GUIDELINES FOR EVALUATION
2007-08 Team and Individual Problem Solving
Future Problem Solving Program International

PURPOSE OF EVALUATION

The primary purpose of Future Problem Solving (FPS) evaluation is to provide coaches and students with feedback that allows them to develop and improve their problem solving skills. Team problem solving is performance-based, and evaluation is authentic assessment of the team's booklet. Because there is no single "right" answer, FPS employs a variety of strategies to review student work. Specific criteria evaluate performance in each step. Skill improvement remains the most important aspect of evaluation; however, FPS also involves competition and competition scoring must be impartial. A secondary purpose of evaluation is to provide a fair and reliable method for comparing teams in a Future Problem Solving competition.

ATTITUDE

It is essential that evaluators maintain a positive attitude throughout the evaluation process. The central purpose of the Future Problem Solving Program International is to assist students in acquiring better thinking, communicating, and problem solving skills. Evaluation is always done with this thought in mind. **The better evaluators offer constructive feedback and make students want to improve their problem solving skills.** Effective feedback praises students for what they did well and encourages them to use their improved skills to tackle the next problem. **Regardless of the quality of the student effort, critical feedback may discourage a team and keep them from improving. This defeats the purpose of the program.**

It is important for evaluators not to confuse their personal expectations and skill level with that of the teams represented in an evaluation sample. Do not confuse the sophistication of the task with that of the students. Consider the age/division of the team and the level of competition (practice or competition) in constructing positive feedback. Once you see an exceptional booklet, it may be easy to expect the same quality from all booklets. Remember, the completion of an FPS booklet is, by itself, a major accomplishment, possibly more demanding than anything else the students have done as part of their educational experience. Students will delight, frustrate, and eventually reward the demanding task of evaluation.

The ability to consistently provide positive and constructive feedback is the goal to which all evaluators must aspire.

PREPARATION

Before evaluating booklets for any topic, evaluators should have knowledge of the topic. The students put tremendous effort into their work. They can really lose respect for anything you say — no matter how valid it is — if you do not show a basic understanding of the material. Reading the topic chapter in *Readings, Research, and Resources*, or several articles from the bibliography, is a good way to gain topic knowledge necessary for evaluation.

Reading, discussing, or contemplating the ideas presented in the future scene and in the topic evaluation notes help bring a high level of consistency to FPSP evaluation.

SCANNING

When possible, before scoring begins, it is recommended the evaluator read through the sample of booklets he or she is evaluating. This review should give the evaluator an idea of the sophistication of the challenges, underlying problems, and solution ideas. It also gives the evaluator an overview of the least common and the most common responses. However, in competitive evaluation, scanning all of the booklets is not always possible.

SCORESHEET COMPONENTS

Three elements require evaluator attention on the Future Problem Solving Program scoresheet:

- *Identification* – An evaluator fills in the identification portion at the top of the scoresheet before evaluating a booklet.
- *Feedback* – Space is provided on the scoresheet for both general and specific feedback. ***Feedback is the most important aspect of the evaluation***.
- *Scoring* – Use the descriptors on the scoresheet for each criterion to determine the numerical score.

FEEDBACK

Feedback helps a team understand its strengths and weaknesses and motivates them to improve their skills. Feedback is ***the most important aspect of the evaluation*** and is given for each step of the FPS booklet. Feedback enables students to focus their learning process and allows the coach to adapt problem solving instruction to meet the needs of the students.

Comments, both general in nature and specific to a single response, are written on the scoresheet.

Using a strategy devised by Edward de Bono (1974) improves the quality of feedback. Edward de Bono suggests that attention be given to the following four areas when responding to students about their problem solving: praise, clarification, criticism and amplification.

Praise: Positive acknowledgement of a team's effort, creativity, and major strengths

- Reinforces positive aspects of a team's performance
- Establishes a good working relationship between the evaluator and problem solvers
- Rewards the team for facing a problem and developing a solution idea
- Reminds the team, even if their score is not high, they did some things right and encourages the team to improve

Clarification: Evaluator comments asking teams to clarify ideas

- Points out statements that may be confusing or unclear and offers suggestions for improvement
- Encourages teams to improve the clarity and elaboration of their work
- Promotes the development of effective communication skills

Criticism: Evaluator suggestions for areas needing improvement

- Specific constructive comments help teams build their skills
- Gives teams examples of ways to use their ideas, research, or the problem solving process more effectively
- Encourages teams to learn from their work to become better problem solvers

Amplification: Evaluator comments that help students expand their ideas, push their thinking even further, and improve the quality of their problem solving

- Points out gaps in information or logic
- Helps teams improve their planning for an FPS booklet
- Identifies other ideas the team might consider
- Prompts teams to consider the possible consequences of their ideas
- Lists positive, constructive ideas for improvement

TYPES OF SCALES

Criteria in FPS evaluation are grouped into four categories of scales:

- **Frequency scales**: Points awarded based on a count of responses that meet specific criteria
- **Rating scales**: Points awarded based on the degree or extent to which a team meets a descriptor
- **Weighted scales**: Bonus points awarded for fluent or relevant responses found infrequently and are indicative of high-quality thought
- **Composite scales**: Points awarded based on a total of points earned on individual elements

The type of scale used for each criterion is identified following the criterion's description.

TYPES OF CRITERIA

The FPSP criteria measure student skills in the following three categories:

- **Content**: These criteria measure the *quality of the content* in students' work. Content-oriented criteria evaluate the *merit of the ideas*.
- **Structure**: Structure-oriented criteria assess how effectively students fit their work into a prescribed format. This measures a team's mechanics in completing their booklet.
- **Process**: These criteria judge how well students use the FPS model.

STEP 1 - IDENTIFY CHALLENGES

Objective: Identify 16 challenges within the future scene

A challenge is an issue, concern, or problem that needs attention or consideration (points of importance). A challenge logically relates to the future scene, tells what the concern is, and why it is a concern.

Fluency 1 2 3 4 5 6 7 8 9 10
Flexibility 1 2 3 4 5 6 7 8 9 10
Clarity 1 2 3 4 5 6 7 8 9 10
Originality ..... 3 x ____ = _____

Fluency (1-10) – The numerical score is based upon the number of challenges awarded. A higher score is earned if most of the challenges show a clear understanding of the future scene.

