

## BRIEF REPORT

# Creation and Initial Validation of the Mayo Outpatient Precepting Evaluation Tool

Deirdre Paulson, PhD; Brandon Hidaka, MD, PhD; Terri Nordin, MD

**AUTHOR AFFILIATION:**

Department of Family Medicine, Mayo Clinic Health System, Eau Claire, WI

**CORRESPONDING AUTHOR:**Deirdre Paulson, Department of Family Medicine, Mayo Clinic Health System, Eau Claire, WI,  
[paulson-odonovan.deirdre@mayo.edu](mailto:paulson-odonovan.deirdre@mayo.edu)**HOW TO CITE:** Paulson D, Hidaka B, Nordin T. Creation and Initial Validation of the Mayo Outpatient Precepting Evaluation Tool. *Fam Med.* 2023;55(8):547–552. doi: [10.22454/FamMed.2023.164770](https://doi.org/10.22454/FamMed.2023.164770)**PUBLISHED:** 21 June 2023

© Society of Teachers of Family Medicine

**ABSTRACT**

**Background and Objectives:** Preceptors in family medicine residencies need feedback to improve. When we found no validated, behavior-based tool to assess the outpatient precepting of family medicine residents, we sought to fill this gap by developing and initially validating the Mayo Outpatient Precepting Evaluation Tool (MOPET).

**Methods:** To develop the MOPET, we applied the Stanford Faculty Development Program (SFDP) theoretical framework for education, more recent work on peer review of medical teaching, and expert review of items. The residency behavioral scientist and a volunteer physician independently completed the MOPET while co-observing a precepting physician during continuity clinic sessions (N=20). We assessed the tool's validity via interrater reliability and cross-validation with the SFDP-26.

**Results:** The tool demonstrated high interrater reliability for the following effective teaching behaviors: (a) allowing the resident to present without interrupting, (b) encouraging the formulation of a goal, (c) checking in on the resident's goal, (d) using multimodal teaching aids, (e) asking to discuss the differential diagnosis, (f) asking to discuss alternative management, (g) encouraging the resident to pursue literature and/or other resources, and (h) reinforcing self-directed learning. The MOPET measures strongly correlated with most items from the SFDP-26, indicating good cross-validity.

**Conclusions:** The MOPET is a theoretically sound, behavior-based, reliable, and initially validated tool for peer review of outpatient family medicine resident teaching. This tool can support faculty development in outpatient clinical learning environments.

**INTRODUCTION**

Effective teaching, requiring feedback and deliberate practice, is central to medical residencies.<sup>1</sup> Skeff's Stanford Faculty Development Program (SFDP) is a fundamental medical teaching framework<sup>2</sup> that has demonstrable positive effects on teaching knowledge, skills, and attitudes,<sup>3</sup> as well as student and resident evaluations of clinical teaching.<sup>4,5</sup> No preceptor assessment tool applies well to all clinical settings.<sup>6</sup> Many tools reflect learners' perceptions rather than observation of precepting.<sup>7–9</sup> Peer review is beneficial and less biased because it sidesteps the issues of power differential.<sup>10,11</sup> Peer teaching observation tools have been validated in settings that do not generalize well to outpatient family medicine (ie, inpatient internal medicine residencies).<sup>12–14</sup> The behavior-based precepting tools for outpatient family medicine were validated on medical student supervision<sup>15</sup> and used preceptor self-assessment,<sup>16</sup> not peer assessment. We therefore

sought to create and start validating an evidence-based, peer-observation instrument to assess precepting behaviors in a family medicine residency outpatient clinic.

**METHODS****Setting**

We developed and initially validated the Mayo Outpatient Precepting Evaluation Tool (MOPET) in the outpatient clinic at the Mayo Clinic Family Medicine Residency–Eau Claire program in Wisconsin. This 5–5–5 program had six core physician faculty members (including the program director), one core behavioral scientist faculty member, and four community preceptors. Mayo Clinic's Education Research Committee (#21-092) and Institutional Review Board (#21-009471) deemed the study exempt from approval.

