

# Pengembangan Rubrik untuk Menilai Proposal Penelitian Peserta Pelatihan Pembentukan Jabatan Fungsional Peneliti

Ijang Permana Sidik<sup>1</sup>, Murti Kusuma Wirasti<sup>1</sup>, Eveline Siregar<sup>1</sup>

<sup>1</sup>Teknologi Pendidikan-Universitas Negeri Jakarta

## INFO ARTIKEL

### Riwayat Artikel:

Diterima: 08-03-2021  
Disetujui: 15-03-2021

### Kata kunci:

*rubric development;*  
*research proposal;*  
*functional;*  
*pengembangan rubrik;*  
*proposal penelitian;*  
*jabatan fungsional*

## ABSTRAK

**Abstract:** The difference of standards used by raters, have become problem during assessing training participant's research proposals. The assessment result become subjective and resulted score with high level of disparity. The purpose of this study is to produce a rubric that can produce objective and consistence scores. Rubric developed by combining *asmawi* development model with *mckeown* and *biss* development model. The rubric stated valid after tested and receiving feedback from evaluation expert, material expert, instructors, training participants, and training committee. The score of rubric's coefficient correlation is 0.78. With this score, the rubric can be stated as a reliable rubric.

**Abstrak:** Perbedaan standar yang dimiliki oleh masing-masing penguji ketika menilai proposal penelitian telah menghasilkan nilai dengan tingkat disparitas yang cukup tinggi. Akibatnya nilai yang dihasilkan menjadi subjektif dan tidak konsisten. Penelitian ini bertujuan untuk menghasilkan sebuah rubrik penilaian yang dapat menghasilkan nilai yang objektif dan konsisten. Penelitian ini termasuk dalam jenis penelitian dan pengembangan. Rubrik dibuat dengan menggabungkan dua model pengembangan yaitu model *zainul* dan model *mckeown and biss*. Rubrik dinyatakan valid setelah diujicoba serta mendapat masukan dari ahli evaluasi, ahli materi, pengajar, peserta, dan penyelenggara pelatihan. Rubrik memiliki nilai koefisien korelasi sebesar 0,78. Dengan skor tersebut maka rubrik dinyatakan reliabel.

### Alamat Korespondensi:

Ijang Permana Sidik  
Teknologi Pendidikan  
Universitas Negeri Jakarta  
Kampus A UNJ, Gedung Bung Hatta, Jl. R.Mangun Muka, Rawamangun, Jakarta  
E-mail: IjangPermanaSidik\_9901818010@mhs.unj.ac.id

Penilaian (*assessment*) adalah bagian yang tidak terpisahkan dari proses pembelajaran. Penilaian dilakukan untuk melihat tingkat keberhasilan (efektivitas) dari suatu proses pembelajaran. Terdapat dua jenis penilaian yaitu penilaian dengan tes dan non tes (Arifin, 2011a). Penilaian baik dengan tes maupun non tes memiliki sifat dan karakteristik yang berbeda. Penilaian dengan tes umumnya dilakukan untuk mengukur hasil belajar pada aspek kognitif sedangkan penilaian non tes umumnya digunakan untuk mengukur hasil belajar yang lebih kompleks. Penilaian dengan bentuk tes pada umumnya dilakukan dengan cara memberikan ujian tertulis menggunakan instrumen baku. Oleh karenanya, penilaian tersebut sering disebut juga sebagai *pencil and paper test* (Arifin, 2011a). Instrumen baku dikembangkan dengan sedemikian rupa dan melewati tahapan lolos uji validitas dan uji reliabilitas. Validitas berkaitan dengan ketepatan alat untuk melakukan penilaian. Reliabilitas berkaitan dengan kemampuan alat dalam memberikan hasil yang konsisten. Penilaian non tes umumnya digunakan untuk menilai kinerja (*performance*) atau produk yang di hasilkan oleh peserta didik. Contoh penilaian non tes adalah untuk menilai kemampuan bernyanyi (Rahmawan, Sumaryanto, & Suproyadi, 2016), membuat karya tulis (Rakedzon & Tsabari, 2017), mengapresiasi karya seni (Tam, 2017), hingga ujian praktik bagi calon dokter (O'Donnell, Oakley, Haney, O'Neill, & Taylor, 2011).

Pelatihan Pembentukan Jabatan Fungsional Peneliti (PPJFP) merupakan sebuah program pelatihan yang diselenggarakan oleh Lembaga Ilmu Pengetahuan Indonesia (LIPI). Program ini diperuntukkan bagi peneliti ASN dari lembaga penelitian (litbang) baik dari kementerian, lembaga pemerintah non kementerian (lpnk), maupun daerah. Program ini bertujuan untuk membekali para peserta dengan berbagai kompetensi yang dibutuhkan untuk duduk pada jabatan fungsional peneliti. Salah satu mata diklat pada program PPJFP adalah mata diklat Penulisan Proposal Penelitian. Mata diklat ini merupakan mata diklat dengan jumlah jam pelajaran (jp) terbanyak dibanding mata diklat lainnya yaitu 14 jp. Banyaknya jp pada mata diklat ini dikarenakan adanya tugas mandiri serta bimbingan (mentoring) menulis proposal penelitian. Di akhir kegiatan, peserta akan diuji untuk mempresentasikan proposal penelitian yang mereka tulis.

Proposal yang dipresentasikan akan dinilai oleh enam orang penguji. Tiga puluh orang peserta diklat akan dibagi menjadi tiga kelas dimana masing-masing kelas terdiri dari dua orang penguji. Pembagian kelas seperti ini bertujuan agar proses penilaian dapat dilakukan dengan cepat, tepat, dan efisien. Namun, dalam praktiknya proses penilaian seperti ini ternyata memiliki permasalahan tersendiri utamanya pada rendahnya tingkat reliabilitas skor penilaian. Rendahnya tingkat reliabilitas skor penilaian disebabkan karena tidak adanya standar yang dapat dijadikan panduan dalam melakukan penilaian. Tiap penilai memiliki standar yang berbeda sehingga nilai yang dihasilkan menjadi subjektif. Dari beberapa kali pelatihan, dapat dilihat ada beberapa penilai yang cenderung memberikan nilai rendah dan ada beberapa penilai yang cenderung memberikan nilai tinggi.

Permasalahan ini ternyata tidak hanya terjadi di LIPI saja. Hal serupa juga dapat ditemukan dalam penelitian O'Donnell, Oakley, Haney, O'Neill, & Taylor (2011). Penelitian tersebut mengungkap adanya permasalahan dalam proses penilaian ujian praktik bagi mahasiswa kedokteran gigi di Amerika Serikat. Dalam penelitian tersebut, diungkapkan adanya ketidakkonsistenan dalam melakukan penilaian. Selain contoh di atas, ternyata permasalahan serupa juga terjadi pada penelitian Furze, Gale, Black, Cochran, & Jensen (2015) dan penelitian Rakedzon & Tsabari (2017). Kedua penelitian tersebut mengangkat permasalahan yang sama persis. Ketidakkonsistenan terjadi karena setiap penilai memiliki standar penilaian yang berbeda.