Each challenge is individually read and classified in one of the following categories:

YES – Challenge has a ***strong*** possibility of existing or occurring if the future scene were to occur. This is a challenge seen as a logical ***cause*** of the future scene **OR** a challenge that ***results*** from the actions within the future scene. **(A challenge does not have to be written as a cause/effect statement.)**

PERHAPS – (1) Challenge has ***some*** possibility of existing or occurring if the future scene were to occur. (2) Challenge is worded poorly or ambiguously and would have been scored Yes had it been stated more clearly.

WHY – (1) Challenge does not appear to have a clear connection to the future scene. -- (2) Statement does not identify a challenge.

SOLUTION – A response that suggests **how** to solve a challenge of the future scene. Identifying issues that might result from a solution to the future scene is not the same as identifying actual challenges in the future scene. Citing challenges resulting from a solution without defining the challenge or indicating why it is a challenge causing the future scene or resulting from it is an elaboration of **how** to implement the solution.

DUPLICATE – Any challenge too contextually similar to another accepted challenge cannot receive additional credit. *Evaluators should not confuse duplicate ideas with duplicate categories – it is acceptable for students to list several different ideas in the same category.*

Each challenge statement is read and marked as **Y** – yes, **P** – perhaps**, W** – why, **S** – solution or **D** – duplicate in the appropriate column on the scoresheet. After evaluating each of the team's challenges, count the number of Yes challenges, indicate the number in the last row of

the Yes column, and circle the fluency score based on the scale provided on the scoresheet. Write the fluency score in the Step 1 score box. *Frequency; Content and Process*

Flexibility (1-10) – The numerical score is based upon the number of categories used in writing the challenges scored as **Yes**. A more varied approach to the future scene allows a more complete picture of the whole situation. Evaluators take each of the following areas into consideration in scoring booklets for flexibility:

- Categorize the **Yes** challenge responses on the scoresheet. Evaluators can use the list on page 16 of *Guidelines for Evaluation* or on the scoresheet. If you choose, you may create your own categories appropriate to the future scene.
- Determine the variety of points of view taken within the **Yes** responses. Does the team look at what caused the challenges in the future scene as well as the effects of the future scene? Do the challenges, even within one category, examine the concerns of a different number of "clients" (e.g., in the "Natural Disaster" future scene there are economic challenges for individuals, businesses, governments, and insurance companies)?

Count the number of different categories and determine the numerical score according to the scale on the scoresheet. *Frequency; Process*

Clarity (1-10) – Clarity evaluates the description of the challenges identified. A challenge with a clear and thorough description of the concern demonstrates good clarity; a clearly written challenge shows effective communication skills. Challenges lacking clarity are more often challenges scored as **Perhaps** or **Why**. Teams who consistently state what the challenge is, why it is a challenge, how it logically relates to the future scene, and the causes/consequences of the challenges should receive a high score. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Structure*

Originality (3x) – This scale rewards the identification of a rare, high-quality challenge that received a **Yes** in fluency**.** Any response found infrequently among responses at that age/grade level and considered of high quality (insightful, indicative of breakthrough thinking) is scored original. Mark **Original** (a check mark or O) in the O column for each **Yes** challenge judged original. Enter the total number of originals on the scoresheet and multiply by 3. *Weighted; Content*

STEP 1 EVALUATION ESSENTIALS:

1. Student work must relate to the future scene. A future scene is a hypothetical, what if, scenario based on current information. In cases where research can be found that contradicts the future scene, the team is still required to problem solve within the boundaries of the given situation.
2. Step 1 challenges are written in statement form. Questions are inappropriate for this step.
3. Students state challenges in terms of *possibility*. Challenges are stated with possible (i.e., non-absolute) terms such as could, may, might, etc. Absolute terms that indicate that *x **will** be a challenge* deny an important element of projecting into the future—it is impossible to know what will or will not occur in the future. We can only make educated guesses as to possible occurrences, based on an investigation of the resources.
4. Students should word challenges clearly. Students must tell what the challenge is, why it is a challenge, and how it logically relates to the future scene.
5. Students should demonstrate flexibility in their thinking and explore challenges from different perspectives or categories.

STEP 2 - SELECT an UNDERLYING PROBLEM

Objective: To identify and state an important part of the future scene to solve

Structure 0 1 2 3 4 5 6 7 8 9 10
Adequacy 1 2 3 4 5 6 7 8 9 10

Focus 1 2 3 4 5 6 7 8 9 10

Structure (0-10) – Structure is an assessment of the composition of the underlying problem (UP). At this point, evaluators do not judge the effectiveness or value of the underlying problem. Instead, structure score simply determines if all the parts are there and are correctly written. A well-structured underlying problem is one that correctly includes all four required elements — condition phrase, stem + key verb phrase, purpose, and future scene parameters.

Condition Phrase (0, 1, <u>or</u> 2 points) — The condition phrase is a lead-in fact or logical assumption from the future scene or related research that is the basis for or cause of the challenge the team chooses as its underlying problem (e.g*., Because little is known about the new natural disaster of sonic shock* ...").

Stem + Key Verb Phrase (0, 1, 2, <u>or</u> 3 points) — The underlying problem includes a **stem** ("How might we" or "In what ways might we") and **one key action verb** in a phrase indicating what to do to solve the challenge. (e.g., "Because little is known about the new natural disaster of sonic shock*, how might we provide financial support for insurance companies* … ?") The key verb phrase (KVP) is one key action verb in a phrase that *mandates* what must occur in Step 3 to solve the underlying problem.