## Tool Development

While using the Patient Centered Observation Form–Clinician Version (PCOF)<sup>17</sup> to observe and provide feedback to residents in the continuity clinic, faculty preceptors inquired about a comparable tool for them. A literature review led to Beckman and colleagues' work.<sup>12,14,18,19</sup> Finding no tool directly applicable, we created the MOPET.

Beckman granted permission to use and modify the various versions of the Mayo Teaching Evaluation Form (MTEF).<sup>12,14,18,19</sup> We selected and revised MTEF items that consistently demonstrated strong validity. We also created new items derived from the SFDP,<sup>2</sup> PCOF,<sup>17</sup> Maastricht Clinical Teaching Questionnaire,<sup>20</sup> and personal experience. Four core faculty members' edits of this amalgam resulted in the studied version of the MOPET (Supplemental Figure 1). All items defined an observable, effective teaching behavior and were made quantitative to increase objectivity and reliability by counting the frequency of each behavior.

## Procedure

Each observation occurred over a half day of outpatient clinic, allowing for observation of multiple residents and presentations. The behavioral scientist and one volunteer faculty physician observed a precepting physician and independently counted the frequency of teaching behaviors. A precepting encounter was defined as a time a resident entered the precepting space to discuss the care of one or multiple patients. An observed behavior could be tallied only once per precepting encounter. At the end of the half day, observers used the MOPET to share feedback with the preceptor. The behavioral scientist also completed the SFDP-26 for each half day.

We estimated needing 20 observations to achieve at least 90% power to detect a correlation of 0.70 with a type-I error rate of 5%. We quantified the interrater reliability using Spearman's rank correlation coefficient. We applied the Mann-Whitney U test to determine whether the behavioral scientist's observations were significantly higher or lower than the physicians'. Using Spearman's correlation, we assessed the cross-validity of the tool between items of the SFDP-26 and the sum of the relevant MOPET items (Supplemental Table 1). We identified relevant MOPET items (or lack thereof) by consensus. We chose nonparametric tests because only two of the MOPET items and none of the SFDP-26 items had evidence of a normal distribution (Shapiro-Wilk Test,  $P > .05$ ). We performed all analyses in BlueSky Statistics (R version 3.6.3).<sup>21</sup>

## RESULTS

Five core faculty members applied the MOPET as an observer alongside the behavioral scientist in 20 precepting sessions of six different preceptors. The most observed behaviors were showing genuine interest, being available and ready to precept, and allowing the resident to present without interruption (Table 1). The rarest behaviors were encouraging the resident to ask questions, acknowledging when the resident taught them something, encouraging the resident to pursue the literature

and/or other resources, and reinforcing self-directed learning.

The high interrater reliability of the MOPET items is presented in Figure 1. However, observers did not agree on when the preceptor was supporting the resident in activities the learner found difficult. We found weak concordance for prioritization of precepting over other tasks and explaining why the resident was correct or incorrect. The behavioral scientist observed more instances of preceptors supporting the resident in activities the learner found difficult and the use of multimodal learning aids, and fewer explanations for why the resident was correct or incorrect (all  $P < .05$ ).

The good cross-validity between the SFDP-26 and relevant MOPET items is presented in Figure 2. We found especially strong correlations between the SFDP-26 and MOPET for items related to resident goal formation, use of multimodal learning aids, evaluation of a learner's ability to analyze and synthesize knowledge, and delivery of feedback. We found no evidence that the MOPET measures correlated with any of the following SFDP-26 items: expressing respect for learners, discouraging external interruptions, explaining relationships in material, or promoting self-directed learning.

## DISCUSSION

Our work began the validation of an objective precepting evaluation tool for outpatient family medicine residencies, filling an important gap in family medicine education. High interrater reliability indicates that the MOPET measures consistently recognizable behaviors. Strong correlations between the MOPET and most SFDP-26 items demonstrate cross-validity. Important to note is that the SFDP-26 is a peer's gestalt of precepting, whereas the MOPET quantifies effective teaching behaviors. Because observers were not formally trained, the interrater reliability was likely lowered; yet our findings suggest that the MOPET requires little training.