Standar penilaian merupakan sesuatu yang penting untuk dirumuskan. Perumusan standar akan sangat berguna baik ketika mengembangkan maupun ketika melakukan sebuah proses penilaian. Apabila permasalahan ini tidak segera diselesaikan maka akan berimbas pada permasalahan keadilan, validitas, dan reliabilitas. Padahal ketiganya merupakan indikator dari baik dan tidaknya kualitas suatu proses penilaian (Mislevy & Knowles, 2002). Masalah yang pertama berkaitan dengan ketidakadilan. Sebagian peserta akan diuntungkan apabila secara kebetulan dinilai oleh penilai yang cenderung mudah memberikan nilai besar. Sebaliknya, peserta yang lain akan dirugikan apabila secara kebetulan dinilai oleh penilai yang cenderung memberikan nilai yang kecil.

Permasalahan selanjutnya berkaitan dengan validitas nilai. Karena setiap penilai menggunakan standar yang berbeda, maka makna angka pada nilai yang diberikan pun akan berbeda. Nilai 80 yang diberikan oleh seorang penilai belum tentu bermakna sama dengan nilai 80 yang diberikan oleh penilai lainnya. Permasalahan yang terakhir berkaitan dengan reliabilitas. Reliabilitas berkaitan dengan tingkat konsistensi dalam melakukan penilaian. Hal ini juga harus diperhatikan dengan baik karena pada praktiknya nanti, penilaian proposal penelitian akan dilakukan oleh beberapa orang yang berbeda. Apabila satu buah proposal penelitian memiliki dua nilai dengan tingkat disparita yang tinggi, maka dapat dipastikan bahwa nilai yang didapat tidak reliabel (tidak konsisten).

Penelitian ini bertujuan untuk menyelesaikan permasalahan yang telah dijelaskan sebelumnya. Tiga permasalahan di atas dapat diselesaikan dengan mengembangkan rubrik penilaian. Rubrik adalah sebuah instrumen penilaian yang biasa digunakan untuk menilai produk dan kinerja peserta didik. Rubrik terdiri dari beberapa dimensi penilaian di mana masing-masing dimensi penilaian dideskripsikan secara bertingkat sesuai dengan tingkat kesulitan pekerjaan (Brookhart, 2013). Setiap tingkatan pekerjaan dideskripsikan dengan spesifik sehingga dapat dijadikan dasar untuk melakukan penilaian. Dengan bantuan rubrik, diharapkan dapat menghasilkan nilai yang adil, valid, dan reliabel.

Sebelum melakukan pengembangan, peneliti telah melakukan kajian terkait model-model apa saja yang digunakan untuk mengembangkan sebuah rubrik penilaian. Berdasarkan kajian tersebut, ditemukan beberapa model yang digunakan antara lain model *4D Thiagarjan, Borg and Gall*, dan *Crusan*. Selain model tersebut, terdapat juga beberapa pengembangan yang tidak secara eksplisit menerangkan perihal model yang digunakan namun dapat dijelaskan secara logis. Masing-masing model dan proses pengembangan pada penelitian di atas memiliki kelebihan dan kekurangannya masing-masing. Selanjutnya, peneliti melakukan pengelompokan berdasarkan kelebihan dan kekurangan dari masing-masing model dan proses pengembangan. Adapun pengelompokan tersebut akan disajikan dalam tabel 1.

Kategori 1 terdiri dari beberapa pengembangan yang dilakukan berdasarkan model pengembangan tertentu. Model yang digunakan adalah model *4G Thiagarjan, Borg and Gall*, dan *Crusan*. Penggunaan Model *4D Thiagarjan* dan *Borg and Gall* pada umumnya diaplikasikan pada pengembangan di dalam negeri. Model *4D Thiagarjan* digunakan oleh Fitriani & Yarmayani (2018); Aji, Hudha, Huda, Nandiyanto, & Abdullah (2018); Febriyanti, Harsiati, & Dermawan (2017); dan Hermawan et.al. (2017). Adapun model pengembangan *Borg and Gall* digunakan oleh Latifa, Rahman, Hamra, Jabu, & Nur (2015). Kedua model tersebut pada dasarnya dapat digunakan untuk mengembangkan sebuah rubrik penilaian, namun perlu diketahui bahwa kedua model tersebut tidak didesain untuk mengembangkan sebuah rubrik penilaian. Kedua model tersebut umum digunakan untuk mengembangkan media pembelajaran. Selain itu, keduanya sudah berumur cukup lawas. Model *4D Thiagarjan* dan *Borg and Gall* masing-masing pertama kali diperkenalkan pada tahun 1974 dan 1989 (Arifin, 2011b).

Contoh pengembangan pada kategori 1 lainnya adalah pengembangan yang dilakukan oleh Rakedzon & Tsabari (2017). Berbeda dengan contoh sebelumnya, pengembangan ini dilakukan di negara israel. Model yang digunakan pun masih terhitung baru yaitu model *Crusan*. Namun demikian, model *crusan* memiliki karakteristik yang sama dengan model *4G Thiagarjan* dan model *Borg and Gall*. Ketiganya sama-sama melakukan tahapan uji validitas namun tidak melakukan uji reliabilitas. Padahal sebagaimana diketahui bahwa reliabilitas adalah salah satu kriteria penting dalam melakukan penilaian.

Kekurangan pada kategori 1 terjawab pada pengembangan di kategori 2. Selain melakukan uji validitas, rubrik juga dikembangkan dengan melakukan uji reliabilitas. Hanya saja tidak seperti pada kategori 1, pengembangan rubrik pada kategori 2 tidak dikembangkan berdasarkan pada model pengembangan tertentu. Pengembangan pada kategori 2 diantaranya dilakukan oleh Tam (2017); Aulet, Moore, Callas, Nicholas, & Hulme (2018); dan Sweeney, Rollins, Gantt, Swanson, & Ravitz (2019).

**Tabel 1. Pengelompokan Pengembangan Rubrik Berdasarkan Kelebihan dan Kekurangan**

|                   | <b>Kelebihan</b>   | <b>Kekurangan</b>  |  |
|-------------------|--|--|--|
| <b>Kategori 1</b> | <ul style="list-style-type: none"> <li>• Fitriani &amp; Yarmayani (2018)</li> <li>• Aji, Hudha, Huda, Nandiyanto, &amp; Abdullah (2018)</li> <li>• Febriyanti, Harsiati, &amp; Dermawan (2017)</li> <li>• Hermawan et al., (2017)</li> <li>• Rakedzon &amp; Tsabari (2017)</li> <li>• Latifa, Rahman, Hamra, Jabu, &amp; Nur (2015)</li> </ul> | <ul style="list-style-type: none"> <li>• Penelitian dilakukan dengan model yang memiliki landasan literatur yang jelas.</li> <li>• Melakukan uji validitas instrumen.</li> </ul> | <ul style="list-style-type: none"> <li>• Model yang digunakan terhitung lama dan tidak didesain secara khusus untuk mengembangkan rubrik [kecuali pada penelitian Rakedzon &amp; Tsabari (2017)]</li> <li>• Tidak melakukan uji reliabilitas instrumen.</li> </ul> |
| <b>Kategori 2</b> | <ul style="list-style-type: none"> <li>• Tam (2017)</li> <li>• Aulet, Moore, Callas, Nicholas, &amp; Hulme (2018)</li> <li>• Sweeney, Rollins, Gantt, Swanson, &amp; Ravitz (2019)</li> </ul>  | <ul style="list-style-type: none"> <li>• Melakukan uji validitas instrumen.</li> <li>• Melakukan uji reliabilitas instrumen.</li> </ul>  | <ul style="list-style-type: none"> <li>• Tidak menjelaskan model/prosedur pengembangan rubrik.</li> </ul>  |

