Teaching continues to be recognized by the Accreditation Council for Graduate Medical Education (ACGME) and Liaison Committee on Medical Education (LCME) as an essential skill for future physicians. Residents as teachers (RAT) and medical students as teachers (MSAT) programs have been described and evaluated in the literature. MSAT programs have demonstrated positive outcomes in learner assessment, knowledge attainment, and development of future physician educators. RAT programs have also demonstrated benefits to both the resident teacher and student learner in attitude, knowledge, communication, and professional development.

As separate entities, both RAT and MSAT programs have been recognized by the Accreditation Council for Graduate Medical Education (ACGME) and Liaison Committee on Medical Education (LCME) as an essential skill for future physicians. Residents as teachers (RAT) and medical students as teachers (MSAT) programs are attracting more attention in graduate and undergraduate medical education as an approach to improving teaching skills, with RAT programs established at 55% of ACGME-accredited residencies across multiple specialties.
supported by statements from the ACGME Sponsoring Institution 2025, whose future vision for residency programs included increased alignment of UME/GME educational methods.18

In 2010, Northwestern University Feinberg School of Medicine (NUF-SM), in collaboration with McGaw Medical Center of Northwestern University (McGaw), aligned teacher-training programs to generate a program across the GME and UME continuum. The goal of the aligned program was for resident and fourth-year medical student (M4) trainees to demonstrate effective techniques for teaching in small groups and to give effective feedback as a clinical teacher. The alignment of teaching programs satisfied a resident desire to teach, decreased faculty time constraints, created consistency in training across the continuum, and provided experiential learning in teaching and feedback to residents and medical students.

We hypothesize that the aligned teacher-training program across the medical education continuum provided educational benefits for both residents and medical students through high-quality feedback and satisfying experiential learning opportunities. Given the potential benefits of the aligned teacher-training program across the medical education continuum, the purpose of this report was to provide a curricular overview and evaluate the early outcomes of the program through analysis of narrative feedback quality and participant satisfaction.

**Methods**

**Program Description**

Since 2005, all M4s at NUSFM have been required to teach M1/M2s in a course called the Teaching Selective. GME-UME program alignment occurred in 2010, with creation of an elective resident teaching program that invited residents to provide feedback to M4s on their teaching skills. Both programs had the same objectives, including giving effective feedback as a clinical teacher; and assessing and improving personal performance through the creation of an individual improvement plan. We chose these two objectives since the relevant accrediting bodies for each group (LCME and ACGME) both have standards in these domains. In designing the RAT program, aligning resources and teaching the same standards for both groups was the key to achieving high-quality feedback across the continuum. Both residents and M4s receive separate, yet the identical training (in-person and video modules) in providing effective feedback and teaching, followed by experiential learning opportunities to develop skills through practice and reflection.

Both residents and M4s participate in multiple experiential learning opportunities for teaching and feedback. In the Teaching Selective, M4s teach and provide feedback on history, communication, and physical exam skills to M1/M2 small groups. Throughout the year, M4s teach an average of 12 hours. The M4s are videotaped while teaching, and all M4s review their teaching videos for self-assessment. Each M4 teaching video is also reviewed by a resident who provides in-person verbal and written feedback to the M4. Residents provide feedback to four to five M4s throughout the academic year. Both residents and M4s complete posttraining satisfaction surveys.

Medical student participation in the aligned program is mandatory and resident participation is voluntary. Residents receive a letter of completion signed by the course director, designated institutional official, and the senior associate dean for medical education to include in their record, but receive no additional incentives for their participation.

**Program Evaluation**

Evaluation of the aligned teacher-training program occurred through analysis of written feedback quality and program satisfaction, providing an evaluation of the program goal to give feedback as a clinical teacher. Sources of data for the program evaluation included written feedback from residents to M4s on their teaching skills, and posttraining satisfaction survey data from residents and M4s. Two raters analyzed written feedback provided during the teaching program using preestablished codes (described below) to assess the quality of feedback. Evaluation of participant reaction to the program occurred through analysis of resident and M4 satisfaction survey data.

Residents provided in-person and written feedback to M4s on their teaching skills over 3 academic years from 2013 to 2016. Only written feedback was captured and analyzed for the purpose of this initial program evaluation. In this time period, a total of 79 residents and 445 M4s participated in the program, including residents from multiple residency programs with varying years of participation (Table 1). Prior to analysis in 2017, written feedback was deidentified, and codes were utilized to distinguish academic year, residency program, and number of years participating in the teacher-training program.