Purpose (0, 1, 2, <u>or</u> 3 points) — The purpose specifies an optimal direction, goal to pursue, or reason for solving the challenge. (e.g., "Because little is known about the new natural disaster of sonic shock, how might we provide financial support for insurance companies *so they will be able to continue to provide the protection property owners need?*") **The purpose should give further information about what will be accomplished if the underlying problem is solved; it is not a repetition of the condition phrase or key verb phrase.**

Future Scene Parameters (0, 1 <u>or</u> 2 points) — The future scene parameters (FSP) place the underlying problem within the confines of the future scene. These parameters include time, place, and topic. (e.g., "Because little is known about the new *natural disaster of sonic shock,* how might we provide financial support for insurance companies so they will be able to continue to provide the protection property owners need *after 2032 in the Sound Tunnel?*")

Determine the numerical score by adding the points awarded in all four components as outlined on the scoresheet. At this point, evaluators are not deciding if the *content* of the team's underlying problem is good or bad. Structure is a question of *composition*. Did the team write the underlying problem correctly? The merit of the underlying problem is evaluated in adequacy.

Condition Phrase

- 2 points: The condition phrase relates to the key verb phrase and uses accurate information from the future scene and/or from related research on the topic.
- 1 point: The condition phrase does not use accurate information or reasonable assumptions, or it does not relate to the key verb phrase.
- 0 points: The condition phrase is missing.

Stem & Key Verb Phrase

- 3 points: The key verb phrase is present and contains a single active verb or verb phrase.
- 2 points: The key verb phrase is present but has two objects or two modifiers.
- 1 point: The key verb phrase is present but has two verb phrases.
- 0 points: The key verb phrase is missing.

Purpose

- 3 points: The purpose is present and has a single focus with a logical relationship to the key verb phrase.

- 2 points: The purpose is present but does not have a clear relationship to the key verb phrase.
- 1 point: More than one purpose is present.
- 0 points: The purpose is missing.

Future Scene Parameters

- 2 points: All 3 parameters of topic, place, and time are present.
- 1 point: Two of the three parameters are present.
- 0 points: Only one or none of the parameters are present.

Focus (1-10) – The challenge identified in the underlying problem should be a smaller part of the entire future scene. It should narrow the future scene without trivializing any part of it. Full credit is awarded to an underlying problem that identifies a significant area of concern of the future scene. A lower score is given if the challenge identified is too broad or too narrow. (The challenge identified in the underlying problem must be derived from a challenge(s) the team generated in Step 1. If this is not the case, score the underlying problem between 1 and 3, based on the quality of its focus.) Use the descriptors on the scoresheet to determine the numerical score. *Rating; Process*

Adequacy (1-10) – Adequacy assesses the importance of the problem area the team chooses as its underlying problem. Adequacy measures two aspects of the underlying problem:

- Has a challenge of the future scene been identified, as opposed to a future scene fact, a non-challenge, the whole future scene, or something outside the future scene?
- Is it important to solve the underlying problem in relation to the major challenges of the entire future scene?

While focus measures the scope of an underlying problem, adequacy measures its significance or merit. Higher adequacy scores are awarded to teams that identify an important issue within the future scene. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Content*

STEP 2 EVALUATION ESSENTIALS:

1. An underlying problem is stated as one question containing four basic components:

   Condition Phrase: A condition phrase is a lead-in fact or logical assumption from the future scene or related research that is the basis for or cause of the challenge the team chooses as the focus of its underlying problem. The condition phrase guides (forces) the team to make a connection to the future scene and the Step 1 challenge(s) it uses as the focus of its underlying problem. (e.g., *Because our town is within the area called the Sound Tunnel where a sonic shock could occur*...") The condition phrase explicitly or implicitly indicates what the challenge is from Step 1 that the team selects to solve in Step 3. The possible occurrence of a sonic shock is the area of concern from the Step 1 challenge(s) the team selected as the focus of its underlying problem.

   Stem + Key Verb Phrase: The obligatory stems include, "How might we…" or "In what ways might we..." The key verb phrase is one key action verb (and/or object) in a phrase that *mandates* what must occur in Step 3 to solve the underlying problem. The KVP provides direction for Step 3. Relevant solutions are those that do what the KVP *mandates*. If the underlying problem begins, *"Because our town is within the area called the Sound Tunnel where a sonic shock could occur, how might we <u>organize a plan for preparing our community</u>…?"* (KVP is underlined), then Step 3 solution ideas must *"organize a plan for preparing our community"* to earn credit as relevant.

   Purpose: The purpose provides direction for the KVP. It is an ultimate goal toward which teams strive as they attempt to solve the underlying problem. Solution ideas must satisfy the

purpose to be considered relevant. For example, if the underlying problem is *"Because our town is within the area called the Sound Tunnel where a sonic shock could occur, how might we organize a plan for preparing our community in order to reduce the number of our citizens who would be injured…?"* (Purpose is underlined), then relevant solution ideas must not simply "organize a plan for preparing our community", but they must do so in such a way that it will *"reduce the number of our citizens who would be injured."*

Leaving out the purpose can negatively affect scores in both Step 2 and Step 3. In Step 2, teams lose 3 points in structure and score from 1-3 in focus. In addition, evaluators impose a purpose that seems logical to the future scene and the KVP. Solution ideas in Step 3 will be scored for relevancy. Solutions are scored very strictly against the KVP, the imposed purpose, and the future scene parameters (see below). In competitive situations such as the affiliate final or International Conference, a team whose UP has no purpose has a very hard time advancing to further rounds of evaluation.

FSP - Future Scene Parameters: Elements that place the underlying problem within the parameters of the future scene, including time, place, and topic

Including the parameters of the future scene in the underlying problem ensures the challenge is one that is a sub-area or subtopic of the future scene. The parameters are found in the future scene and include place (geographic location involved), topic (major focus of the future scene), and time (the date or time reference).

Relevant solution ideas to the underlying problem do not contradict any part of the FSP. If your underlying problem is "*Because our town is within the area called the Sound Tunnel where a sonic shock could occur, how might we organize a plan for preparing our community in order to reduce the number of our citizens who would be injured if this natural disaster were to occur here in 2032?"*, relevant solutions must *organize a plan for preparing our community* ***and*** do so in such a way that it *will reduce the number of our citizens who would be injured.* They must do so, however, in relation to the place – Sound Tunnel, topic – natural disasters and time – 2032, identified in the future scene. Even if a team forgets to include the FSP in their underlying problem, solutions still must fall within these parameters to be scored relevant in Step 3.