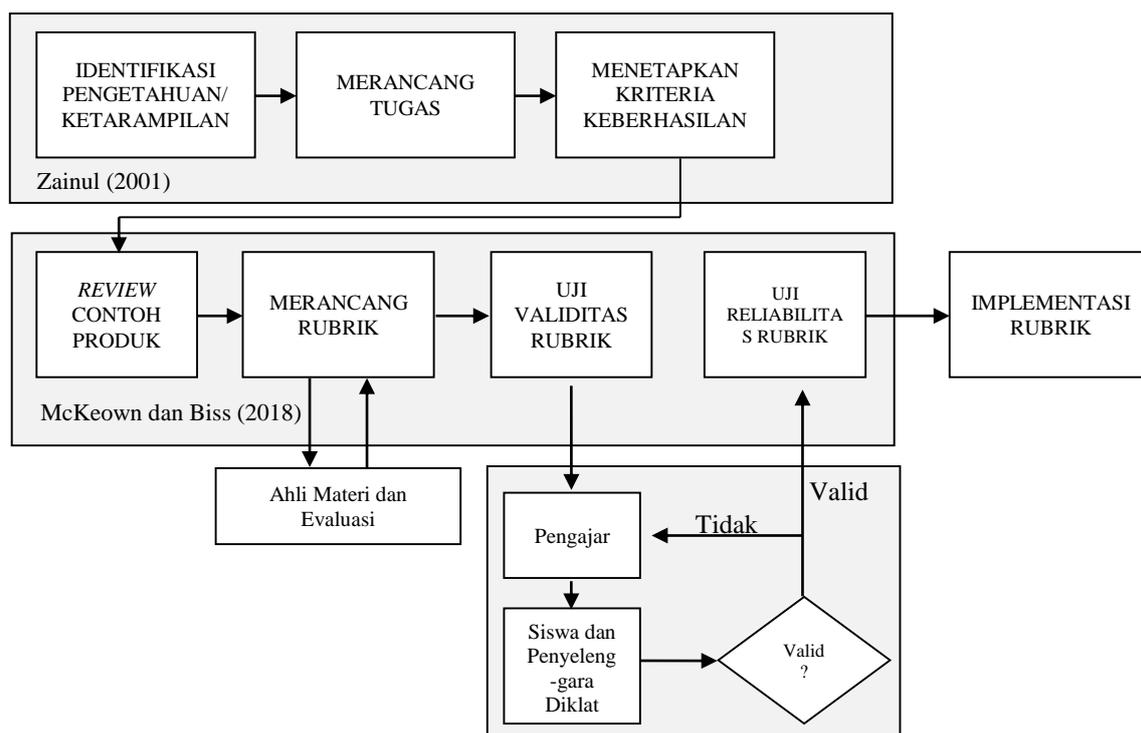
Berdasarkan kajian di atas, maka perlu dilakukan sebuah pengembangan rubrik yang mencakup kelebihan pada kategori 1 dan 2. Pengembangan rubrik harus dilakukan berdasarkan model pengembangan tertentu. Selain itu, prosedur pengembangan juga harus meliputi uji validitas dan uji reliabilitas.

### METODE

Pengembangan rubrik dilakukan dengan menggunakan dua buah model pengembangan. Model yang digunakan adalah model Zainul (2001) dan Model Mckeown dan Biss (2018). Model pertama merupakan model yang digunakan untuk menilai kinerja peserta didik. Sementara model kedua merupakan model yang secara spesifik digunakan untuk mengembangkan sebuah rubrik penilaian.

Penggunaan model Zainul (2001) bertujuan untuk merumuskan proses penilaian alternatif. Model ini dipilih karena penilaian proposal penelitian pada program PPJFP termasuk dalam rumpun penilaian alternatif. Adapun prosedur dalam model ini adalah sebagai berikut. *Pertama*, mengidentifikasi pengetahuan dan ketrampilan yang diharapkan dapat dimiliki siswa. *Kedua*, merancang tugas yang memungkinkan siswa dapat menunjukkan kemampuan berpikir & ketrampilan. *Ketiga*, menetapkan kriteria keberhasilan yang akan dijadikan tolok ukur untuk menyatakan bahwa seorang siswa telah mencapai tingkat *mastery* pengetahuan atau ketrampilan yang diharapkan.

Model selanjutnya adalah model Mckeown dan Biss (2018). Model ini secara spesifik digunakan untuk mengembangkan sebuah rubrik penilaian. Model ini dipilih sebagai jawaban atas kekurangan pada penelitian sejenis yang telah dibahas pada bagian sebelumnya. Tidak seperti model lainnya, model ini termasuk model yang mutakhir karena baru diperkenalkan pada tahun 2018. Selain itu, model ini juga meliputi prosedur uji validitas dan uji reliabilitas. Adapun prosedur pengembangan pada model ini adalah sebagai berikut. *Pertama*, review contoh rubrik yang sudah ada. *Kedua*, membuat rancangan rubrik. *Ketiga*, menguji validitas rubrik. *Keempat*, menguji reliabilitas rubrik. *Kelima*, implementasi rubrik.



**Gambar 1. Model Pengembangan Rubrik Penilaian Proposal Penelitian Menggabungkan model Asmawi dan model Mckeown and Biss**

Langkah pertama adalah dengan melakukan penelaahan (*review*) contoh rubrik penilaian yang telah ada sebelumnya. Penelaahan dilakukan terhadap rubrik sejenis yang sudah teruji validitas dan reliabilitasnya. Dengan melakukan penelaahan tersebut diharapkan dapat memudahkan untuk merancang rubrik yang valid dan reliabel.

Tahapan selanjutnya adalah membuat rancangan rubrik penilaian. Perancangan meliputi tujuan, dimensi, skala, deskripsi dimensi, dan bobot penilaian. Apabila ada ketersediaan waktu dan dana proses perancangan dapat juga melibatkan para ahli (Mckeown & Biss, 2018). Ahli yang terlibat dapat berasal dari ahli evaluasi, materi substansi, atau keduanya. Dari tahapan ini akan didapatkan data berupa masukan dari para ahli. Setelah rancangan selesai dibuat, maka rancangan tersebut siap untuk diuji tingkat validitasnya.

Merujuk pada prosedur pengembangan Mckeown dan Biss (2018), uji validitas dilakukan dengan pendekatan subjektif yaitu dengan cara mencari kesepakatan (konsensus) di antara para Pengampu kebijakan. Pengampu kebijakan yang dimaksud adalah pengembang rubrik, ahli substansi, pengajar, asisten pengajar, peserta didik, dan pihak terkait lainnya. Uji validitas dilakukan dengan metode *multiple round test*. Pengujian dengan metode *multiple round test* dilakukan secara berulang-ulang hingga dicapai kesepakatan dari seluruh *stakeholder*.