Quality ratings of written feedback were devised from a previously established operational definition of feedback in clinical education developed by van de Ridder et al, that describes quality feedback as “specific information about the comparison between a trainee’s observed performance and a standard, given with the intent to improve.”19 Using the three components of quality feedback defined by van de Ridder, a single numeric quality coding scheme was developed by the authors (Figure 1). Those three components of quality feedback include: specific information (ie, specific observation), comparison between observed performance and a standard (ie, performance gap) and given with an intent to improve (ie, actionable item for improvement). This is in alignment with family medicine program director and resident identification of quality feedback as containing actionable, specific information.20
Feedback containing any of the qualities defined by van de Ridder was deemed as strong with further coding quantification depending on the content. Specifically, strong feedback was coded on a scale from 1 to 3, with 3 being the highest quality rating. With the goal in feedback to ultimately improve or modify the performance of the learner, we gave the highest quality code (code=3) to feedback stating an actionable item for improvement. Identifying a performance gap (code=2) captured modifying feedback without a specific item for improvement. Finally, making a specific observation (code=1) captured reinforcing feedback and specific performance traits without identifying a performance gap or an area for improvement. Feedback containing none of these qualities was deemed weak in the quality coding scheme. Specifically, weak feedback (code=0) included nonspecific, blank, or irrelevant feedback. Specific examples of each type of feedback and the accompanying codes are provided in Table 2.

Two independent coders (R.B. and T.U.) pilot coded eighteen feedback narratives to discover discrepancies among coding definitions and establish concordance between the two independent coders. Coders and an expert panel (R.B., T.U., K.W., C.P., E.R.) discussed ratings to reach a consensus on code definitions before completing full data analysis. After independent qualitative coding of all 445 resident narratives, discrepancies between the two independent coders (R.B. and T.U.) were discussed with a panel of experts (R.B., T.U., C.P., E.R.) to reach consensus on the final code.

Quantitative analysis of participant reaction to the program was based on resident and M4 posttraining satisfaction survey data from 2011 to 2016. Upon completion of the teacher-training program, both groups were emailed a link to a final course satisfaction survey. The survey asked residents and M4s to rank on a scale of 1 to 6 (1=extremely dissatisfied, 6=extremely satisfied) their “overall satisfaction participating in the Teaching Selective program,” and “perceived importance of this program to your experience/education as a learner.”

We analyzed data using SPSS version 24 for Windows. All tests used an α level of 0.05. We calculated descriptive statistics including means and standard deviations, and we calculated grand means by averaging overall quality ratings across residents in each residency program. To determine whether or not quality ratings varied by residency program, we used a mixed-model analysis to account for repeated observations by the same resident. Northwestern University Feinberg School of Medicine Institutional Review Board determined this program evaluation to be exempt from further review (ID STU00203143).

### Table 1: Resident Participation by Residency Program From 2013 to 2016

<table>
<thead>
<tr>
<th>Residency Program</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anesthesia</td>
<td>10</td>
<td>12.7</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>9</td>
<td>11.4</td>
</tr>
<tr>
<td>Family medicine</td>
<td>5</td>
<td>6.3</td>
</tr>
<tr>
<td>Internal medicine</td>
<td>19</td>
<td>24.1</td>
</tr>
<tr>
<td>Neurology</td>
<td>8</td>
<td>10.1</td>
</tr>
<tr>
<td>Orthopedic surgery</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>26</td>
<td>32.9</td>
</tr>
<tr>
<td>Physical medicine and rehabilitation</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td><strong>Total resident participants</strong></td>
<td>79</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year(s) of Participation</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>38</td>
<td>48.1</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>26.6</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>22.8</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2.5</td>
</tr>
</tbody>
</table>

We analyzed data using SPSS version 24 for Windows. All tests used an α level of 0.05. We calculated descriptive statistics including means and standard deviations, and we calculated grand means by averaging overall quality ratings across residents in each residency program.

Quantitative analysis of participant reaction to the program was based on resident and M4 posttraining satisfaction survey data from 2011 to 2016. Upon completion of the teacher-training program, both groups were emailed a link to a final course satisfaction survey. The survey asked residents and M4s to rank on a scale of 1 to 6 (1=extremely dissatisfied, 6=extremely satisfied) their “overall satisfaction participating in the Teaching Selective program,” and “perceived importance of this program to your experience/education as a learner.”

We analyzed data using SPSS version 24 for Windows. All tests used an α level of 0.05. We calculated descriptive statistics including means and standard deviations, and we calculated grand means by averaging overall quality ratings across residents in each residency program.

We analyzed data using SPSS version 24 for Windows. All tests used an α level of 0.05. We calculated descriptive statistics including means and standard deviations, and we calculated grand means by averaging overall quality ratings across residents in each residency program.