The underlying problem used in the example above has been broken down into the four basic components as shown below:

> Because our town is within the area called the Sound Tunnel (FSP) where a sonic shock could occur (C), — how might we organize a plan for preparing our community (S + KVP) — in order to reduce the number of our citizens who would be injured (P) — if this natural disaster (FSP) were to occur here in 2032 (FSP)?

2. The underlying problems may be composed of ***one challenge*****,** ***one area or category of concern, or a compilation of several related challenges*** the team has identified in Step 1 and wishes to solve in Step 3.
3. An underlying problem is never as large as the future scene. Encourage a team to identify an underlying problem focusing on one concern or area of concern. An underlying problem that restates the entire future scene is inappropriate. The following are examples of how an underlying problem restates the future scene:
   - *How might we overcome (or develop remedies for, decrease) the challenges created by the sonic shocks in our country in the year 2032?*
   - *How might we reduce the problems resulting from sonic shocks in our country in the year 2032?*

   An underlying problem equivalent in context to these examples is a restatement of the future scene. Without narrowing the future scene, a team not only misunderstands the FPS process, but it also places itself in a competitively advantageous position because "relevancy" in solution ideas is more easily obtained for an ***unqualified*** underlying problem. To balance this inequity, ***booklets with underlying problems that restate the entire future scene, as described above, receive a score of one (1) in focus and a score of***

one (1) in adequacy in Step 2. Evaluators are instructed to consider this when scoring creative strength because it is a critical error in the creative process.

4. The underlying problem addresses only one issue. Using more than one verb or verb phrase results in a score of 1 for stem + KVP. Examples of multiple verbs follow.
   - *How might we (HMW) select and educate jurors…?*
   - *HMW protect the natural habitats of Antarctica's animals while at the same time allowing tourists to take advantage of ecotours…?*
5. Some underlying problems contain multiple objects or descriptors in combination with a single verb in the key verb phrase and receive a score of 2 for stem + KVP in structure. Examples of this follow.
   - *HMW solve the problems of child abuse and drug abuse…?*
   - *HMW provide economic and counseling opportunities…?*
   - *HMW distribute food and jobs…?*
6. If an underlying problem contains two verbs or verb phrases, evaluators refer only to the first verb when scoring focus and adequacy in Step 2. **Solution ideas, however, must be relevant to *everything mentioned* in the underlying problem to be scored relevant in Step 3**.
7. A score of 1 is awarded for focus and for adequacy if, in its essence, the underlying problem:
   a. ***Restates*** the entire future scene (e.g., "How might we solve the challenges related to sonic shocks?") or
   b. ***Broadens*** the future scene (e.g., "How might we solve the challenges resulting from natural disasters?"); or
   c. ***Does not identify a challenge related*** to the future scene (e.g., an unrelated underlying problem, such as, "How might we increase participation in the Red Cross ...?" or an underlying problem that does not identify *any* challenge, such as, "How might we make Dundee a model for how to recover from a sonic shock?").

   In addition, teams that restate, broaden, or do not identify a challenge in the underlying problem score lower in effectiveness (Step 6) and creative strength (Overall). It would be difficult for any solution idea to *effectively* solve such a broad underlying problem. Using any of these types of underlying problems is a critical error in a strategic step of the creative problem solving process.

   If the underlying problem broadens the future scene or does not identify a challenge related to the future scene, the evaluator imposes the future scene parameters on the underlying problem when scoring relevance in Step 3. This means the evaluator scores Step 3 as if the team had *restated* the future scene.
8. The use of an absolute verb in the KVP, such as stop, prevent, and eliminate, receives a score of one to three (1-3) in focus and one to three (1-3) in adequacy. If a team's KVP is …prevent drug abuse…, then Step 3 relevant solution ideas must totally negate all drug abuse, which would be a very difficult task, particularly in a global setting where many of the future scenes are positioned. A team that uses …reduce drug abuse… in the KVP will be more successful in writing relevant solution ideas in Step 3.
9. If the underlying problem has no purpose, evaluators:
   a. Score 0 points for purpose in structure (see structure pages 5-6)
   b. Score focus on a scale of 1 to 3 (a score of 3 is the highest possible score on focus)
   c. Impose a purpose that seems logical to the future scene and KVP and evaluate solution ideas for relevance based on the imposed purpose. It is helpful if an evaluator writes the imposed purpose on the scoresheet so the team knows what it is.
   d. If a solution idea does not address the imposed purpose, it cannot be scored relevant.

e. Provide feedback that explains the usefulness of the purpose to help focus the underlying problem. Help the team understand the connection between the purpose in the UP and relevance of solution ideas in Step 3.

f. A low score in focus and adequacy is awarded if the purpose repeats the condition phrase or key verb phrase. The purpose should give further information about what will be accomplished if the underlying problem is solved; it is not a repetition of the condition phrase or key verb phrase.

STEP 3 - PRODUCE SOLUTION IDEAS

Objective: To create 16 varied and unusual solution ideas responding to the underlying problem

Fluency 1 2 3 4 5 6 7 8 9 10
Elaboration 1 2 3 4 5 6 7 8 9 10
Flexibility 1 2 3 4 5 6 7 8 9 10
Originality ..... 3 x ____ = _____

Fluency (1–10) – Fluency measures the number of solution ideas relevant to the underlying problem. All solution ideas are individually read and put into one of the following categories:

RELEVANT**:** A solution idea, if implemented, achieves the goal of the underlying problem. More specifically, a relevant solution idea addresses the issues of the condition, answers the key verb phrase and purpose, and occurs within the future scene parameters. **If either the KVP or purpose is multiple, all the stated goals must be addressed in the solution idea to earn credit as relevant.**

At this point, you are not evaluating whether a solution idea is good or bad, humane or inhumane, cost efficient or expensive. You are determining whether the solution idea is relevant to the underlying problem.