Tahapan selanjutnya adalah menguji tingkat reliabilitas rubrik. Uji reliabilitas dilakukan dengan pendekatan *inter rater reliability (IRR)*. Pengujian dilakukan dengan cara melibatkan dua orang penilai. Masing-masing penilai akan diminta untuk menilai sebuah proposal penelitian yang ditulis oleh peserta pelatihan. Kedua buah nilai selanjutnya dianalisis tingkat korelasinya dengan menggunakan rumus *product moment* (Arifin, 2011a). Selanjutnya, nilai koefisien korelasi dibandingkan dengan keterangan yang terdapat pada tabel 2.

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

$r_{xy}$  = koefisien korelasi

$\sum xy$  = jumlah produk x dan y

**Tabel 2. Tingkat Kekuatan Berdasarkan Nilai Koefisien Korelasi**

| Nilai Koefisien Korelasi | Makna                       |
|--------------------------|-----------------------------|
| Lebih dari 0,9           | Sangat kuat                 |
| 0,8 s.d. 0,89            | Kuat                        |
| 0,7 s.d. 0,79            | Cukup, dapat diterima       |
| Kurang dari 0,7          | Lemah, tidak dapat diterima |

Setelah melewati seluruh tahapan pengembangan, selanjutnya rubrik penilaian siap untuk digunakan. Apabila rubrik akan digunakan dalam skala penilaian yang lebih besar, maka lakukan kembali penyamaan persepsi dengan para penilai terlebih dahulu.

## HASIL

Proses pengembangan rubrik meliputi dua tahapan. Tahapan pertama adalah mendesain sistem penilaian alternatif sebagaimana dijelaskan oleh Zainul (2001). Tahapan selanjutnya adalah mengembangkan rubrik penilaian dengan mengacu pada model Mckeown dan Biss (2018).

Hasil pada tahapan pertama meliputi tiga hal, yaitu pengetahuan/keterampilan yang harus dikuasai oleh peserta, bentuk penugasan yang akan diberikan guna melihat kemampuan yang telah dikuasai oleh peserta, serta kriteria yang dapat digunakan sebagai acuan keberhasilan. Identifikasi pengetahuan/keterampilan dilakukan dengan cara menganalisis kurikulum pelatihan yang telah ada sebelumnya. Berdasarkan analisis tersebut didapatkan delapan jenis pengetahuan yang harus dikuasai oleh peserta pelatihan. Untuk melihat kemampuan peserta, selanjutnya peserta akan ditugaskan untuk menulis sebuah proposal penelitian. Selain menganalisis kurikulum pelatihan dilakukan juga analisis terhadap bahan ajar untuk menetapkan kriteria keberhasilan yang akan dijadikan tolak ukur keberhasilan peserta pelatihan. Pengetahuan/keterampilan beserta kriteria keberhasilannya akan disajikan dalam tabel 3.

**Tabel 3. Pengetahuan/Keterampilan dalam Menulis Proposal Penelitian beserta Kriteria Keberhasilannya**

| Pengathuan/Keterampilan           | Kriteria Keberhasilan   |
|-----------------------------------|---|
| Menentukan Topik Penelitian       | <ul style="list-style-type: none"> <li>• Topik penelitian menarik dan penting untuk dilakukan</li> <li>• Memiliki kebermanfaatan <i>science for science, science for community, science for stakeholders</i>.</li> <li>• Dianalisis menggunakan <i>theory of change</i> dan mulai mengarah pada pendekatan interdisiplin.</li> </ul>  |
| Menyusun Rumusan Masalah          | <ul style="list-style-type: none"> <li>• Dilakukan dengan terlebih dahulu melakukan identifikasi masalah.</li> <li>• Identifikasi meliputi juga akibat dan sumber masalah.</li> </ul>   |
| Merumuskan Tujuan Penelitian      | <ul style="list-style-type: none"> <li>• Sejalan dengan agenda penyandang dana penelitian.</li> </ul>   |
| Menemukan <i>State of The Art</i> | <ul style="list-style-type: none"> <li>• Mengikuti perkembangan dari topik yang akan diteliti. Tidak hanya pada level nasional, tetapi juga pada level internasional.</li> <li>• Menemukan celah kontribusi di antara penelitian-penelitian yang sudah ada.</li> <li>• Menunjukkan <i>novelty</i> penelitian, baik dari sisi hasil maupun metode yang akan digunakan.</li> <li>• Terhindar dari duplikasi atau pengulangan penelitian.</li> </ul> |
| Merumuskan Manfaat Penelitian     | Dianalisis dengan menggunakan <i>research impact pathway</i> . Peneliti dapat menjelaskan <i>output, outcome</i> , dan menghubungkannya dengan <i>impact</i> yang mungkin terjadi.  |
| Menentukan Metode Penelitian      | <ul style="list-style-type: none"> <li>• Metode didasarkan pada literatur yang kredibel.</li> <li>• Metode penelitian dapat dijelaskan dengan logis.</li> <li>• Ketepatan dalam memilih tempat penelitian, sumber data/narasumber, teknik pengumpulan data, teknik pengolahan data, teknik analisis dan penarikan kesimpulan.</li> </ul>  |
| Memilih Literatur yang Relevan    | <ul style="list-style-type: none"> <li>• Literatur yang digunakan harus relevan.</li> <li>• Tahun terbit buku tidak lebih dari 20 tahun.</li> <li>• Tahun terbit artikel ilmiah tidak lebih dari 5 tahun.</li> </ul>  |
| Pemilihan Tata Bahasa             | <ul style="list-style-type: none"> <li>• Padat dan ringkas</li> <li>• Jelas</li> <li>• Sederhana. Kosakata tidak muluk-muluk, sintaksis jelas.</li> <li>• Koheren. Tulisan utuh dan logis.</li> </ul>   |

Delapan acuan penilaian di atas selanjutnya dikembangkan menjadi sebuah rubrik penilaian. Pengembangan rubrik dilakukan dengan menggunakan model Mckeown & Biss (2018). Proses pengembangan dimulai dengan melakukan penelaahan terhadap contoh-contoh rubrik yang sudah ada sebelumnya. Hal ini dilakukan agar memudahkan proses pengembangan pada tahapan selanjutnya. Setelah selesai menelaah, selanjutnya dilakukan proses penyusunan *draft* rubrik. Rancangan rubrik selanjutnya dikonsultasikan pada dua orang ahli, yaitu ahli materi dan ahli evaluasi. Setelah melewati tahapan pengembangan yang pertama, selanjutnya dilakukan tahapan kedua yaitu uji validitas. Uji validitas dilakukan dengan cara mencari

kesepakatan/konsensus di antara pengajar serta melihat persepsi dan masukan dari peserta dan penyelenggara diklat. Kronologis perancangan dan uji validitas rubrik akan disajikan dalam tabel 4. Adapun persepsi peserta akan disajikan pada tabel 5 dan 6.

