted research output using PubMed, Clarivate Web of Science, and institutional databases to identify English language or English translated journal articles that explicitly recorded CS:GO performance indicators, or comparable FPS esports, and/or empirical articles that discussed performance indicators within those electronic domains. Keywords were listed prior to the literature search (see Appendix 1); these were used in varying groupings (e.g., Counter-Strike and performance), in addition to their associated abbreviations (e.g., CS:GO and FPS), to identify a broad range of associated literature. As third-party markets (i.e., Steam) utilize tags to categorize video games, we also used these search terms to broaden our empirical search (e.g., fast-paced, tactical). Given the relative age of the field of esports, we conducted a targeted review of any research using the listed databases up to the date we concluded our search (November 6, 2021). Our search identified 1,908 articles that incorporated our preestablished key terms. Providing fulfillment of the validation process (peer review), article titles were retained if they appeared relevant to the broad domain of esports performance. Retained article abstracts were then screened for relevance to our study’s objective (i.e., determine a means to appropriately indicate individualized esports performance). The final process required the full text review of 226 articles with 79 being rejected. Articles were only retained if CS:GO, or comparable FPS, performances were defined, indicated, and/or discussed. This step resulted in a total of 127 articles (see Figure 1).

The selected articles included performance variables that encapsulated a wide range of attributes in cognitive psychology (e.g., reaction time, percentage accuracy, expertise classification, response time). As such, it was possible to group these performance variables into broad themes (see Appendix 2), including: game metrics, categorized by the variable of choice being a provided metric included within the esport (e.g., kill/death ratio); skill, referring to variables that recorded the proficiency of the individual in character control (e.g., mouse control); cognition, referring to any variable that was commonly adopted within the field of cognitive or neurocognitive science (e.g., reaction time); strategy, including markers that recorded an individual’s ability to achieve a game-specific goal or action (e.g., strategy success/failure);
awareness, any variable that demonstrated an understanding of the game’s differing mechanics (e.g., weapon awareness) or environment (e.g., map awareness); knowledge, categorized by the variable testing an individual’s game-specific knowledge (e.g., game economy); experience, referring to the time invested within specific aspects of FPS game(s) (e.g., video game experience); and vigilance, categorized by any indicator that recorded an individual’s ability to remain attentive over a specified period (e.g., concentration). We would like to highlight that some variables may have overlapped with individualized and team-based video game performance (e.g., team experience). A critical friend (an esports researcher who had no further involvement in the project) was used to assess variable selection, which ultimately led to inclusion of all variables. Critical friends are commonly adopted to provide insight where evidence is lacking while providing the authors with critical feedback on methodological designs (Reinard, 2006; Smith & McGannon, 2018). Furthermore, the titles given to each theme may not encapsulate the formal definition of the term (e.g., cognition). Theme titles were provided to facilitate variable grouping and minimize confusion during the panel discussions.

Technical Expert Panel

The study adopted a modified Delphi methodology (Brown, 1968) to establish consensus over the primary research question in two phases (individual rating of variables and online focus groups). This method allowed the TEP to review, interpret, and consider available evidence (or the lack thereof) before participating in a moderated discussion of results within a focus group (Hasson & Keeney, 2011; Thangaratinam & Redman, 2005). To establish the TEP, we invited 67 individuals from a range of esports backgrounds (e.g., researchers, practitioners) and 40 retired or current professional CS:GO athletes while attempting to maintain balance wherever possible (e.g., gender, geographical location). The final count included 28 individuals with a range of esports roles and experiences. Of the esports researchers that joined the panel, individuals held expertise in performance and the influence of human factors, mental ill health, stress and coping, and player evaluation among many more. Esports practitioners held expertise as performance coaches, team coaches, clinical psychologists, and analysts. Esports professional athletes, whether retired or current, had all relied on the role for a primary source of income and had competed internationally. All panelists were English-speaking individuals based in a range of continents.

The two phases required the panelists to consider the value of each variable highlighted earlier. Specifically, what variable(s) may best encapsulate individualized performance in CG:GO (e.g., reaction time). The panelists were first introduced to the variables of interest using an online Qualtrics questionnaire (Phase 1) wherein detail was intentionally minimized to avoid potential biases prior to the second stage. Here, the experts were individually presented with each variable and were asked to grade the relevance to capturing performance (1 = highly important, 2 = less important, 3 = unsure, and 4 = not important). Panelists were prompted throughout to consider how, why, and should each indicator be adopted within empirical research. A consensus was reached if at least 70% of the entire sample voted “highly important” on a variable. If individuals selected “unsure” they were removed from the percentage calculation (Hasson et al., 2000). The intention of this initial phase was to build an overview of the thoughts and opinions of the panelists to determine whether any specific groups (e.g., coach vs. athlete) held any substantial differences in opinion and provided the experimenters with a direction of emphasis during the focus groups.