We analyzed data using SPSS version 24 for Windows. All tests used an α level of 0.05. We calculated descriptive statistics including means and standard deviations, and we calculated grand means by averaging overall quality ratings across residents in each residency program.
### Table 2: Examples of Coded Feedback Narratives

<table>
<thead>
<tr>
<th>Quality Code</th>
<th>Feedback Narratives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak feedback Code=0</td>
<td>No comment.</td>
</tr>
<tr>
<td>Specific observation Code=1</td>
<td>I think [student] gave really meaningful feedback to the students about their encounters and really made an effort to include specific suggestions on how they could have changed their counseling. It’s great she also reinforced the things they did well.</td>
</tr>
<tr>
<td>Identifies performance gap Code=2</td>
<td>Had good, direct feedback for students (real, medical suggestions for ROS) as well as good practical experience related to clinical rotations. Soft spoken and often drowned out by standardized patient.</td>
</tr>
<tr>
<td>Actionable item for improvement Code=3</td>
<td>Good introduction: made student feel welcome, comfortable and made it a less formal setting. Good how you started by saying “what do you think well or not well?” This is a great way to hear their perspective on their experience and helps you target where you can focus your teaching... When the student did something well, you often said “you did a great job there.” Try to help them understand WHY they did it well by further explaining... At the end, I recommend asking the student to give a summary of what they can do next time to improve their H&amp;P skills; this helps solidify the teaching points made.</td>
</tr>
</tbody>
</table>

### Results

**Resident to M4 Narrative Feedback**

Seventy-nine total residents participated in the aligned program from 2013 to 2016, developing 445 feedback narratives, of which all were included for a 100% response rate. The overall mean quality rating of resident feedback was 2.72 with a standard deviation of 0.71 (Table 3). Among the narratives 1.8% (8 of 445 total) were coded as weak feedback, 10.1% (45/445) were coded as providing a specific observation, and 2.2% (10/445) were coded as identifying a performance gap. The remaining 85.8% (382/445) received the highest quality code of providing an actionable item for improvement.

We analyzed quality ratings of written feedback by academic year, years of resident participation in the aligned program, and residency program (Table 3 and Figure 2). The quality of feedback provided by residents to M4s remained consistently high over time. The highest quality code of actionable item for improvement (code 3) was assigned to 84% (126/150) of narratives in 2013-14, 85.7% (138/161) in 2014-15, and 88.1% (118/134) in 2015-16. The frequency of weak feedback narratives also remained consistently low over the years, ranging from 3.3% (5/150) in 2013-14, to 1.2% (2/161) in 2014-15, and to 0.7% (1/118) in 2015-16.

Differences by academic year were not statistically significant ($P=0.87$). When analyzed by years of participation, feedback with an actionable item for improvement occurred in 84% of narratives (131/156) with 1 year of participation, 89% (131/147) with 2 years of participation, 83.3% (105/126) with 3 years of participation, and 93.8% (15/16) with 4 years of participation. Differences by years of participation were not statistically significant ($P=0.74$). When divided by subspecialty, the mean quality coding for each residency program was greater than 2.5 (Figure 2).

**Program Satisfaction**

Postparticipation resident and medical student satisfaction surveys evaluated the program as a whole. Resident mean satisfaction with participation in the aligned teacher-training program ranged from 4.7 to 5.1 on a Likert scale of 1 to 6 (1=extremely dissatisfied, 6=extremely satisfied) from 2011 to 2016. Medical student mean satisfaction with participation in the program ranged from 4.7 to 5.2 on the same scale. Resident perceived importance of the aligned program to their experience or education as a learner ranged from a mean of 4.1 to 5.6 on a scale of 1 to 6 (1=extremely unimportant, 6=extremely important) from 2011 to 2016. Medical student mean perceived importance of the program ranged from 4.4 to 4.7.

### Discussion

Evaluation of the aligned teacher-training program found evidence of high-quality written feedback from resident trainees in all programs and satisfaction by all participants, while providing experiential learning opportunities in teacher-training across the UME-GME continuum. High-quality written feedback was provided across all of the residency specialties. M4 trainees received high-quality feedback and all trainees received experiential learning through recurrent practice in teaching and feedback. Our study demonstrated the aligned RAT and MSAT program achieved high-quality written feedback, which is one component of effective teaching. However, future research is needed to evaluate the magnitude of effect of high-quality feedback on learners. The curriculum also fostered new near-peer opportunities to socialize with residents and medical students outside of their own training programs through meetings established within the aligned program, although this was not directly studied in this initial program evaluation.