**Tabel 4. Alur Kegiatan Penyusunan *Draft*, Uji Validitas, dan Urutan Revisinya**

| Tanggal          | Kegiatan   | Urutan Revisi |
|------------------|--|---------------|
| 13 Oktober 2020  | Penyusunan rubrik penilaian (tanpa masukan ahli) selesai dibuat. | Revisi ke-1   |
| 25 Oktober 2020  | Diskusi dengan Ahli Materi                                       | Revisi ke-2   |
| 30 Oktober 2020  | Diskusi dengan Ahli Evaluasi                                     | Revisi ke-3   |
| 2 November 2020  | Diskusi dengan Ahli Evaluasi                                     | Revisi ke-4   |
| 4 November 2020  | Diskusi dengan Ahli Evaluasi                                     | Revisi ke-5   |
| 11 November 2020 | Diskusi dengan Ahli Evaluasi                                     | Revisi ke-6   |
| 18 November 2020 | FGD dengan Pengajar sesi I                                       | Revisi ke-7   |
| 23 November 2020 | Masukan Peserta dan Penyelenggara Diklat sesi I                  | Revisi ke-8   |
| 27 November 2020 | FGD dengan Pengajar sesi II                                      | Revisi ke-9   |
| 28 November 2020 | Masukan Uji Coba Penyelenggara Diklat Peserta sesi II            | Revisi ke-10  |

**Tabel 5. Persepsi Peserta terhadap Rubrik Penilaian Proposal pada Uji Coba Sesi I**

| Dimensi Penilaian          | Istilah<br>(Terminology) |       | Dimensi Penilaian |       | Membantu peserta menyelesaikan tugas |       | Sistem Penilaian |       |
|----------------------------|--------------------------|-------|-------------------|-------|--------------------------------------|-------|------------------|-------|
|                            | Jelas                    | Tidak | Jelas             | Tidak | Membantu                             | Tidak | Jelas            | Tidak |
| Pemilihan Topik Penelitian |                          |       | 100%              | 0%    | 95%                                  | 5%    | 95%              | 5%    |
| <i>State of The Art</i>    |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Rumusan Masalah            |                          |       | 100%              | 0%    | 95%                                  | 5%    | 100%             | 0%    |
| Tujuan Penelitian          | 100%                     | 0%    | 100%              | 0%    | 95%                                  | 5%    | 100%             | 0%    |
| Manfaat Penelitian         |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Metode Penelitian          |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Kajian Literatur           |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Tata Bahasa                |                          |       | 100%              | 0%    | 95%                                  | 5%    | 95%              | 5%    |

**Tabel 6. Persepsi Peserta terhadap Rubrik Penilaian Proposal pada Uji Coba Sesi I**

| Dimensi Penilaian          | Istilah<br>(Terminology) |       | Dimensi Penilaian |       | Membantu peserta menyelesaikan tugas |       | Sistem Penilaian |       |
|----------------------------|--------------------------|-------|-------------------|-------|--------------------------------------|-------|------------------|-------|
|                            | Jelas                    | Tidak | Jelas             | Tidak | Membantu                             | Tidak | Jelas            | Tidak |
| Pemilihan Topik Penelitian |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| <i>State of The Art</i>    |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Rumusan Masalah            |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Tujuan Penelitian          | 100%                     | 0%    | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Manfaat Penelitian         |                          |       | 100%              | 0%    | 0,89%                                | 0,11% | 100%             | 0%    |
| Metode Penelitian          |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Kajian Literatur           |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |
| Tata Bahasa                |                          |       | 100%              | 0%    | 100%                                 | 0%    | 100%             | 0%    |

Tahapan selanjutnya adalah uji reliabilitas. Uji reliabilitas dilakukan dengan cara membandingkan dua buah nilai yang didapat dari dua orang penilai. Nilai yang didapat kemudian diolah sebagaimana tertera pada tabel 7.

**Tabel 7. Perhitungan Koefisien Korelasi**

| Peserta   | X<br>(Nilai Pengajar 1) | Y<br>(Nilai Pengajar 2) | x      | y      | x <sup>2</sup> | y <sup>2</sup> | xy     |
|-----------|-------------------------|-------------------------|--------|--------|----------------|----------------|--------|
| Peserta 1 | 90,00                   | 80,00                   | 2,40   | 0,20   | 5,76           | 0,04           | 0,48   |
| Peserta 2 | 98,00                   | 94,00                   | 10,40  | 14,20  | 108,16         | 201,64         | 147,68 |
| Peserta 3 | 72,00                   | 66,00                   | -15,60 | -13,80 | 243,36         | 190,44         | 215,28 |
| Peserta 4 | 84,00                   | 78,00                   | -3,60  | -1,80  | 12,96          | 3,24           | 6,48   |

**Tabel 7. Perhitungan Koefisien Korelasi (Lanjutan)**

| Peserta    | X<br>(Nilai Pengajar 1) | Y<br>(Nilai Pengajar 2) | x     | y     | x <sup>2</sup> | y <sup>2</sup> | xy     |
|------------|-------------------------|-------------------------|-------|-------|----------------|----------------|--------|
| Peserta 5  | 94,00                   | 96,00                   | 6,40  | 16,20 | 40,96          | 262,44         | 103,68 |
| Peserta 6  | 84,00                   | 82,00                   | -3,60 | 2,20  | 12,96          | 4,84           | -7,92  |
| Peserta 7  | 86,00                   | 70,00                   | -1,60 | -9,80 | 2,56           | 96,04          | 15,68  |
| Peserta 8  | 88,00                   | 80,00                   | 0,40  | 0,20  | 0,16           | 0,04           | 0,08   |
| Peserta 9  | 90,00                   | 74,00                   | 2,40  | -5,80 | 5,76           | 33,64          | -13,92 |
| Peserta 10 | 90,00                   | 78,00                   | 2,40  | -1,80 | 5,76           | 3,24           | -4,32  |
| $\Sigma$   | 876,00                  | 798,00                  | 0,00  | 0,00  | 438,40         | 795,60         | 463,20 |
| rata-rata  | 87,60                   | 79,80                   |       |       |                |                |        |

Selanjutnya data diatas diolah dengan menggunakan rumus *product moment*. Setelah dihitung didapat nilai koefisien korelasi sebesar 0,78.

### PEMBAHASAN

Tabel 3 menunjukkan hasil identifikasi terhadap pengetahuan/keterampilan yang harus ditunjukkan oleh peserta ketika menulis sebuah proposal penelitian. Delapan pengetahuan/keterampilan tersebut selanjutnya akan dijadikan acuan dalam menilai sebuah proposal penelitian. Delapan pengetahuan/keterampilan yang dimaksud, meliputi topik penelitian, rumusan masalah, tujuan penelitian, *state of the art*, manfaat penelitian, metode penelitian, kajian literatur, dan pemilihan tata bahasa. Masing-masing pengetahuan/keterampilan disertai dengan kriteria keberhasilan yang harus dipenuhi oleh peserta pelatihan. Peserta akan dinyatakan lulus apabila berhasil memenuhi kriteria keberhasilan tersebut.