The second stage of the study involved several groups being formed to match availability and accommodate individual requests (Phase 2). As such, three groups were formed. Group A (n = 13) comprised exclusively of past and current professional CS:GO athletes (M = 6.01, SD = 1.92 years game-specific professional experience) including three women and 10 men. Group A, termed the athletes panel, lived in a range of countries including Denmark (n = 2), Norway (n = 1), Ukraine (n = 1), United States (n = 5), and United Kingdom (n = 4) with ages ranging from 18 to 28 (M = 23.23, SD = 3.31 years). Group B (n = 7) comprised three female and four male esports researchers (n = 5) and practitioners (n = 2) (M = 8.94, SD = 3.51 years of experience in such fields). Finally, Group C (n = 8) comprised eight male esports researchers (n = 2) and practitioners (n = 6) (M = 5.16, SD = 2.89 years of experience in such fields).

The newly formed groups each participated in a face-to-face webinar and focus group that contained the aims of the project, a review of broad themes identified in the literature review, and the subsequent unrestricted discussion of each variable of interest. The groups were asked to repeat the rating process, as completed prior, or input their expert opinion (e.g., disagreement with a particular variable). The goal of the focus groups was to develop a consensus, or further questions, about each variable of interest. Focus groups were encouraged to consider whether specific variables could be included, considered further, or removed from future discussions. The experimenter allowed the panelists freedom to discuss the
topics at their own rate and only maintained the broad direction of themes. Once each variable was discussed, the experimenter revealed the data gathered from the first stage of the study (e.g., differences in opinions between practitioners and active athletes). This was included to further promote discussion. However, this final stage did not alter the scores of the second rating process. All focus groups lasted between 40 min and 2 hr. The intention of this final step was to generate an open conversation across the panelists in a medium that allowed individuals to draw consensus or debate topics of interest.

The results from the literature review and modified Dephi method are presented next. Irrespective of the study’s findings, the authors would like to emphasize that the intended outcome of the current study was not to draw definitive conclusions on how researchers should indicate performance but, instead, present the start of an open discussion among the esports community.

Results

Literature Review

A total of 127 peer-reviewed articles contributed to our analysis. Among these articles, 32 variables were discussed or tested to indicate individualized performance. These variables of interest fell into a range of broad themes that are now discussed. We note that some articles included numerous variables, which represented different broad themes. As such, the total number of articles per theme does not equate to the total number of articles used in the final analysis. Initially, a review of 63 articles presented seven variables of interest that appeared to represent the umbrella term of game metrics. Each variable was a common aspect of FPS video games that was typically provided to the player. It appeared that most of these variables were adopted in a broad view of player ranking (i.e., domain-specific expertise) or performance proficiency. The majority (49) of these studies required participants to simply input data from their prior match or record an average over a series of tests. Particularly, a proportion of these studies (13) used kill/death ratio (the difference between kills and deaths) to determine expert and novice groupings, whereas others (17) categorized high and low performance levels through percentage accuracy (the percentage of successful on-target shots). When determining within-subject (e.g., stress) or repeated-measure effects (e.g., an intervention, supplementation), several articles (27) adopted the metric of game score (the overall score of the game after a match) or expertise classification (the provided in-game rank) when determining differences to performance.

Three variables were extracted from 23 articles that were categorized under the umbrella term of skill. Some of these articles (16 out of 23) claimed to utilize such variables to determine their participant’s proficiency to perform within the game, whereas others adopted these variables as general performance indicators. Most commonly (15), mouse control (mouse proficiency) was recorded as a performance indicator to determine the influence of various variables of interest (e.g., distractibility, dual tasking, sustained attention). A substantial proportion of articles (48) utilized commonly adopted cognitive metrics to encapsulate FPS performance. Collectively, four variables were categorized under the theme of cognition. The most prominent variable (46) of interest included reaction time (speed to respond to a stimulus). Our literature review appeared to suggest that reaction time was the most popular variable adopted to determine the influence of a group or investigate the unique contribution of a specific variable of interest. The umbrella terms strategy, awareness, and knowledge included three, two, and four variables of interest, respectively. Of these broad themes, variables were extracted from 19 articles; however, only a small proportion of these articles (5) tested such metrics. Specifically, these experimental studies generated strategy-based rulesets (e.g., successfully planting the bomb) to determine a player’s ability to perform under varying scenarios. Of the remaining articles (13), we descriptively suggest the potential benefit of their inclusion (e.g., testing map knowledge) when determining player performance and expertise.

Many of the included articles refer to or utilized experience (e.g., hours spent playing a particular esport) as a performance indicator. However, as mentioned prior, this means of recording performance was questioned during the literature review process (e.g., experience vs. expertise). Nonetheless, across 71 articles, six variables of interest were included that were categorized under experience. The most adopted variables included game-specific experience (56) and competitive experience (32), referring to the total time spent playing a specific game and the total time spent competing, respectively. Finally, three variables were extracted from nine articles that were categorized under the term vigilance. Each variable of interest claimed to determine player performance yet incorporated other prominent aspects of the game (e.g., influence of time, impact of stress). The authors wish to note that these articles appeared to range in their semantic adoption of the term performance, and hence, this latter category was approached with caution during the second stage.

