Inter-rater Reliability
Student Teaching Observation Tool
Fall 2019

All candidates in the initial licensure programs at Purdue University Northwest are required to be evaluated by the Student Teaching Observation Tool (STOT). The School of Education and Counseling adopted the STOT in Fall 2018. This instrument was developed by the North Dakota Association of Colleges of Teacher Education which has established measures of reliability and validity. A developmental approach is used through the program to measure impact of candidates in the field. The Modified Student Teaching Observation Tool (M-STOT) is used during the early and mid-field courses. Items were chosen based on what was developmentally possible based on placement in the program. All items (34) are assessed in student teaching.

Training
On October 11, 2019 an inter-rater reliability training was conducted on the Student Teaching Observation Tool (STOT).

During the training, there were thirteen (13) participants consisting of faculty, university supervisors, and field supervisors.

The training consisted of a power point presentation and videos. The participants watched two videos on candidates teaching a lesson plan. Video 1 was 4th grade language arts watching for evidence related to InTASC 5 Application of Content. Video 2 was 7th grade language arts watching for evidence related to InTASC 1 learner development. Following the videos, the participants scored two rubrics. The scores were collected and compared to the initial scorers from the video.

	Rater 1 Average	Rater 2 Average	Agreement	% Absolute	% Adjacent
Video 1	3.00	2.77	7/13	54%	46%
Video 2	3.00	2.88	8/13	62%	38%

Feedback
- There is a lot to assess
- There is a choice of not observed if you don't always see something
- Assess what candidates are prepared to demonstrate
- No time for collaborative conversation

What recommendations do you have?

- Introduce better progression, too many indicators at mid-level

- Sharing completed evaluations mid-way can help identify strengths and areas for growth

Inter-Rater Reliability Exercise
During the training, four volunteers agreed to participate in an inter-rater reliability exercise during the Fall 2019 semester. Volunteers were field supervisors for Elementary EDCI 35500 Teaching and Learning in K-12 Classroom (Early) and Elementary EDCI 36203 K-3 Literacy (Mid-Level). University Supervisors for Elementary Student Teaching (Late); and Secondary Education Student Teaching (Late).

Elementary Early M-STOT
Below 50% Agreement
InTASC 2 Exhibits fairness and belief that all students can learn

Elementary Mid M-STOT
Below 70% Agreement
InTASC 2 Exhibits fairness and belief that all students can learn
InTASC 3 Structures a classroom environment that promotes student engagement
InTASC 3 Clearly communicates expectations for appropriate student behavior
InTASC 3 Responds appropriately to student behavior
InTASC 3 Guides learners in using technologies in appropriate, safe, and effective ways
InTASC 5 Connects core content to relevant, real-life experiences and learning tasks
InTASC 5 Accesses content resources to build global awareness
InTASC 6 Provides students with meaningful feedback to guide next steps in learning
InTASC 6 Uses appropriate data sources to identify student learning needs
InTASC 6 Engages students in self-assessment strategies
InTASC 7 Uses assessment data to inform planning for instruction
InTASC 8 Instructional practices reflect effective communication skills
InTASC 9 Uses feedback to improve teaching effectiveness
InTASC 9 Uses self-reflection to improve teaching effectiveness
InTASC 9 Demonstrates commitment to the profession

Elementary STOT
Below 50% Agreement
InTASC 2 Uses knowledge of students' socioeconomic, cultural and ethnic differences to meet learning needs
InTASC 3 Guides learners in using technologies in appropriate, safe, and effective ways
InTASC 4 Integrates culturally relevant content to build on learners' background knowledge
InTASC 5 Accesses content resources to build global awareness
InTASC 6 Uses appropriate data sources to identify student learning needs
InTASC 6 Engages students in self-assessment strategies
InTASC 10 Collaborates with parents/guardian/advocate to improve student performance