Acuan penilaian yang pertama adalah kemampuan dalam memilih topik penelitian. Topik yang dipilih harus menarik dan penting untuk dilakukan. Penelitian dapat dinyatakan penting apabila dapat memberikan kontribusi bagi pengembangan disiplin ilmu pengetahuan (*science to science*), komunitas ilmiah (*science to community*), dan hasilnya dapat dijadikan rujukan oleh pengambil kebijakan/pemerintah (*science to stake holders*). Selain itu, topik penelitian juga harus aktual dan peka dengan perkembangan yang sedang terjadi. Lebih jauh lagi, seorang peneliti juga harus dapat melihat topik penelitian yang dia pilih dengan menggunakan kaca mata multidisiplin.

Acuan penilaian selanjutnya adalah rumusan masalah dan tujuan penelitian. Rumusan masalah harus diawali dengan identifikasi masalah terlebih dahulu. Selain itu peneliti juga harus melakukan identifikasi sumber dan dampak yang diakibatkan oleh masalah tersebut. Adapun untuk tujuan penelitian, seorang peneliti harus memahami visi dan misi lembaga penyandang dana penelitian yang akan ia tuju. Dengan demikian, peneliti harus dapat menyelaraskan tujuan penelitiannya dengan visi dan misi lembaga tersebut. Peneliti harus sadar bahwa salah satu tujuan dari penulisan proposal penelitian adalah untuk mencari dana penelitian. Dengan demikian, diharapkan peluang pendanaan juga semakin besar.

Acuan penilaian selanjutnya adalah *state of the art*. Untuk dapat memenuhi kriteria ini, peneliti harus mengikuti perkembangan topik penelitian yang akan ia lakukan hingga pada tataran internasional. Dengan demikian ia dapat menempatkan posisi penelitiannya di antara penelitian-penelitian yang lain. Tidak hanya itu, penelitian yang akan dilakukan juga harus memiliki nilai kebaruan dibanding dengan penelitian lainnya. Kebaruan dapat dilihat baik dari sisi produk maupun cara/metode yang ia gunakan. Selanjutnya, penelitian yang akan dilakukan juga harus terhindar dari duplikasi/pengulangan penelitian.

Acuan penilaian selanjutnya berkaitan dengan manfaat penelitian. Manfaat penelitian dapat dianalisis dengan menggunakan metode *research impact pathway*. Melalui metode tersebut, peneliti diharapkan dapat menjelaskan *output*, *outcome*, serta menghubungkannya dengan *impact* yang mungkin terjadi. Acuan penilaian yang keenam adalah metode penelitian. Metode yang digunakan harus didasarkan pada literatur yang kredibel. Selain itu, peneliti juga harus dapat menjelaskan metode penelitian dengan logis. Peneliti juga harus dapat memilih tempat penelitian, sumber data/narasumber, teknik pengumpulan data, teknik pengolahan data, teknik analisis, dan penarikan kesimpulan.

Dimensi penilaian yang ketujuh berkaitan dengan penggunaan literatur yang digunakan dalam penelitian. Literatur yang digunakan hendaknya harus relevan dengan penelitian yang akan dilakukan. Apabila literatur yang digunakan berupa buku, maka hendaknya diterbitkan tidak lebih dari 20 tahun. Apabila literatur yang digunakan berasal dari artikel ilmiah maka sebaiknya diterbitkan tidak lebih dari lima tahun. Dimensi penilaian yang terakhir adalah penggunaan bahasa. Bahasa yang digunakan dalam sebuah proposal penelitian hendaknya padat dan ringkas, jelas, serta sederhana. Kosakata yang digunakan tidak muluk-muluk, serta memiliki sintaksis yang jelas, koheren, utuh, dan logis.

Pada tabel 4 diperlihatkan kronologis pengembangan rubrik mulai dari tahapan perancangan sampai dengan uji validitas. Tahapan perancangan dilakukan dengan melibatkan ahli materi. Tahapan uji validitas dilakukan dengan mencari kesepakatan (konsensus) di antara pengajar serta disempurnakan dengan cara melihat persepsi yang diberikan oleh peserta pelatihan. Diskusi dengan ahli materi dilakukan sebanyak satu kali. Pertemuan tersebut dilaksanakan pada tanggal 2 Oktober 2020 bertempat di Bhumi Satwa, *Cibinong Science Center (CSC)*, Kabupaten Bogor. Ahli materi Profesor Riset di Pusat Penelitian Biologi, Lembaga Ilmu Pengetahuan Indonesia. Beliau memberikan masukan agar peneliti melakukan perbaikan rubrik pada dimensi penilaian *state of the art*, metode penelitian, kajian literatur, dan tata bahasa.

Masukan yang pertama berkaitan dengan penentuan skala kinerja (*leveling*) pada dimensi penilaian *state of the art*. Peneliti mendapatkan masukan agar memperbaiki deskripsi kinerja pada tiap skala kinerja. Dengan demikian, terdapat batasan yang tegas untuk membedakan kualitas *state of the art* mulai dari level terendah (1) sampai dengan level tertinggi (5). Peneliti juga mendapat masukan untuk memperbaiki kalimat yang digunakan dalam dimensi penilaian ini. Hal ini perlu dilakukan agar kelak rubrik dapat dipahami dengan jelas oleh penilai yang akan datang.

Masukan untuk dimensi penilaian metode penelitian meliputi perbaikan redaksional pada deskripsi kinerja mulai dari level terendah (1) sampai dengan level tertinggi (5). Masukan yang pertama adalah bahwa pada level 1 peserta cukup menjelaskan metode secara umum saja. Pada level 2 peserta sudah mulai menjelaskan dengan detail, tetapi masih terdapat kesalahan pada beberapa aspek. Pada level 3 peserta dapat menjelaskan metode dengan tepat namun konsep dan alat bukan sesuatu yang mutakhir (terkini). Pada level 4 peserta sudah menggunakan konsep dan alat yang mutakhir (terkini). Pada level 5, peserta benar-benar memahami metode yang akan ia gunakan ditunjukkan dengan kemampuannya dalam menjelaskan metode penelitian dengan logis. Perlu peneliti sampaikan, bahwa perbaikan yang peneliti dapatkan pada tahapan ini bukanlah perbaikan final. Selanjutnya peneliti akan mendapatkan masukan dari pengajar lain melalui kegiatan uji validitas yang pertama dan kedua. Masukan selanjutnya berkaitan dengan dimensi penilaian kajian literatur. Peneliti mendapatkan masukan tentang batas lawas sebuah literatur. Semula peneliti menuliskan bahwa batas maksimal usia buku dan artikel ilmiah masing-masing adalah 20 dan 5 tahun. Ahli materi menyarankan agar peneliti menghilangkan satuan waktu tersebut. Menurut penilai, perkembangan ilmu pengetahuan bisa terjadi lebih cepat mengingat perubahan dalam bidang ilmu pengetahuan bisa terjadi dengan sangat dinamis.