Phase 1: Preliminary Grading

Mentioning only those that were considered highly important, six out of the possible 32 were highlighted by 50% or more of the expert panelists. Across the entire sample, over 70% of the sample reported reaction time and response time as the most valuable variables of interest. Likewise, 60% of the sample also considered game fatigue (i.e., ability to maintain and regulate fatigue throughout a game), concentration (i.e., ability to maintain and regulate concentration throughout a game), and stress (i.e., ability to maintain and regulate response to stress throughout a game) as variables that should be considered further. Finally, 63% considered processing speed (i.e., the time it takes a person to do a mental task) as a metric that may enable researchers to indicate individualized performance. Of the 32 variables, the panelists considered more than half (19) to be less important on average. However, four of the 32 variables of interest were considered by more than 50% of the sample as not important for determining individual performance. Specifically, these variables included expertise classification (i.e., the provided in-game rank), total kills (i.e., number of kills during a game/task), total deaths (i.e., number of deaths during a game/task), and shooting distance (i.e., the proficiency to hit a target at distance). The remaining three variables of interest were scored as unsure (i.e., coach-led experience, practice time, and competitive team experience).

From this initial stage of our study, numerous differences in opinions were present between disciplines. Specifically, the current and retired professionals (n = 13) determined mouse control, keyboard proficiency, reaction time, and response time to be highly important. Similarly, those who primarily considered themselves esports practitioners (n = 4) considered percentage accuracy, mouse control, response time, and kill/death ratio as the most valuable variables of interest. However, the esports researchers (n = 11) incorporated numerous alternative variables of interest to encapsulate individual performance, including reaction time,
response time, percentage accuracy, fatigue, concentration, and stress. It may be suggested that the variables highlighted by each discipline were indicative of their role and environment (e.g., a coach spending more time with players).

Phase 2: Focus Groups

Group A

Coach-Determined Metrics and Player Role. The panelists highlighted at the onset of the focus group that their performance was continuously monitored throughout a season through scrimmages (referring to matched games with competitive teams outside of competition). Specifically, panelists noted that their coaches provided a breakdown of performance metrics per scrimmage based on accuracy, kill/death ratio, economy efficiency (i.e., in-game cash flow and usage), bombs planted, and overall rounds won. Likewise, the panel noted that this occurred after competitive matches during a competitive season. It was discussed that team coaches would ask members to rate (0–100) the performance of each other based on their ability to complete their objective (i.e., role within the team). Of interest, the specific role (e.g., Lurker) did not arise during the literature review as a metric that would influence how an athlete’s performance may be indicated. Panelists noted that although they held an opinion on the variables that may best encapsulate individualized performance, the group did not feel that this battery of variables would be appropriate for all roles, particularly roles that were not primarily focused on typical objectives (e.g., planting a bomb) and, instead, focused on directing the team (e.g., in-game leaders). Irrespective, the group concluded on three variables of priority: mouse control, reaction time, and response time. Panelist AD noted, “The ability to demonstrate these three basic variables can be a quick way to distinguish the good players from the bad,” and member AA added, “When you get to a professional level though, maybe top 100, then other things come into play like game sense or competitive experience.” Panel member AA added, “I guarantee that a fragger at any level beyond nova would have a better reaction time to a professional with a different role,” and AC added “I’d confidently say an AWPer at any level would perform better at a response time task than most professionals of different roles.” Here, both panel members highlighted the potential challenge associated with not considering CS:GO role.

Game Sense. When prompted to define game sense, the group collectively referred to it as “knowing what to do in a certain situation” whether that was from “prior experience or from just understanding the timings of the game perfectly.” When asked how game sense might be measured, the panel had no suggestions; however, panelist AK suggested, “I’d be able to tell you after it happened and why I did what I did.” Finally, the experimenter revealed the data from phase one of the study and outlined the differences in opinions based on discipline. Of interest, the panel agreed with the esports researcher decision to include stress, but not fatigue and concentration, as a primary variable. Here, the group noted how their ability to deal with “the stress of the tournament” (e.g., loud crowds, money on the line, and not letting the team down) “separated those from the best of the best to just good competitors.” When prompted, the group suggested that they rarely felt fatigued from playing, nor did they feel that their concentration was hampered during competitive play. In fact, panel member AC stated, “I think the reason we are all professionals is because we can concentrate on these video games without becoming fatigued like a typical player.” The topic of tournaments also generated the statement, “Surely the score of the game is the key variable here.” When prompted, panelist AK elaborated, “Irrespective of the level of players the only thing that’s discussed at the end of the game is game score, so why wouldn’t that be the primary measurement tool adopted by researchers.” The entire panel agreed with this statement, suggesting that although they considered numerous variables of value, the score of the game(s) must always be considered.