Secondary STOT
Below 50% Agreement
InTASC 3 Clearly communicates expectations for appropriate student behavior
InTASC 6 Engages students in self-assessment strategies
InTASC 8 Differentiates instruction for a variety of learning needs

Below agreement on multiple inter-rater reliability exercises…
- InTASC 6 Engages students in self-assessment strategies (3)
- InTASC 2 Exhibits fairness and belief that all students can learn (2)
- InTASC 3 Clearly communicates expectations for appropriate student behavior (2)
- InTASC 3 Guides learners in using technologies in appropriate, safe, and effective ways (2)
- InTASC 5 Accesses content resources to build global awareness (2)
- InTASC 6 Uses appropriate data sources to identify student learning needs (2)
- InTASC 6 Engages students in self-assessment strategies (2)

InTASC Standards flagged
- InTASC 2 Learning Differences
- InTASC 3 Learning Environments
- InTASC 5 Application of Content
- InTASC 6 Assessment

Revisions to M-STOT for Spring 2020
Data was shared at both the October 16, 2019 Secondary Program Data Dialogue Day and the October 22, 2019 Elementary Program Data Dialogue Day. Both committees felt that there were too many indicators assessed at the mid-level. Early Childhood also had discussion regarding concern about too many indicators on the M-STOT mid-level at their October 3rd Data Dialogue Day. Both the Early and Mid-Level M-STOTs were revised for Spring 2020. The M-STOT went from 30 indicators down to 23 indicators.

Early M-STOT
- Removed… demonstrated commitment to the profession InTASC 9
- Added… uses feedback to improve teaching effectiveness

Mid-Level M-STOT
Removed
- Guides learners in using technologies in appropriate, safe, and effective ways InTASC 3
- Accesses content resources to build global awareness InTASC 5
- Uses appropriate data sources to identify student learning needs InTASC 6
- Engages students in self-assessment strategies InTASC 6
- Uses assessment data to inform planning for instruction InTASC 7
- Adjust instructional plans to meet students' needs InTASC 7
- Collaboratively designs instruction InTASC 7
- Demonstrates commitment to the profession InTASC 9

Added
- Uses technology appropriately to enhance instruction InTASC 8

Positivity: The field supervisors shared something positive they have seen in the field. Good group of candidates in the field this semester. Only three weeks in the field and the candidates are growing in confidence.

Reviewed the revised M-STOT (Early and Mid-Level) Spring 2020 requirements for field.

STOT Scoring:
- Scoring levels were explained with examples
- A score of 3.0 is considered proficient. If you are giving your candidates a score of 3.5 or 4.00, give them explicit feedback on why they received that score.

Fall 2019 Inter-Rater Reliability Data
F19 STOT Inter-rater reliability data was shared and standards that fell below 50% in agreement between the field supervisor and the cooperating teachers were InTASC **Standard 2** exhibits fairness and belief that all students can learn; **Standard 3** learning differences, **Standard 5** accesses content resources to build global awareness and **Standard 6** assessment.

Training
Standards that were focused on at the training were Standards 3 **Learning Differences** and 6 **Assessment** since they had the most indicators falling below 50%.

Inter-Rater Reliability Exercise
- *Video 1:* Standard 3 Learning Differences: Secondary Math classroom and then scored based on the five standard indicators.
- *Video 2:* Standard 6 Assessment: Elementary lesson focusing on English Language point of view and then scored based on the four standard indicators. *(Source: WIDA)*

Turn and Talk Discussion
- Results were shared and discussed
- How can you have difficult conversations with candidates and/or cooperating teachers?
- Have conversations with your cooperating teachers to improve the percentage of agreement on these weaker areas.
- How are cooperating teachers evaluated and concerns handled?
    o Candidates evaluate their cooperating teachers and feedback is reviewed and discussions are held with the schools
- Use technology to avoid confusion; as simple as a docucam
- Candidates struggle with how do you give meaningful feedback

Future Plans:
- Conduct training with cooperating teachers
- Share videos with candidates during appropriate method courses