Feedback is essential in clinical learning and has been shown to change a physician’s clinical performance. For that reason, it is vital...
that our trainees receive high-quality feedback in the aligned teacher-training program as it can impact their performance as a learner and future clinician. The quality of resident to M4 feedback remained consistently high through each academic year, years of participation, and residency program, thus ensuring M4s received this essential feedback regardless of the resident with whom they were paired. For the resident participants, this continued to satisfy a desire to teach while reinforcing the practice of providing high-quality feedback without the need for significant faculty intervention. Lastly, as an organization, the creation of the program aligned UME and GME educational resources while fulfilling shared ACGME and LCME standards. This high-quality feedback was likely multifactorial and cannot be fully credited to the aligned program. Outside influences included an institutional focus on giving feedback and teaching opportunities outside of the aligned program.

Kirkpatrick's model of program evaluation encourages evaluation beyond learner satisfaction. This program evaluation incorporated multiple levels of Kirkpatrick's model, including reaction via satisfaction surveys and behavior via feedback quality analysis. A high level of resident satisfaction was essential to the sustainability of the aligned teacher-training program given that resident participation was optional. This sustainability was evident as a secondary finding in the multiple years of participation and wide range of resident participants from various residency programs (Table 1).

We utilized several strategies to increase validity and reliability in this program evaluation. We addressed the validity of feedback quality by utilizing an operational definition of feedback from the literature, independent rater coding, and use of an expert panel to establish coding consensus. Although a singular definition of feedback is not universally accepted in the education literature, feedback as defined by van de Ridder et al has been frequently referenced in recent literature. In analysis of satisfaction surveys and feedback quality,
attempts were made to decrease construct underrepresentation by including all resident data from several years. By increasing the sample size, we achieved a more reliable reflection of the population mean quality feedback and satisfaction ratings.

Despite efforts to minimize threats to validity and reliability, a limitation to this study is self-selection bias, as resident participation in the program was optional. Secondly, a ceiling effect limits the ability to distinguish the degree of feedback quality with the majority of learners receiving the highest quality code. Further, we suspect the feedback quality ratings may actually underrepresent the quality of feedback through construct-irrelevant variance, as written feedback did not reflect all narrative (verbal and written) feedback given during the feedback session. For example, in-person verbal feedback may have a higher level of detail and feedback quality than that captured in the coded written feedback as “no comment.”

Conclusion
The aligned teacher-training program demonstrated evidence of success across the educational continuum in providing high-quality written feedback, providing experimental learning opportunities, and contributing to accreditation standards for both learner groups. Medical students received high-quality feedback from residents, and trainees reported satisfaction with the aligned program. Further investigation is needed to demonstrate the impact of high-quality feedback on learners and the long-term impact of the aligned teacher-training program on future faculty development and professional identity formation.

ACKNOWLEDGMENTS: The authors thank senior leadership from Northwestern University Feinberg School of Medicine (Marianne Green, MD, Diane B. Wayne, MD, John X Thomas, PhD, Raymond Curry, MD), McGaw Medical Center of Northwestern University Designated Institutional Official Joshua Goldstein, MD, and Program Directors Loussanne Carabini, MD, John Bailitz, MD, Mike Gisondi, MD, Deborah Edberg, MD, Deborah Smith Clements, MD, Aashaish Dwivania, MD, Danny Bega, MD, Monica Rho, MD, Matthew Beal, MD, John Sullivan, MD, Richard Dsida, MD, Susan Gerber, MD, Magdy Milad, MD, Ronell Malkani, MD, Tanya Simuni, MD and Sharon Unti, MD, for supporting the alignment of the resident and student teaching programs.

CORRESPONDING AUTHOR: Address correspondence to Dr Robyn Bockrath, Department of Pediatrics and Medical Education, Northwestern University Feinberg School of Medicine, 225 E. Chicago Ave, Box 152, Chicago, IL 60611. 312-227-6145. rbockrath@luriechildrens.org.

FUNDING/SUPPORT: Funding for the Teaching Selective course was internally provided by the Augusta Webster Office of Medical Education at Northwestern University Feinberg School of Medicine. The authors report no external funding for this study.

PREVIOUS PRESENTATION: Data in this manuscript was presented by Robyn Bockrath, MD, MEd, at the 2017 AAMC Learn Serve Lead conference in Boston, Massachusetts, Educating Physicians and Scientists session titled “Highlights in Medical Education Innovations: Education for Learners.”

References

FAMILY MEDICINE VOL. 52, NO. 5 • MAY 2020 351