Summary. Group A reached 70% consensus on four variables (mouse control, reaction time, response time, and game score) holding the potential to encapsulate CS:GO performance.

Group B

Semantics. The initial portion of the focus group was centered on semantics of the term “performance.” Panelist BC stated, “Performance is a fraught subject,” pointing toward the apparent difficulty in categorizing player or team performance in addition to the various philosophical approaches that could be used. The panel first drew consensus on their own definition of performance before discussing the variables of focus, defining the term as “the ability to accomplish task.” Panelist BB suggested, “Irrespective of the tasks chosen to categorise performance, it may be of value to select numerous variables that can be used to provide a composite score,” pointing toward the latent variable approach commonly adopted in cognitive psychology (e.g., Corbett et al., 2015; Metzler-Baddeley et al., 2017).

Complexity of Indicating Performance. Each theme was presented to the panel with discussion centered on how each variable of interest may be tested and whether they truly contributed to successful player performance. Three variables gathered substantial consensus throughout this stage that were considered most valuable for researchers to record (processing speed, reaction time, and mouse control). BB suggested, “A player that can respond to a game specific mental challenge efficiently and most rapidly, whilst having the motor ability to control their character to achieve the new goal, surely has the greatest advantage.” The final theme to be presented, vigilance, also stimulated considerable debate. Consensus was reached that such variables (i.e., fatigue, concentration, and stress), as defined for the present study, may be best adopted as moderator variables (e.g., variable that may influence the level, direction, or presence of a relationship) instead of a primary mechanism to encapsulate performance. Panelists BB and BE suggested, “Reaction time may distinguish player’s basic skill level, yet their capacity to handle considerable pressure may enable researchers to separate those most ready to handle the pressures of elite competitive esports.” However, two further variables gathered substantial consensus throughout these discussions that were considered most valuable for researchers (mouse control and processing speed). Unfortunately, panelists chose not to elaborate beyond the statement, “If you can’t process information quickly and respond with your mouse accurately, then you will consistently be outplayed by others” (outplayed referring to the opposition performing better than the individual). In the closing stages of this conversation, the discussion led toward different theoretical approaches to best capture the complexity of performance as an adaptive system, with BC stating, “Ecological dynamics theory may allow us to take a more holistic approach to player performance.” Here, the panelist refers to looking beyond solely the identification of performance and, instead, reviewing how we interact with available information within certain contexts.
or environments (e.g., Araújo et al., 2006; Davids et al., 2015). As the focus group naturally concluded, the experimenter revealed the data from phase one of the study (i.e., opinions based on discipline and conversation points from Group A).

**Game Sense.** Of particular focus, the group focused on “game sense” and shared their experiences associated with the term (e.g., hearing players use the term repeatedly). When prompted to expand on the term’s meaning, numerous explanations arose, including advanced cue utilization, prediction, anticipation, imminent awareness, and forward reasoning. However, the panel drew consensus on one method to approach future research (think aloud protocol). The panel, although noting no interest in recording “game sense” to encapsulate performance, did note the value in further research (i.e., why do some individuals claim game sense as a beneficial tool).

**Summary.** Although stating the inherent challenges associated with the term “performance,” Group B reached consensus on three variables (mouse control, reaction time, and processing speed) that they considered may collectively encapsulate CS:GO performance.

**Group C**

**Game Score.** During the review of the first theme of variables, panelist CG stated, “Irrespective of the variables discussed today, surely the primary game metrics are fundamental to capturing performance individually or across a team.” The remaining panelists disagreed, suggesting that other variables, such as reaction time, surely differentiated performance levels better than game score. Panelist CG disputed this by noting how game score was the sole metric that distinguished players or team performance at a professional level and suggested that “authors must be more explicit in what they mean by performance.” The remaining panel agreed that the semantics concerning “performance” must be adjusted. Panelist CA pointed toward literature that “claim to encapsulate esports performance with tasks unrelated to game of focus.” The panel opted to continue with the study’s task at this stage.