PERHAPS**:** A solution idea does not have a clear connection to the goals of the underlying problem

WHY**:** A solution idea is unrelated to the underlying problem or the statement does not identify a solution idea

DUPLICATE**:** A solution idea too contextually similar to another solution idea previously scored as relevant is considered a duplicate idea. **Contextually similar or trivial solution ideas may be combined and several counted together as one relevant solution idea.** *Evaluators are cautioned not to confuse duplicate solution ideas with duplicate categories. It is acceptable for students to list several solution ideas in the same category.*

Each solution idea is read and marked as **R** – relevant, **P** – perhaps, **W** – why or **D** – duplicate in the appropriate column on the scoresheet. After evaluating each solution idea, count the number of relevant responses, and determine the numerical score according to the scale on the scoresheet. *Frequency; Process*

Elaboration (1–10) – An elaborated solution idea is any relevant solution idea that includes at least three of the who, what, why, and how elements. Where and when may be counted toward elaboration only if they are of a substantive nature. Teams do not have to write a paragraph in order to earn elaboration points, nor should they earn elaboration points just for writing a paragraph. **Solution ideas elaborated by simply adding on the key verb phrase and/or purpose are not considered for elaboration credit more than three times in a booklet**. Evaluators should provide feedback that contains suggestions on ways to specifically elaborate solution ideas. Mark an E next to each relevant solution idea that qualifies as elaborated. If no credit is awarded for elaboration, leave the space blank. Count the number of solution ideas marked as elaborate and determine the numerical score according to the scale on the scoresheet. *Frequency; Structure*

Flexibility (1–10) – Refers to the diversity of thought in relevant solution ideas. If the solution idea is not relevant, it is not scored for flexibility. Categorize the relevant solution ideas in the team's booklet. (The category list appears on the scoresheet.) Count the number of different categories and determine the numerical score according to the scale on the scoresheet. *Frequency; Process*

Originality (3x) – Reward teams for thinking that is especially insightful or creative. Students have a chance to show their creativity and futuristic thinking through unique solution ideas. Although evaluators encourage creativity, wildly futuristic ideas are not necessarily original. A relevant solution idea must have substance for a team to receive originality points. Mark "ORIGINAL" (a check mark or O) next to each relevant solution idea judged original. Enter the total number of originals on the scoresheet and multiply by 3. *Weighted; Content*

STEP 3 EVALUATION ESSENTIALS:

1. Solution ideas are written in statement form as definite proposals.
   a. A relevant solution idea must address the condition, the key verb phrase, and the purpose of the underlying problem. It must also occur within the confines of the future scene parameters.
   b. A team does not have to repeat the condition, key verb phrase, and/or purpose verbatim for a solution idea to be relevant. A solution idea can be relevant even if the purpose is not specifically mentioned.
   c. Does the solution idea solve the underlying problem? A relevant solution idea does not have to work perfectly, be humane, be tried and true, or be new. How well the solution idea solves the underlying problem is judged in Step 6.
   d. If the underlying problem has no purpose, the evaluator imposes a purpose that seems logical to the future scene and the key verb phrase. Evaluators use the imposed purpose to judge the relevance of each solution idea. In competitive evaluation, solutions are judged strictly on meeting the terms of the key verb phrase, the imposed purpose, and the future scene parameters. In other words, a team whose underlying problem has no purpose has a very hard time advancing to further rounds of evaluation.
2. On practice problems, Step 3 feedback should emphasize the connection between the purpose in the underlying problem and relevance of solution ideas in Step 3.
3. Teams should elaborate each solution idea. They may explain who, what, how, and why to gain elaboration points. While it is helpful to include when and where, these will only be counted toward elaboration if substantive in nature *("in the year 2007 in the US"* is not substantive).

   Who indicates who will implement the solution idea; the "who" should identify the person or group capable of carrying out the solution. A pronoun such as *we*, *they*, *he/she* is not sufficient to count as a "who."

   What states the solution idea.

   How describes how the solution idea will work.

   Why explains why the solution idea will solve the underlying problem (KVP and purpose).

   Here is an example of an elaborate (*over* elaborate, for the sake of illustration) solution idea utilizing who, what, how, why, and a substantive where and when:

   Who: The local unit of the Red Cross

   What: will prepare a temporary shelter

   How: by stocking supplies to feed and care for the people of our community

   Why: so people will have a safe and secure place to relocate

   Where: in a large strong building, such as a warehouse or school, in a nearby community

   When: during the three-day warning time before the sonic shock occurs.

5. Students should demonstrate flexibility in their thinking and explore solution ideas from different perspectives or categories. Imaginative inventions are fun, but must include how or why explanations. **An invention can't happen just because someone says it will.**

STEP 4 – GENERATE and SELECT CRITERIA to SOLUTION IDEAS
Objective: To develop yardsticks that evaluate (measure) the creative potential or the importance of solution ideas

A criterion **measures** how well the solution idea does what the underlying problem mandates. Teams are encouraged to generate and select criteria that measure the comparative quality of solution ideas.

Correctly Written……0 1 2 3 4 5
Applicability and Relevance……..0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Correctly Written (0-5) – Correctly written is a matter of structure. Each criterion focuses on a single dimension, demonstrates a measure of degree, and indicates the desired direction. Criteria not scored as correctly written **are** evaluated for applicability and relevance.

At this point, the evaluator is not deciding the ***value*** of the team's criteria in ranking the solution ideas. Correctly written is a question of ***structure.*** Place a check mark in the appropriate box on the scoresheet to indicate if a criterion is correctly written. Award one point for each correctly written criterion. *Composite; Structure*

Applicability and Relevance (0–15) – This scale assesses the merit or value of each criterion and determines the extent to which each criterion is an important consideration in evaluating the underlying problem.

CRITERIA EVALUATION
Adopted June 2007

It turns out that criteria may be much more complex than some of us thought! Many also feel that criteria are much more important than sometimes perceived, and are especially applicable to real life situations. Criteria can be used to lead us to the ideal best solution, or they can lead us to a solution that is practical to implement. The information here is a start to help us understand criteria more thoroughly.

STEP 4 GLOSSARY
Quite a few terms have been used to describe criteria, sometimes with completely different meanings depending on who is using the term. Here is a list of the terms and definitions we will be using for international evaluation.

A) RELEVANCE
- ***RELEVANT***: A criterion that is appropriate for evaluating solutions to this underlying problem.
- ***NOT RELEVANT***: A criterion that is not appropriate for evaluating solutions to this underlying problem.