Though this study was not powered to observe behavior change among individual preceptors, the behavioral scientist and observed preceptors expressed a strong sentiment of subjective improvement. Additionally, feedback from faculty members during MOPET development and unsolicited praise from observed preceptors regarding the tool's value during the end-of-session feedback all signified face validity.

The lack of agreement between how frequently the behavioral scientist and peer observer saw the preceptor support a resident in activities the learner found difficult was surprising, but could be improved by including more concrete examples of this behavior in future versions. Another limitation of this study was the element of human error; for example, the correlation among observers for the number of residents in each precepting session was not 1.0 (possibly because additional residents not working in clinic came in to discuss a prescription request or patient case). Although measuring behavior frequency increases the reliability of measurement, it may not be an optimal approach for evaluating precepting behavior for constructive feedback; the best teaching involves using the right combination of behaviors for the specific

learner at specific times. Furthermore, evidence-based precepting behaviors are tallied with this tool even when used at inopportune times, which could lead to a less effective educational experience. Lastly, which observable behaviors are most important to target is unclear.

Future research could include (a) re-evaluating after refinement via a modified Delphi method or factor analysis; (b) measuring use of the tool's effect on preceptor behavior, learner experience, and patient outcomes; (c) assessing external validity in other family medicine residency programs, primary care clinics in other specialties, or other learners; and (d) comparing use to other preceptor improvement tools.

In conclusion, we created and initially validated an evidence-based, peer-observation instrument to support faculty development in the quintessential learning environment of family medicine residency: the continuity clinic.

## PRESENTATION

Poster presentation, STFM Annual Conference, Indianapolis, IN, May 2022.

## FUNDING

This study was partially funded by the 2022 Endowment for Education Research Award from the Mayo Clinic College of Medicine and Science.

## ACKNOWLEDGMENTS

The authors want to thank all those who supported and participated in this study, including faculty, preceptors, and residents.

## REFERENCES

- Bowen JL, Irby DM. Assessing quality and costs of education in the ambulatory setting: a review of the literature. *Acad Med.* 2002;77(7):621–680.
- Skeff KM. Enhancing teaching effectiveness and vitality in the ambulatory setting. *J Gen Intern Med.* 1988;3(2):26–33.
- Skeff KM, Stratos GA, Berman J, Bergen MR. Improving clinical teaching: evaluation of a national dissemination program. *Arch Intern Med.* 1992;152(6):161–161.
- Litzelman DK, Stratos GA, Marriott DJ, Skeff KM. Factorial validation of a widely disseminated educational framework for evaluating clinical teachers. *Acad Med.* 1998;73(6):688–695.
- Litzelman DK, Westmoreland GR, Skeff KM, Stratos GA. Factorial validation of an educational framework using residents' evaluations of clinician-educators. *Acad Med.* 1999;74(10):25–27.
- Fluit CR, Bolhuis S, Grol R, Laan R, Wensing M. Assessing the quality of clinical teachers: a systematic review of content and quality of questionnaires for assessing clinical teachers. *J Gen Intern Med.* 2010;25(12):337–338.
- Castiglioni A, Shewchuk RM, Willett LL, Heudebert GR, Centor RM. A pilot study using nominal group technique to assess residents' perceptions of successful attending rounds. *J Gen Intern Med.* 2008;23(7):60–61.
- Huff NG, Roy B, Estrada CA. Teaching behaviors that define highest rated attending physicians: a study of the resident perspective. *Med Teach.* 2014;36(11):991–996.
- Smith CA, Varkey AB, Evans AT, Reilly BM. Evaluating the performance of inpatient attending physicians: a new instrument for today's teaching hospitals. *J Gen Intern Med.* 2004;19(7):766–771.
- Hinrichs L, Judd D, Hernandez M, Rapport M. Peer review of teaching to promote a culture of excellence: A scoping review. *J Phys Ther Educ.* 2022;36(4):293–302.
- Olivet DM, Willey JM, Bird JB, Rabin JM, Pearlman RE, Brenner J. Third year medical students impersonalize and hedge when providing negative upward feedback to clinical faculty. *Med Teach.* 2021;43(6):700–708.
- Beckman TJ, Lee MC, Rohren CH, Pankratz VS. Evaluating an instrument for the peer review of inpatient teaching. *Med Teach.* 2003;25(2):131–135.
- Mookherjee S, Monash B, Wentworth KL, Sharpe BA. Faculty development for hospitalists: structured peer observation of teaching. *J Hosp Med.* 2014;9(4):244–250.
- Beckman TJ, Lee MC, Mandrekar JN. A comparison of clinical teaching evaluations by resident and peer physicians. *Med Teach.* 2004;26(4):321–325.
- Huang WY, Dains JE, Monteiro FM, Rogers JC. Observations on the teaching and learning occurring in offices of community-based family and community medicine clerkship preceptors. *Fam Med.* 2004;36(2):131–136.
- Brink D, Power D, Leppink E. Results of a preceptor improvement project. *Fam Med.* 2020;52(9):647–652.
- Mauksch L. Patient centered observation form-clinician version. *University of Washington Department of Family Medicine.* 2016. <https://depts.washington.edu/fammed/pcof>.
- Beckman TJ, Mandrekar JN. The interpersonal, cognitive and efficiency domains of clinical teaching: construct validity of a multi-dimensional scale. *Med Educ.* 2005;39(12):229–229.
- Beckman TJ, Cook DA, Mandrekar JN. Factor instability of clinical teaching assessment scores among general internists and cardiologists. *Med Educ.* 2006;40(12):209–210.
- Stalmeijer RE, Dolmans DH, Wolfhagen IH, Muijtjens AM, Scherpbier AJ. The Maastricht clinical teaching questionnaire (MCTQ) as a valid and reliable instrument for the evaluation of clinical teachers. *Acad Med.* 2010;85(11):732–733.
- R Foundation. The R project for statistical computing. 2020. <https://www.R-project.org>.