**Global Scores.** Each theme continued to be presented to the panel with discussion centered on how each variable of interest may be tested and whether they truly contributed to successful player performance. Six variables gathered substantial consensus throughout this stage that were considered most valuable for researchers to record (percentage accuracy, reaction time, pattern recall, keyboard proficiency, map knowledge, and mouse control). Discussions began by referring to the development of a comprehensive framework of CS:GO performance, pointing toward a game-specific competences structure (e.g., Nagorsky & Wiemeyer, 2020). In line with this, considerable discussion was focused on reaction time. Panelist CB noted, “Can you be a professional without good reaction time? No.” Continuing from this, panelist CA suggested, “You would probably find the best player in the world just by checking their reaction time,” whereas CC argued, “It’s not that simple, a more holistic approach is needed. Perhaps numerous variables related to speed need to be collated to provide a global score.” CG added that “a composite score of task performance variables, alongside outcome variables, may demonstrate the best performers.” When prompted to expand on this, CG discussed how researchers focusing solely on the mechanisms that allowed athletes to perform well, instead of the outcome of their performance, would struggle to distinguish those who are professional or novice. CB agreed with this statement, suggesting, “Someone with the best reaction time may be the worst gamer.” This statement ended with the panelists directing focus toward “game sense.” Unlike the prior panel groups, Panel C suggested that this term was used when players had no explanation of how they performed a certain way. Panelist CE noted, “My players refer to game sense in any situation where an athlete performs better than typical.” Here, panelists noted that the term may refer to “the coming together of numerous mechanisms perfectly [CB].” The panelists suggested that game sense could be the term adopted for “performance composite score” (i.e., a score derived from recording numerous performance variables). The panel concluded by stating that a fruitful avenue of investigation would be to test a performance composite score against professional, semiprofessional, and amateur athletes to determine whether it could successfully discriminate expertise levels.

**Summary.** Group C reached 70% consensus on four variables that they considered may collectively encapsulate CS:GO performance (percentage accuracy, reaction time, keyboard proficiency, and mouse control). Disagreement remained regarding the value of game score (i.e., win or lose).

**Discussion**

The aim of the present study was to harness an expert panel to glean expert opinions on the indicators of CS:GO performance. Given that the field has begun to acknowledge the challenges associated with recording performance (Nagorsky & Wiemeyer, 2020; Pedraza-Ramírez et al., 2020), our study is the first to systematically capture the opinions of international experts to inform the direction of future research. By gathering real-world evidence from domain-specific experts within the esports community, we were able to gather such opinions from a representative sample. In line with the aim of the present study, the panel collectively stated the negative implications of adopting disparate indicators to encapsulate performance. Specifically, panelists called for the value of adopting explicit terminology and extending the study’s conversation of performance indicators into the wider community.

Numerous points of disagreement between panelists or panel groups were presented, with many experts providing abstract, contradicting, or supporting statements based on their own personal accounts. Panelists, when pushed, did not propose any variable of interest that was not drawn from the literature review. We observed a high level of consensus across the discussed variables. Specifically, speed to respond to stimuli (i.e., enemy), the time it takes a person to do a mental task (e.g., respond to a game-specific challenge), the proficiency to utilize a mouse, and the speed to respond to strategy (e.g., bomb planted) were all discussed at length by panelists. Panelists even noted that such variables all followed a similar theme linked to reacting at speed (e.g., actioning a mental process or motor response to a stimulus or change in stimuli quickly). It may be speculated that variables that received the strongest consensus are indicative of the category associated with CS:GO (i.e., fast paced, action). Perhaps if the focus was on differing competitive games, such as online chess, the highlighted variables of interest would differ. In relation to the study’s TEP-highlighted theme, the panelists collectively pointed toward the value of a latent variable approach to incorporate numerous variables that may share commonality of function, as commonly adopted in cognitive literature (Miyake et al., 2000), namely, using several metrics to provide an uncontaminated means to indicate performance (e.g., reduce measurement error). Despite
the advocated use of such composite scores to capture performance, however, there is an absence of research that has supported such an approach in CS:GO, and the field has yet to determine empirically justified performance indicators.

To the authors’ surprise, panel groups held distinct viewpoints on the most valuable indicators of performance; however, some overlap was present. The athletes panel was confident that game score was a clear performance indicator that may set apart differing levels of expertise. Likewise, the panel appeared to share the view that numerous outcome indicators, including kill/death ratio, were valuable indicators of performance success. The athletes’ views were reported to reflect those of their coaches, who noted such in-game metrics after each competitive game. In contrast, Group B made limited references to game score or any other form of in-game outcome indicator, whereas the practitioners of Group C shared the view that outcome indicators (e.g., win or loss) were less valuable for indicating individual performance compared with more traditional performance indicators (e.g., reaction time). It may have been speculated that competitors and researchers may share differing viewpoints based on their competitive, or lack of, sport experience. However, given that a substantial proportion of practitioners appeared to prefer performance indicators, such as reaction time and mouse control, perhaps our findings demonstrate a disconnect between players and practitioners. Given that practitioners are typically informed by researchers, and practitioners utilize evidence-informed approaches to benefit their players, our findings may simply reflect the age of the field (e.g., the recent growth of esports research has yet to reach all players). As such, it may be anticipated that a future iteration of the current study may find more similarities than differences as the field grows.