B) TYPES OF CRITERIA
- ***UP-BASED: KVP or UP-BASED: PURPOSE***: A criterion that is based on the ideas from the underlying problem, either from the key verb phrase or from the purpose.

- ***GENERIC***: A criterion whose core idea can be applied to nearly every solution to nearly every underlying problem for nearly every topic. ***A generic criterion does NOT become specific by adding future scene information to it.***
- ***SPECIFIC***: A criterion whose core idea can only be applied more narrowly, to this underlying problem and/or for this topic, or for other closely related topics, but not to a wide variety of areas. It may be based on the research for a topic. *A specific criterion does NOT need to have future scene information in it. However, sometimes it is hard to tell if a criterion is generic or specific without modifying information.*

C) "FLAVORS" FOR GENERIC AND SPECIFIC TYPES

- ***PLAIN***: A criterion with no details or information coming directly from the future scene.
- ***MODIFIED***: A criterion with adequate details, such as important stakeholders, to place it firmly within this underlying problem and/or future scene. Parameters alone (time, place, topic) are not enough for modification. Try to avoid adding a complete KVP or purpose as a modification.
- ***JUSTIFIED***: A criterion that gives a condition from the future scene as a rationale for its use. The justification must have a logical connection to the criterion.

STEP 4 EXAMPLES

Underlying Problem: Because there is a controversy regarding Navitas users receiving SI transplants, HMW in 2032 in the US maintain the equitable distribution of SI transplants so that potential donors continue to donate their organs? (*taken from the Evaluation Guidelines*)

Not relevant	Which solution will prevent identity theft the most? (0 points)

UP-Based: KVP	Which solution will provide the fairest method of determining who gets the SI organs? (3 points)
UP-Based: Purpose	Which solution will most convince people to continue donating their organs? (3 points)

Plain Generic	Which solution will be most accepted by the people? (1 point)
Modified Generic	Which solution will be most accepted by the Navitas users who need SI transplants? (2 points)
Justified Generic	Since 4% of the population continues to use Navitas that can lead to SI failure, which solution will be most accepted. (3 points)

Plain Specific	Which solution will be the most legally binding? (3 points)
Modified Specific	Which solution will be the most legally binding on the organ donors? (3 points)
Justified Specific	Since ethically and legally all patients must be treated without discrimination, which solution will be the most legally binding on the organ donors? (3 points)

STEP 4 SCORING

****International Terms**: **As shortcuts, the terms generic, modified, and advanced will be used to designate 1, 2, or 3 points.** There are multiple ways to earn 3 points, all part of the "advanced" group of criteria. Note that ALL specific criteria receive 3 points, but generic criteria must be "justified" in order to receive 3 points. If a justification does not come directly from the future scene or have a logical connection to the criterion, the justification is ignored and the criterion is scored on its own.

Evaluation Shortcut Terms	Type & Flavor	Points
NR - Not relevant	Not relevant	0
G - Generic	Plain Generic	1

M – Modified	Modified Generic	2
A – Advanced	Justified Generic	3
A – Advanced	UP-Based: KVP	3
A – Advanced	UP-Based: Purpose	3
A – Advanced	Plain Specific	3
A – Advanced	Modified Specific	3
A – Advanced	Justified Specific	3

STEP 5 – APPLY CRITERIA to SOLUTION IDEAS

Objective: To develop an evaluation matrix to determine the best solution idea to use in developing an action plan

The team selects its eight most intriguing solution ideas and applies the criteria by using the evaluation matrix (grid) in determining the best solution idea to use for the Step 6 action plan.

Correctly Used 1 2 3 4 5

Step 5

Correctly Used (1–5) – Examine the evaluation matrix (grid) to ensure the numbers in it have been used correctly.

a. Each vertical column should include the numbers 1 (low) to 8 (high) or 1 through the number of solutions entered in the grid.

b. Each number should be used only once with one exception: if a team feels two solution ideas satisfy a criterion equally, the two ranks that would have been given are averaged. Therefore, half-points are used for "ties."

c. A single criterion may be weighted if a team feels that criterion is especially important. It is essential, however, that numbers are still entered correctly. If a criterion is double weighted, the numbers 2 (low) through 16 (high) would be used in increments of two.

d. Correct use of the grid includes ranking the solution ideas and accurate addition in determining totals. Totals for each solution idea are determined by adding the ranks across the columns.

e. An automatic one (1) point is assigned if a team does not use the solution idea that scored highest on the grid in developing its action plan.

Evaluators should subtract 1 point for every error in the grid (i.e., column rank-ordering or addition error). More specifically, a perfect grid equals 5 points, one error equals 4 points, two errors equal 3 points, three errors equal 2 points, and four or more errors equal 1 point. *Composite; Structure*

STEP 4 & STEP 5 EVALUATION ESSENTIALS:

1. Each criterion focuses on a single dimension, demonstrates a measure of degree, and indicates the desired direction. A correctly written criterion satisfies all three areas. Although many teams write criteria in a question format, doing so is not required.
   a. **Single dimension –** Each criterion focuses on only one concern or dimension.
   b. **Measure of degree –** Each criterion uses a superlative such as most, least, fewest, greatest, etc. These superlatives must be measurable.
   c. **Desired direction –** Criteria are phrased in the direction of the desired outcome.
   d. **Use superlatives** – When phrasing criteria, sue least, most, greatest, fewest, criteria (for example, "Which solution is *the easiest to implement*?" or "Which solution provides *the greatest social benefit*?").
3. Word criteria to indicate a *desired direction*. Ask, "Which solution is the *easiest* to implement?" instead of "Which solution is the *hardest* to implement?" "Which solution provides the greatest social *benefit*?" instead of, "Which solution provides the greatest social *harm*?"
4. Half-points are used for ties in rank-ordering numbers in the grid. A team may choose to weigh certain columns. This is acceptable as long as numbers are entered correctly.