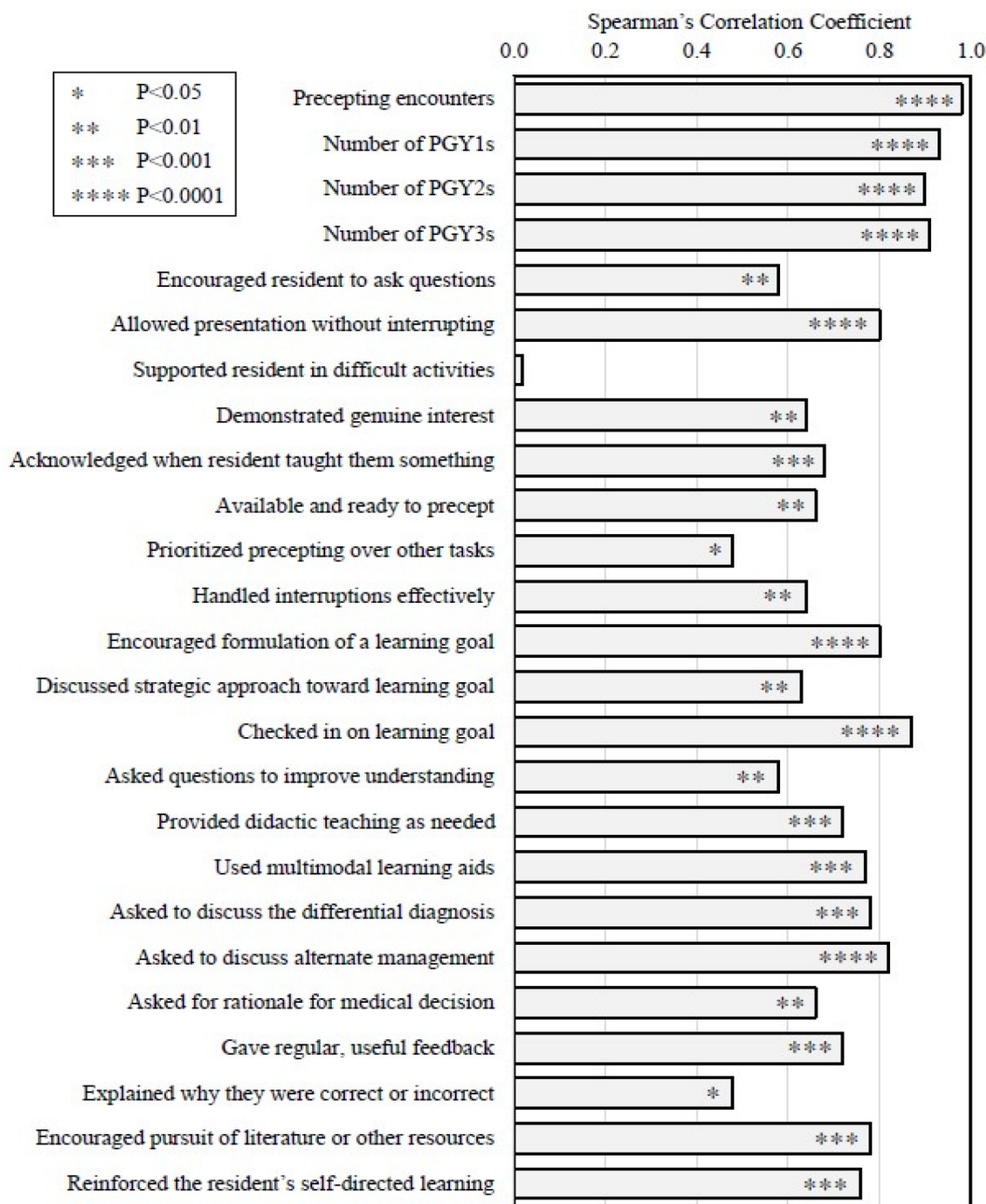
**TABLE 1.** Distribution of MOPET Observations (N=20)