Although the majority of the TEP panelists did suggest the value of traditional cognitive measures of performance, other panelists did express value in using game outcomes (e.g., game score). The esports athletes and several researchers raised numerous concerns relating to what variables (e.g., win or lose) should be considered the gold standard of performance among empirical investigations. Specifically, one panelist stressed that game score provides a direct marker of success. Panelists extended this view by noting that outcome indicators in a single CS:GO game are ultimately the best-of-30 rounds, whereas a single round is based on planting the bomb, defusing the bomb, or eliminating the entire team. It was the opinion of a limited proportion of TEP panelists that some outcome indicators (e.g., game score) should be used as the gold standard for capturing performance, although no strict consensus was reached. Although maintaining this view, panelists did propose some caution when adopting game outcomes as individual performance indicators. Here, panelists suggested that game score may not be of value for capturing individual performance. For example, irrespective of performance capability, an individual may not be able to redirect the outcome of play given the team nature of the sport. Perhaps, in line with the proposed value of outcome indicators, researchers may wish to direct their focus more narrowly to an individual’s role within the game (e.g., an entry fragger’s primary role is to get the first kill), particularly when the role an individual holds has a substantial influence on their individual game outcome (e.g., kills, bombs planted). In line with our literature review, no literature incorporating role-specific performance indicators was found.

Finally, many panelists were unable to initially define the term performance but provided invaluable comments for future direction. Although numerous attempts have been made to explore individual game-specific performance within esports (e.g., Hopp & Fisher, 2017; Liu et al., 2021; Neri et al., 2021), limited consideration has yet to be presented across esports literature to clarify the use of the term performance in a unified manner. Reflecting this, panel members did point toward the challenges associated with differing terminology within the field of sport performance. This was demonstrated firsthand with numerous disputes concluding upon discovering that both sides of the debate were using the same terminology (e.g., game statistics) for different concepts (e.g., win/loss vs. accuracy). Clarification of terminology is clearly critical for effective discourse, and consequently, the wider esports community must continue to contribute to such a discussion. We now provide some possible considerations for future practice.

**Semantics and Future Directions**

Attempts to address semantics within the field have been outlined briefly (e.g., esports performance, mechanical expertise, cognitive performance; Pedraza-Ramirez et al., 2020); however, further consideration is needed. As seen in occupational literature, authors suggest that when conceptualizing performance, the field must first differentiate the function of interest (e.g., action vs. outcome; Campbell, 1990). Performance as a function of action (i.e., behavior) may refer to the steps that are required by the role to achieve success (Ilgen & Schneider, 1991). Within CS:GO, action performance metrics could encapsulate carrying the bomb, providing utility (e.g., flashbangs), and/or directing teammates. Accordingly, performance may relate to the commonly adopted term skill. Skill can be defined as a mechanism that represents an individual’s capacity to perform an action (e.g., mouse control) but does not represent the outcome (e.g., beating an opponent; Davids et al., 2003). Likewise, skill does not relate to individual expertise level (e.g., amateur vs. professional) and solely represents the proficiency of a behavior (Davids et al., 2003). It could be argued that the adoption of action-based performance tasks may not be used to represent outcome goals (e.g., test the benefit of an intervention) and may, perhaps, more appropriately be used to determine the mechanisms that allow performers to achieve greater task-specific outcomes (i.e., winning). For the benefit of clarity, it could be suggested that this conceptualization of performance may be used to categorize methodologies that evaluate individual processes (e.g., reaction time) from those that record an outcome.

When referring to the function of outcome, the conceptualization is the opposite of action performance. Instead of the process, steps taken, or skill of an individual, outcome performance may be operationally defined as the result of actions (Campbell et al., 1993). Within an esports environment, this may encapsulate the number of rounds won, the ratio of competitive wins to losses, or the outcome of a single match. It is outcome performance, however, that appears to be commonly overlooked by researchers. Perhaps as prior literature so commonly demonstrates that specific behaviors (i.e., action performance variables) lead toward varying outcomes (e.g., Gonçalves et al., 2018; Spratford & Campbell, 2017), some authors may assume that action performance and outcome performance variables are one and the same. However, we wish to express concern over this commonality given that the key determinant of success within any domain is inherently an outcome performance variable (win or loss).

Let us consider if a researcher adopts reaction time as a tool to determine the best CS:GO athlete within a varied sample. If the sample contains both traditional sporting and esports athletes, it may be the case that the highest performing individual holds no CS:GO experience. Particularly, if an individual requires the capacity

(Ahead of Print)
to respond to changes in stimuli rapidly within their own domain (e.g., Formula One driver), it may be argued that this limitation could be avoided by testing in-game reaction time (e.g., an esports-specific reaction time task). Nevertheless, it may remain that an individual without in-game experience may hold near (e.g., another sport) or far (e.g., military surveillance) transfer effects from their own experiences. However, without any in-game experience (i.e., knowledge on the mechanics or rules of the game), it would be surprising if an inexperienced individual had a better outcome performance than an esports athlete within their chosen game.