5. Correct use of the grid includes accurate addition in determining totals.
6. An automatic one point is assigned if a team does not use the solution that scored the highest in developing its action plan.
7. Subtract one point for every error in the grid (i.e., column rank-ordering error or addition error). More specifically, a perfect grid = 5 points, one error = 4 points, two errors = 3 points, three errors = 2 points and four or more errors = 1 point. *Composite; Structure*
8. If there is a tie for best solution in the grid, teams must choose to use one or the other. The team is not required to explain to evaluators the reasoning behind its choice. (See *Key Tips* for suggestions to teams about how to break ties.)
9. If a team combines two or more unrelated solutions to develop the action plan, the grid should receive an automatic one (1) point for correctly used. The highest scoring solution in the grid must become the main focus of the action plan development in Step 6.

STEP 4 & STEP 5 SCORING

Correctly Written 0 1 2 3 4 5
Applicability and Relevance……..0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Correctly Used 1 2 3 4 5

STEP 6 - DEVELOP an ACTION PLAN

Objective: To develop an action plan based on the highest scoring solution idea in Step 3 to explain and demonstrate its relevance and importance to the UP and the future scene

Relevance 1 2 3 4 5
Effectiveness..... 1 2 3 4 5
Impact 1 2 3 4 5
Humaneness..... 1 2 3 4 5
Development of Action Plan..... 1 2 3 4 5 6 7 8 9 10

Relevance (1-5) – Almost identical to the relevance criterion in Step 3, relevance in Step 6 measures the extent to which the action plan is relevant to the underlying problem. It is determined by comparing the action plan to goals stated in the underlying problem. Evaluators determine the relevance of the action plan on a 1-5 scale. If the relationship is excellent, a score of 5 is awarded. Lower scores are given to solution ideas off target. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Process*

Effectiveness (1-5) – This scale measures the potential effectiveness of the action plan in relation to the goals stated in the underlying problem. In differentiating between relevance and effectiveness, **relevance asks whether the action plan addresses the underlying problem; effectiveness asks whether the action plan successfully solves the underlying problem.** Evaluators rank the effectiveness of the action plan on a 1-5 scale. An effective action plan is one that does much to solve the underlying problem. A low score is given to an action plan that does very little or nothing to achieve the goals stated in the underlying problem. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Content*

Impact (1-5) – This scale measures the positive impact of the action plan on the future scene. The evaluator must assess the impact of the action plan on the future scene. **An effective action plan for an underlying problem that scores well in focus and adequacy in Step 2 usually receives a high score. Lower scores are assigned to action plans not as good in one or both of these areas.** Use the descriptors on the scoresheet to determine the numerical score. *Rating; Content*

Humaneness (1-5) – This scale measures the productive, positive potential of the action plan as opposed to its destructive, negative potential. To score this section, an evaluator anticipates the practical consequences in implementing the action plan. Humaneness of an action plan is scored independently of relevance, effectiveness, and impact. While an action plan may score poorly in the other Step 6 criteria, it can still score well in humaneness. Evaluators score humaneness on a 1-5 scale. A score of 3 on the scale represents a neutral solution idea. Higher numbers are awarded if the solution idea actively seeks to be constructive; lower scores are assigned if it is actively destructive. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Content*

Development of Action Plan (1-10) – The criterion for development of the action plan measures the degree to which a team creates a strategy for addressing the underlying problem. An action plan that scores high in this area would fully describe the action to be taken and outline the steps necessary to complete the plan. The action plan may also explain the challenges that must be overcome to achieve its goal. The idea is to paint a complete picture of the plan. An action plan that simply restates the solution idea from Step 3 would score on the low end of this scale. Evaluators score the action plan on a 1-10 scale. A well-developed action plan is fully explained and elaborated. The action plan describes any obstacles to overcome in implementing the plan and explains *why* and *how* the plan has a positive impact on the future scene, topic, and society. Use the descriptors on the scoresheet to determine the numerical score. *Rating; Content*

STEP 6 EVALUATION ESSENTIALS:

1. An action plan is a *proposal* for solving the underlying problem. The action plan should ***explain in detail*** the *who, what, how, why, where,* and *when* of the solution idea. Developing an action plan involves moving from creative ideas into action; a new idea is incomplete until it is a workable idea. The action plan demonstrates how it addresses the problem area in Step 2 and how it affects the future scene.
2. Teams may add new facets to their action plan at this point, as long as each addition represents a subpart of the Step 3 solution idea. Combining two or more solution ideas is not recommended.
3. If a team combines two or more unrelated solution ideas in their action plan, the evaluator should score only the first solution for relevance, effectiveness, impact, and humaneness.
4. Reserve a score of one (1) on relevance, effectiveness, and impact for action plans completely "off" the underlying problem and the future scene.

OVERALL

Objective: To combine content (research) and process (creative problem solving) to effectively work from a future scene to a focused action plan

Research Applied..... 1 2 3 4 5 6 7 8 9 10
Creative Strength..... 1 2 3 4 5 6 7 8 9 10
Futuristic Thinking..... 1 2 3 4 5 6 7 8 9

Research Applied (1–10) – This scale measures a team's use of research throughout the booklet. Student work in each step of the process is examined for connection to the research available on the topic, as well as knowledge of issues and trends in general. Concepts from the research, terms from the research, and noticeable flexibility are all indications of research applied. Using the scoresheet criteria, evaluators rank the overall research applied on a 1-10 scale. *Rating; Content*

Creative Strength (1–10) – This scale assesses the overall creative productive thinking evidenced in the booklet. Skillful use of the problem solving process is also indicative of creative thinking. Responses showing creative strength are those requiring intellectual energy to make mental leaps beyond obvious or commonplace responses. Evaluators should look in any or all steps for innovative or unconventional thinking and for ideas indicating fresh insights and perceptions. High scores on the creative scales of fluency, flexibility, elaboration, and originality are also signs of creative strength. Using the scoresheet criteria and the overall creativity of the booklet, evaluators assign a score on a 1-10 scale. *Rating; Content*

Futuristic Thinking (1–10) – This scale assesses the ability of the team to:

1. Put themselves in the time frame of the future scene, and
2. Extrapolate relevant trends and technologies from its research as it identifies futuristic challenges and create workable, futuristic solution ideas. Evaluators should note there is a difference between creative, futuristic solutions and those solution ideas that are trivial and/ or "cutesy." In each step, teams should show an understanding of how their work could impact future society.