Preceptor behaviors	Median	IQR	Range
Number of precepting encounters during session	13	9–16	2–22
Number of PGY1s	1	0–1	0–2
Number of PGY2s	1	1–2	0–3
Number of PGY3s	1	1–2	0–3
Encouraged the resident to ask questions	0	0–1	0–4
Allowed the resident to present without interrupting	6	4–10	1–22
Supported the resident in activities they found difficult	2	1–4	0–11
Demonstrated genuine interest via body language	10	8–14	2–22
Acknowledged when the resident taught them something	0	0–0	0–2
Available and ready to precept	8	7–12	1–17
Prioritized precepting over other tasks	5	3–7	1–17
Handled interruptions effectively	1	0–2	1–5
Encouraged formulation of a learning goal	1	0–3	0–4
Collaboratively discussed how to achieve the learning goal	0.5	0–1	0–3
Checked in on the resident's achievement of the learning goal	1	0–1	0–3
Asked the resident questions aimed at increasing their understanding	4	3–5	0–12
Provided didactic teaching as needed	4	2–5	0–14
Used multimodal learning aids (eg, whiteboard, video, book)	2	0–3	0–6
Asked the resident to discuss the differential diagnosis	2	1–3	0–7
Asked the resident to discuss alternate management	3	1–4	0–7
Asked the resident to provide a rationale for their medical decision	1.5	1–3	0–6
Gave regular, useful feedback on the resident's performance	2	1–3	0–7
Explained to the resident why they were correct or incorrect	2	0–2	0–5
Encouraged the resident to pursue the literature or other resources	0	0–1	0–3
Reinforced the resident's self-directed learning	0	0–1	0–2

Note: The table represents the number of times a behavior was observed while a preceptor was supervising residents during a half-day clinic session. A behavior could be tallied once per precepting encounter, which was when an individual resident came into the precepting space to discuss one or more patients.