A fruitful avenue would be to first determine the unique contribution of various action performance metrics (e.g., response time, reaction time, or keyboard proficiency) in the prediction of outcome performance (e.g., winning a game). Once completed, the field may then adopt a biomechanist’s approach (e.g., generating hierarchical technique models) to determine the clear line of relationships between variables that contribute to a successful outcome (e.g., Hughes & Bartlett, 2002). The field may wish to extend this further by determining the directed dependencies among the variables of interest (e.g., structural equation modeling). This approach may enable the field to have more confidence when generating assessments or training interventions to benefit these action performances. As such, future researchers should consider their approach to performance by explicitly outlining the function of performance in which they are interested. Alternatively, researchers may wish to adopt variables that encapsulate both functions (e.g., Behnke et al., 2022).

There is substantial value in identifying action performance metrics from a strategic standpoint (Robertson et al., 2015; Vandorpe et al., 2012; Waleniarzczuk & Stolarski, 2021). Determining the mechanisms that may directly influence outcome performance provides the mechanism to predict success, discriminate experts from novices, and implement training interventions to develop an individual’s capacity to perform effectively (e.g., Blascovich et al., 2004; Jin et al., 2020). The present study provides insight into potentially the most important variables to examine; however, readers should note that not all aspects of individual performance may directly contribute to the primary outcome. In some instances, individuals may become effective or highly proficient at skills that do not support the end product (e.g., winning or losing) but may support the psychological environment (e.g., informational support; as discussed in Trotter et al., 2021). However, the authors note that in some instances, particular skills may directly contribute to the primary outcome without being evident in the game statistics (e.g., efficient motor control). The differences associated with these mechanisms are often distinguished by task performance or contextual performance (see Figure 2; Borman & Motowidlo, 1997; Van Scotter et al., 2000). Task performance, in the context of esports, may include the variables noted within the current paper (e.g., reaction time, mouse control, response time). As such, task performance could be considered an umbrella term for an individual’s capacity to perform skills that collectively contribute to a performance outcome (see Figure 2). To date, a substantial proportion of literature appears to fall into this category (Bickmann et al., 2021; Dykstria et al., 2021; Pluss et al., 2020).

Conversely, contextual performance skills are behaviors that do not directly influence performance outcome (Organ, 1988), such as personal initiative, sportsmanship, and being conscientious of teammates. Prior literature has suggested that skills related to task performance appear to vary between job roles, whereas contextual performance skills appear similar across roles (Motowidlo & Schmit, 1999). Furthermore, contextual performance may not relate exclusively to skills or abilities but, instead, the individual differences in personality (Borman & Motowidlo, 1997). As such, the distinction between task and contextual performance may provide a means for researchers to clarify findings to wider audiences. Researchers must be mindful of how their work may be misinterpreted by others. For example, if researchers conclude that an individual’s capacity to maintain motivation influences in-game performance, researchers may seek to generate interventions that do just that. However, if researchers are simply referring to contextual performance, then such interventions may not lead to improved outcomes. As such, the value of adopting specific terminology may facilitate researchers in further clarifying the investigation’s primary aim.

The authors wish to note that we are not questioning the value of prior investigations utilizing variables that may fit such categories (i.e., task or contextual performance). In fact, the current paper is suggesting that the use of outcome performance may increase the reproducibility of such research when determining individual or team CS:GO performance and beyond. This would allow the field to extend current findings without the concern that a particular variable may not have a direct association with outcome performance. However, given that some researchers may wish to explicitly focus on action performance, the current paper encourages esports researchers to use terminology that truly captures the underpinning aim of their investigation (as opposed to using terms such as game performance).

**Limitations**

The modified Delphi methodology may not only provide insight into numerous aspects of CS:GO performance but could also have implications for the wider esports ecosystem. The methodology allowed for consensus and provided the opportunity for a world leading esports expert panel to share their opinion on an area that has yet to adopt a unified approach to capturing performance. However, the methodology is not without its limitations. Although incorporating global leaders in esports, the invited panel represents a small percentage of the esports community. As such, the authors note that not all experiences can be translated widely, and a different selection of experts may have elicited differing consensus or discussion. The literature review, although extensive, may suffer from the same limitation common across fields that are growing rapidly in that relevant literature may have become available since the conclusion of the data collection period. We further emphasize that the aim of the present study was to advance discussions surrounding CS:GO performance and not to suggest a unified taxonomy.
Conclusion

The aim of the present study was to gather the opinions of esports experts on what indicator(s) may best encapsulate CS:GO performance, with a specific focus on direction for future practice. The study utilized a modified Delphi method to provide direction where evidence was limited and attempted to find consensus where possible. The expert panelists presented numerous opinions on what encapsulates performance, considerations for best practices, and concerns associated with the semantics of performance. Based on the suggested semantics, outcome performance variables (e.g., game score) may represent a robust indicator of esports performance. Numerous action performance indicators were also considered to encapsulate individualized CS:GO performance (reaction time, response time, keyboard proficiency, and mouse control). These action performance variables may contribute to successful outcomes (i.e., task performance) for CS:GO athletes and provide an avenue for future researchers. We hope the present paper provides food for thought for the esports community (i.e., players, practitioners, researchers, etc.) and encourages continued discussion within this space.