Using the scoresheet criteria and based on the overall futuristic appeal of the booklet, evaluators make their determination of futuristic thinking on a 1-10 scale. *Rating; Content*

CATEGORY LIST

Use these categories in evaluating fluency in Step 1 challenges and Step 3 solution ideas.

1. Arts & Aesthetics
2. Basic Needs
3. Business & Commerce
4. Communication
5. Defense
6. Economics
7. Education
8. Environment
9. Ethics & Religion
10. Government & Politics
11. Law & Justice
12. Miscellaneous
13. Psychological Health
14. Physical Health
15. Recreation
16. Social Relationships
17. Technology
18. Transportation

GENERAL COMMENTS

A team that attempts work on a particular section of the booklet must receive a minimum score of one (1) point for each criterion in that section. Only sections that contain no student work can be scored zero (0).

Exceptions to the rule are: structure in Step 2, originality in Steps 1 and 3, correctly written in Step 4, and applicability and relevance in Step 4.

Upon completing the evaluation of each step, an evaluator should write specific comments on the scoresheet. Although there is only enough space to make a few brief observations, the comments are extremely important. This is the evaluator's chance to encourage the students and give them pointers to improve their problem solving skills. Initially, students look at the score to determine how they performed on a booklet. Shortly thereafter, however, they read the evaluator comments on the scoresheet for the true determination of their performance. It is the evaluator's insights that make the last impression.

After an evaluator completes an evaluation, writes comments, and assigns scores for each section, the total number of points should be determined and entered in the appropriate space. Double-checking addition is a must!

Ranking

The best way for comparing booklets is with a ranking system. Using such a system, each evaluator scores an equal number of booklets and then ranks each booklet according to the total points each booklet receives. Booklets are ranked from 1 (best) to the number of booklets scored. A rank of 1 goes to the highest scoring booklet, a rank of 2 goes to the second highest scoring booklet, etc. Ranking booklets eliminates scoring differences between tough and lenient evaluators. It also creates a "common language" for comparing booklets from different samples. The rank of each booklet is recorded in the appropriate space on the scoresheet.

The 22 criteria in evaluation identify the key elements in an FPS booklet. The evaluation of these elements helps students improve their problem solving skills. While improvement of student work is the primary intent of the evaluation system, it is also designed so students who do the best work receive the highest ranks. Therefore, when an evaluator finishes a sample of booklets, he or she should review the booklets to ensure teams that did the best work receive the highest ranks.

Competitive Scoring

For FPS competitions, such as an affiliate final/bowl or the International Conference, students do not see the future scene in advance. In these situations, evaluators reward students for responding directly to the future scene. Doing so recognizes teams who use their creativity to respond spontaneously to a situation, thereby furthering FPSP's educational goal of preparing students to respond to real world challenges.

For clarification, consider a team preparing for an affiliate final/bowl or the International Conference. The students spend time researching the topic and developing ideas that might be relevant to the future scene; however, the team does not see the future scene until the two-hour competition begins. The students must analyze its contents and determine what part of their research and information on the topic applies to the future scene and what does not. Unfortunately, teams sometimes rely too much on their preparation and do not use their creativity to respond directly to the future scene. The result may be a booklet that is "flat," does not pertain to the future scene, or that seems prepared in advance (canned). Research skills are important and FPS strives to take students to the next level, asking them to apply their knowledge to a specific, focused situation.

FPSP creates future scenes for the affiliate final/bowl and the International Conference with these thoughts in mind. The future scenes emphasize preparation but also contain elements that emphasize creativity:

1. Future scenes revolve around an imaginary, yet realistic, futuristic scenario. The imagined and futuristic elements of the future scene allow FPSP to use its own creativity in producing the scenarios. The FPS program intends for students to build upon the creative elements of the future scene and showcase their own creativity.
2. Future scenes concentrate on only a portion of the topic. Not all of the student's research and information is applicable to the future scene, and the team members must adjust to utilize appropriate information relevant to their work in the booklet.

An example of this is the topic of drugs used for the affiliate final/bowl one year. Instead of describing traditional addictions and the effects of drugs on society, the future scene detailed a virtual reality program with drug-like effects. Using the virtual reality program as the basis of the future scene required teams to use only their background on the effects of drugs rather than their knowledge of drugs as a whole.

Evaluators should reward students for their creative, spontaneous responses to the future scene. Students should not be rewarded for writing canned booklets – booklets prepared in advance and composed mostly of ideas not specific to the future scene. By evaluating with these thoughts in mind, evaluators enhance the educational experience for all students, who learn it takes more than just good research to score well in competitive circumstances.

INTERNATIONAL CONFERENCE PHILOSOPHY OF EVALUATION

The International Conference (IC) winners are those teams that exhibit the best Future Problem Solving skills in response to the future scene. More specifically, evaluators look for top-quality work in regard to the following three areas:

1. Research,
2. Use of FPS model, and
3. Spontaneous response to specifics of the International Conference future scene.

A TIP ABOUT THE INTERNATIONAL CONFERENCE FUTURE SCENE

Future scenes become more difficult as the FPS season progresses. Early in the year, future scenes are open-ended and allow teams to develop and enhance their skills. In an effort to get students to think and to help evaluators distinguish teams that memorize from teams that think, the International Conference future scene applies only to a portion of the research available. Therefore, in IC competition, it is up to the students to analyze the future scene and determine what portion of their research is relevant and what is not. The best teams then apply relevant research to the specifics of the future scene.

CONCLUSION

Evaluation is a highly rewarding experience. Evaluators expend considerable mental energy during a day of evaluation; however, they are always re-energized and inspired by the ideas of creative students. **Evaluation is the lifeblood of the Future Problem Solving Program, and evaluators should take great pride in knowing their evaluation makes a significant contribution to FPSP and the hundreds of thousands of students who participate.**

2015 Grant Place
Melbourne, FL 32901
Phone: 321.768.0074 Fax: 321.768.0097
E-mail: mail@fpspi.org
www.FPSPI.org