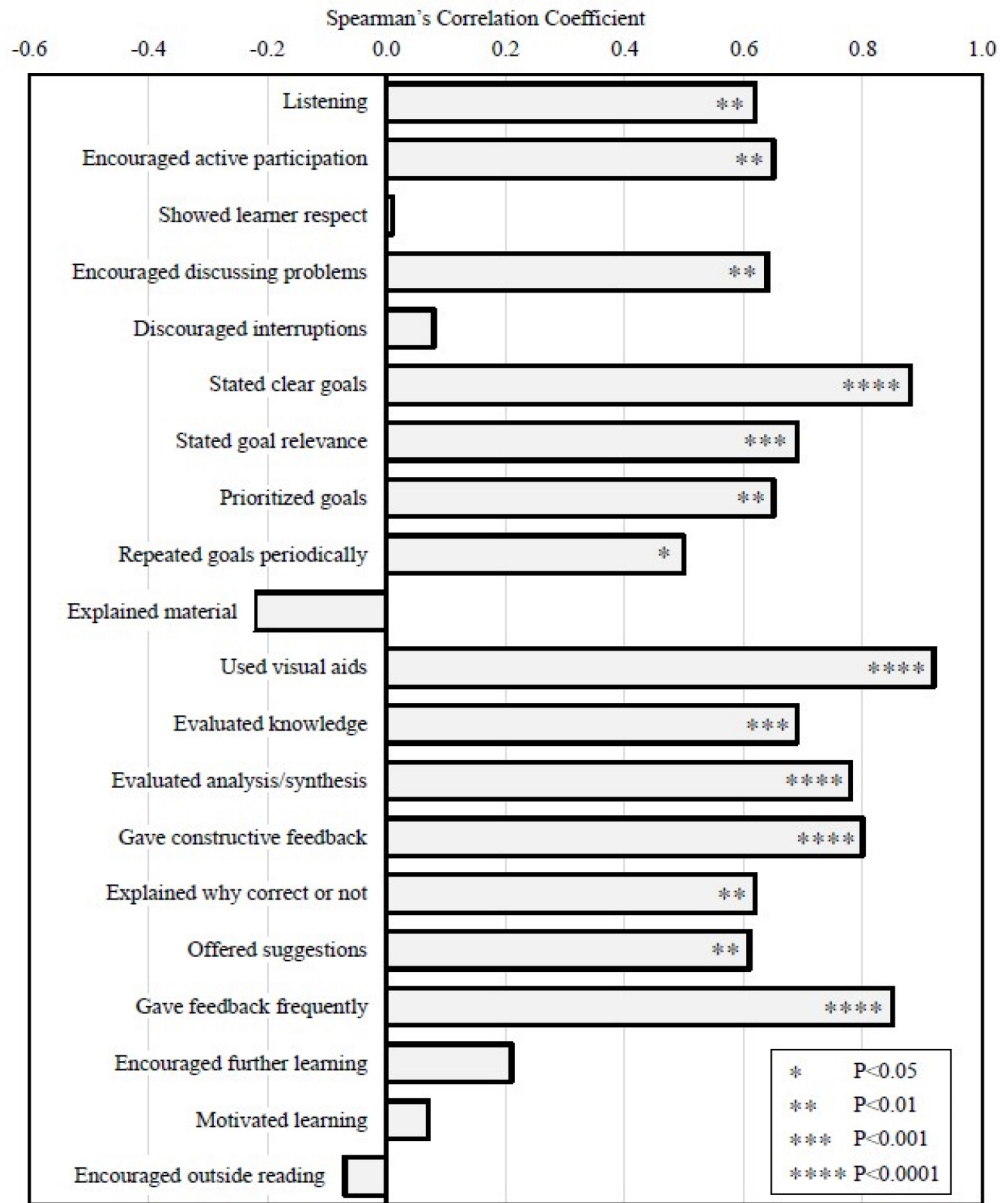
Abbreviations: MOPET, Mayo Outpatient Precepting Evaluation Tool; IQR, interquartile range; PGY, postgraduate year.

**FIGURE 1.** Interrater Reliability of Mayo Outpatient Preceptor Evaluation Tool



Each bar represents the strength of the correlation between the number of times a behavior was observed by the behavioral scientist and a different faculty member while a preceptor was supervising residents for a half-day session in the residency's continuity clinic.

**FIGURE 2.** Correlation Between SFDP-26 and MOPET Measures



This graph illustrates the cross-validity of the Mayo Outpatient Preceptor Evaluation Tool (MOPET). Each bar represents the strength of the correlation between Stanford Faculty Development Program Questionnaire (SFDP-26) measures of effective teaching behaviors and the sum of the relevant MOPET measures.