References


Thangaratinam, S., & Redman, C.W. (2005). The Delphi technique. The Obstetrician & Gynaecologist, 7(2), 120–125. https://doi.org/10.1576/toag.7.2.120.27071
## Appendix 1: Key Search Terms

<table>
<thead>
<tr>
<th>Categories</th>
<th>Key terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>Electronic gaming, video gaming, online gaming, first person shooter, Counter Strike: Global Offensive</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Esports, CS, CS:GO, FPS</td>
</tr>
<tr>
<td>Performance</td>
<td>Performance, skill, expertise, ability, individual, competitive, score</td>
</tr>
<tr>
<td>Measurement</td>
<td>Measure, measurement, index, categorize, rate, composite, cognitive</td>
</tr>
<tr>
<td>Other</td>
<td>Action, tactical, first-person, team-based, fast-paced, strategy</td>
</tr>
</tbody>
</table>
Appendix 2: Themes, Variable Summaries, and Consensus Rankings

<table>
<thead>
<tr>
<th>Themes</th>
<th>Variable</th>
<th>Summary*cf</th>
<th>Articles</th>
<th>Highly important (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game metrics</td>
<td>Total “kills”</td>
<td>Number of kills during a game/task</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Total “deaths”</td>
<td>Number of deaths during a game/task</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Kill/death ratio</td>
<td>The difference between kills and deaths</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Percentage accuracy</td>
<td>The percentage of successful on-target shots</td>
<td>17</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Shooting distance</td>
<td>The proficiency to hit a target at distance</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Game score</td>
<td>The overall score of the game after a match/task (win/loss)</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Expertise classification</td>
<td>The provided in-game rank (e.g., Gold Nova)</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Skill</td>
<td>Pattern recall</td>
<td>The proficiency to navigate a map/task</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Mouse control</td>
<td>Mouse proficiency (i.e., accuracy, speed)</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Keyboard proficiency</td>
<td>Keyboard proficiency (i.e., accuracy, speed)</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>Cognition</td>
<td>Reaction time</td>
<td>Speed to respond to stimuli (i.e., enemy)</td>
<td>46</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Multiple object tracking</td>
<td>Ability to successfully track multiple stimuli simultaneously (e.g., map and player)</td>
<td>2</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Speed–accuracy trade-off</td>
<td>Total speed (i.e., reaction time/misses)</td>
<td>3</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Processing speed</td>
<td>The time it takes a person to do a mental task (e.g., respond to a game-specific challenge)</td>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>Strategy</td>
<td>Strategy success</td>
<td>Number of “successful” attempts during the game</td>
<td>2</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Strategy failure</td>
<td>Number of “failed” attempts during the experiment</td>
<td>3</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Response time</td>
<td>Speed to respond to strategy (e.g., bomb planted)</td>
<td>3</td>
<td>71</td>
</tr>
<tr>
<td>Awareness</td>
<td>Map awareness</td>
<td>Understanding of the areas of greatest interest</td>
<td>11</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Weapon awareness</td>
<td>Recoil control</td>
<td>1</td>
<td>47</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Map knowledge</td>
<td>Map knowledge</td>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Game economy</td>
<td>Economy knowledge</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Game-specific planning ability</td>
<td>Mentally anticipate the right way to carry out a task or reach a specific goal</td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Game-specific reasoning ability</td>
<td>Effectiveness of drawing inferences, reaching conclusions, and making decisions based on available evidence</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td>Experience</td>
<td>Video game experience</td>
<td>Total time playing video games</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Game-specific experience</td>
<td>Total time playing specific game</td>
<td>56</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Competitive experience</td>
<td>Total time playing competitively</td>
<td>32</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Coach-led experience</td>
<td>Total time spent with coach</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Practice time</td>
<td>Total time spent intentionally practicing</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Team experience</td>
<td>Total competitive experience of the team</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>Vigilance</td>
<td>Game fatigue (tiredness)</td>
<td>Ability to maintain and regulate fatigue throughout a game</td>
<td>6</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Concentration (vigilance)</td>
<td>Ability to maintain and regulate concentration throughout a game</td>
<td>3</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Stress</td>
<td>Ability to maintain and regulate response to stress throughout a game</td>
<td>6</td>
<td>69</td>
</tr>
</tbody>
</table>

*Bold indicates focus group consensus (Phase 2). All summaries were defined by the original manuscripts or, when definitions were lacking, critical friends (i.e., individuals with expertise in CS:GO commentary or research who were not authors or part of the TEP). cf = information provided by a critical friend; CS:GO = Counter-Strike: Global Offensive; TEP = technical expert panel.

*Preliminary average grading (Phase 1).