Masukan yang terakhir berkaitan dengan dimensi penilaian pemilihan tata bahasa. Pada dimensi ini tidak banyak terjadi perubahan redaksional. Perubahan hanya terjadi pada kinerja level ke-3 saja. Semula tertulis “Proposal penelitian bisa difahami dengan baik walaupun masih terdapat kekurangan dalam penyusunan kalimat dan paragraf.” menjadi “Proposal penelitian bisa dipahami dengan baik walaupun masih terdapat kekurangan dalam pengungkapan bahasa.” Setelah dilakukan perbaikan berdasarkan masukan dari ahli materi, selanjutnya dilakukan diskusi dengan ahli evaluasi. Diskusi dilakukan sebanyak empat kali. Seluruh pertemuan dilakukan secara daring dan dilaksanakan pada tanggal 30 Oktober 2020, 2 November 2020, 4 November 2020, dan 11 November 2020. Melalui empat pertemuan tersebut, peneliti mendapatkan masukan untuk perbaikan rubrik. Masukan yang diberikan mengenai konsistensi kalimat pada rubrik, perbaikan deskripsi kinerja, dan bobot penilaian.

Konsistensi kalimat yang digunakan pada rubrik berkaitan dengan objek yang dinilai oleh rubrik. Pada beberapa dimensi penilaian terkesan bahwa rubrik dipakai untuk menilai proposal, namun pada dimensi yang lain muncul kesan bahwa yang dinilai adalah penulis proposalnya. Oleh karenanya diberikan saran agar rubrik fokus menilai proposal penelitiannya saja. Perbaikan selanjutnya berkaitan dengan deskripsi kinerja. Deskripsi kinerja yang disusun harus dapat memperlihatkan gradasi. Deskripsi kinerja disusun mulai dari tingkatan pekerjaan terendah sampai dengan tingkatan pekerjaan terbaik. Kalimat yang digunakan juga harus mudah dimengerti agar tidak menyulitkan calon penilai. Perbaikan banyak dilakukan pada dimensi penilaian rumusan masalah, metode penelitian, dan tata bahasa.

Saran yang terakhir berkaitan dengan pembobotan nilai. Pembobotan nilai harus dilakukan seadil mungkin. Dimensi penilaian dengan jenis pekerjaan lebih sulit harus diberikan bobot penilaian yang lebih besar dibanding bobot pada dimensi penilaian yang lebih mudah. Setelah dilakukan perbaikan selanjutnya dilakukan tahapan uji validitas. Uji validitas dilakukan sebanyak dua sesi. Tiap sesi melibatkan pengajar, peserta, dan pihak penyelenggara diklat. Sesi I diikuti oleh dua orang pengajar, seluruh peserta PPJFP gelombang 6 (20 orang), dan pihak penyelenggara diklat. Sesi II diikuti oleh enam orang pengajar, seluruh peserta PPJFP gelombang 7 (9 orang), dan pihak penyelenggara diklat. Masing-masing tahapan uji coba pada tiap sesi ditindaklanjuti dengan perbaikan berdasarkan masukan dari seluruh pihak yang terlibat.

Pada tanggal 18 November 2020, dilaksanakan *focused group discussion* (FGD) untuk menguji coba serta menggali masukan sebagai bahan perbaikan rubrik. Kegiatan ini diikuti oleh peneliti bersama dengan dua orang pengajar secara daring. Setelah dilakukan ujicoba selanjutnya para pengajar diminta untuk memberikan saran dan komentar. Masukan dan komentar yang didapatkan selanjutnya akan dijadikan dasar dalam melakukan perbaikan. Terdapat tiga masukan yang peneliti dapatkan ketika melakukan FGD bersama dengan pengajar di atas. Masukan yang pertama berkaitan dengan pemilihan kata pada rubrik. Pengajar memberikan masukan agar seluruh frase “tidak melakukan” pada rubrik diganti dengan “belum melakukan”. Hal ini berkaitan dengan salah satu sifat rubrik yaitu bahwa rubrik harus bersifat memotivasi bukan demotivasi. Selain itu, terdapat juga beberapa masukan terkait penentuan deskripsi kinerja pada dimensi penilaian metode penelitian. Masukan terakhir berkaitan dengan dimensi penilaian sumber literatur. Keduanya berpendapat bahwa ada sebagian bidang penelitian seperti penelitian sejarah yang justru membutuhkan sumber rujukan lawas sebagai bahan kajiannya. Oleh karenanya, rubrik juga harus disesuaikan dengan kondisi tersebut. Setelah mendapatkan masukan kemudian rubrik penilaian diperbaiki kembali. Rubrik yang sudah diperbaiki selanjutnya akan diperlihatkan pada peserta pelatihan. Selanjutnya peserta diminta untuk mengisi instrumen penelitian.

Tujuan dari dilibatkannya peserta pelatihan dalam proses uji validitas rubrik menurut Mckeown & Biss (2018) adalah untuk melihat empat kriteria yang harus ada pada sebuah rubrik. Empat kriteria tersebut berkaitan dengan persepsi siswa/peserta didik terhadap sebuah rubrik penilaian. Rubrik dinyatakan baik apabila telah memenuhi empat kriteria tersebut. Keempat kriteria tersebut, meliputi (1) memiliki kejelasan terminologi (istilah-istilah yang ada pada rubrik); (2) memiliki *clarity of criteria* (kejelasan dimensi penilaian pada rubrik); (3) dapat membimbing peserta dalam menyelesaikan tugasnya; (4) memiliki sistem penilaian yang mudah dimengerti.

Data yang ditampilkan pada tabel 5 merupakan persepsi peserta terhadap rubrik penilaian proposal. Persepsi tersebut meliputi kejelasan istilah, kejelasan dimensi penilaian, kemampuan rubrik dalam membantu menyelesaikan tugas, dan kejelasan sistem penilaian. Berdasarkan data di atas diketahui bahwa seluruh peserta menganggap istilah-istilah yang ada pada rubrik dapat dimengerti dengan jelas. Tidak ada istilah yang aneh atau asing bagi peserta. Istilah yang dimaksud meliputi seluruh istilah yang digunakan pada seluruh dimensi penilaian mulai pemilihan topik penelitian sampai dengan pemilihan tata bahasa. Istilah yang dipergunakan pada rubrik harus istilah yang dapat dipahami. Hal ini berkaitan dengan salah satu sifat rubrik yaitu dapat mendorong perbaikan kinerja/produk yang dihasilkan oleh peserta didik. Hasil penilaian rubrik dapat digunakan sebagai masukan (*feedback*) yang dapat digunakan untuk memperbaiki kinerja/produk yang mereka hasilkan (Jönsson & Panadero, 2017). Oleh karena itu, harus dipastikan bahwa setiap istilah yang digunakan dapat dipahami dengan jelas.

Selain menampilkan aspek kejelasan istilah, tabel 5 juga menampilkan persepsi tentang kejelasan dimensi penilaian, kemampuan rubrik dalam membimbing pekerjaan, serta kejelasan penilaian. Berdasarkan data tersebut, diketahui bahwa seluruh peserta (100%) memahami seluruh dimensi penilaian yang ada pada rubrik. Selain itu, sebagian besar peserta menganggap bahwa rubrik dapat membimbing mereka dalam menulis proposal penelitian. Hanya 5% peserta (1 orang) yang berpendapat bahwa rubrik tidak membantu dalam hal menentukan topik penelitian, membuat rumusan masalah, membuat tujuan penelitian, dan menentukan tata bahasa. Sebagian besar peserta juga memahami sistem penilaian yang digunakan oleh rubrik. Hanya 5% peserta (1 orang) yang tidak memahami sistem penilaian yaitu masing-masing pada dimensi penilaian pemilihan topik penelitian dan tata bahasa. Persepsi negatif dari peserta menjadi catatan yang berharga bagi peneliti. Ketidakpuasan tersebut (walaupun masing-masing hanya berjumlah 5%/1 orang peserta) mengindikasikan letak perbaikan yang harus peneliti lakukan. Pada tahapan selanjutnya kembali diadakan FGD dengan pengajar. Rubrik yang dibahas merupakan rubrik yang telah diperbaiki setelah melalui tahapan uji coba bersama peserta. FGD kali ini diikuti oleh enam orang pengajar. FGD dilaksanakan pada tanggal 27 November 2020 bertempat di Pusat Pembinaan, Pendidikan, dan Pelatihan LIPI. Terdapat beberapa catatan yang peneliti dapatkan dari kegiatan FGD ini.

Pada FGD yang kedua, peneliti mengundang enam orang pengajar. Jumlah ini lebih banyak dibandingkan dengan peserta FGD sebelumnya yang hanya diikuti oleh dua orang pengajar. Pelibatan peserta FGD yang lebih banyak bertujuan untuk melihat tanggapan yang mungkin akan lebih beragam. Peneliti berasumsi bahwa semakin banyak yang terlibat, maka akan semakin beragam juga respons yang didapatkan. Asumsi peneliti ternyata benar. Pada FGD yang kedua terjadi kesalahpahaman ketika memaknai salah satu deskripsi pada dimensi penilaian topik penelitian. Ada salah satu pengajar yang salah dalam memaknai kalimat pada rubrik. Setelah membaca kalimat yang ada pada rubrik, pengajar tersebut menyangka bahwa peserta harus melakukan penelitian multidisiplin. Padahal yang dimaksudkan oleh rubrik bukanlah demikian. Penelitian yang dilakukan tidak harus penelitian multidisiplin, namun peneliti harus mampu melihat berbagai dampak permasalahan dengan menggunakan kacamata yang lebar seperti kacamata ekonomi dan sosial. Hal ini menjadi catatan bagi peneliti untuk memperbaiki dan meluruskan maksud kalimat tersebut.

Pada FGD yang kedua ini peneliti juga mendapatkan masukan terkait dimensi penilaian rumusan masalah. Peneliti mendapatkan masukan bahwa rubrik harus dapat menjelaskan pada penilai bahwa identifikasi dampak dan penyebab masalah boleh dilihat dari bagian latar belakang. Hal ini untuk meminimalisir kemungkinan kesalahan dalam menilai. Dikhawatirkan akan ada penilai yang menilai dimensi rumusan masalah hanya dengan melihat bagian rumusan masalah saja tanpa melihat bagian latar belakang. Masukan yang terakhir berkaitan dengan dimensi penilaian tujuan penelitian. Pada bagian ini peneliti mendapatkan masukan bahwa calon penilai kelak harus juga peka dengan latar belakang tempat bekerja penulis proposal. Penilai yang akan datang harus meyakinkan bahwa tujuan penelitian pada proposal penelitian telah sesuai dengan lembaga tempat peneliti tersebut bekerja.

Setelah mendapatkan masukan di atas selanjutnya peneliti kembali melakukan perbaikan rubrik penilaian. Setelah diperbaiki, rubrik kembali diperlihatkan pada peserta pelatihan. Peserta yang dimintai tanggapan berbeda dengan peserta pada pengujian rubrik sesi pertama. Data pada tabel 6 merupakan persepsi peserta terhadap rubrik penilaian proposal. Berdasarkan tabel tersebut kita dapat melihat perbedaan yang signifikan dibandingkan pada pengujian yang pertama (tabel 5). Persepsi negatif hanya ada pada dimensi penilaian manfaat penelitian saja. Hanya 0,11% peserta (1 orang) yang berpendapat bahwa rubrik tidak dapat membantu dalam merumuskan manfaat penelitian. Sementara sisanya beranggapan bahwa rubrik dapat membantu mereka dalam menyusun proposal penelitian. Selain itu, dapat dilihat juga bahwa seluruh peserta dapat memahami istilah-istilah yang ada pada rubrik, memahami seluruh dimensi penilaian, serta dapat memahami sistem penilaian yang ada pada rubrik. Setelah melalui serangkaian uji validitas sebagaimana dijelaskan di atas, peneliti dapat menyimpulkan bahwa rubrik dinyatakan valid. Validitas rubrik didapatkan setelah mendapatkan kesepakatan/konsensus dengan para pengajar. Validitas rubrik juga diperkuat dengan persepsi yang diberikan oleh peserta baik pada sesi pertama maupun kedua. Rubrik yang telah diperbaiki juga tidak mendapatkan penolakan dari pihak penyelenggara pelatihan.

Tabel 7 memperlihatkan proses uji reliabilitas. Kolom 1 memperlihatkan jumlah proposal yang telah dinilai yaitu 10 proposal penelitian. Sepuluh proposal penelitian kemudian dinilai oleh dua orang pengajar. Nilai yang diberikan untuk tiap proposal dapat dilihat pada kolom 2 dan 3. Kedua nilai kemudian diuji tingkat korelasinya. Setelah dihitung dengan menggunakan rumus *product moment*, didapatkan nilai korelasi sebesar 0,78. Dengan nilai tersebut maka tingkat reliabilitas rubrik dinyatakan cukup/dapat diterima (Arifin, 2011a).

Uji reliabilitas dilakukan dengan pendekatan *inter rater reliability* atau sering disebut IRR. IRR dilakukan dengan cara membandingkan nilai yang didapat dari dua orang atau lebih pengajar. Uji reliabilitas dengan pendekatan IRR umum digunakan untuk menguji reliabilitas rubrik. Pendekatan semacam ini dapat ditemukan juga pada penelitian rubrik lainnya.

### SIMPULAN

Latar belakang dari dilakukannya penelitian dan pengembangan ini adalah karena adanya perbedaan standar antar masing-masing penilai. Hal tersebut berakibat pada hasil penilaian yang subjektif dan tidak konsisten. Melalui penelitian ini, didapatkan delapan dimensi penilaian yang dapat dijadikan standar dalam menilai proposal penelitian. Delapan dimensi tersebut, meliputi pemilihan topik penelitian, *state of the art*, rumusan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, sumber literatur, dan pemilihan tata bahasa.

Delapan dimensi penilaian kemudian dikembangkan menjadi sebuah rubrik penilaian. Setelah dilakukan uji coba, rubrik dinyatakan valid dan dapat diterima serta dipahami dengan baik oleh pengajar, peserta, dan penyelenggara diklat. Rubrik juga dinyatakan reliabel setelah melalui tahapan uji statistik. Setelah dinyatakan valid dan reliabel, rubrik dapat diandalkan untuk menghasilkan nilai yang objektif serta konsisten walaupun dipakai oleh penilai yang berbeda-beda.

Berdasarkan penelitian yang telah dilakukan terdapat beberapa saran bagi berbagai pihak yang terkait. Bagi Pusbindiklat LIPI, agar dapat menggunakan rubrik ini sebagai alat untuk menilai kemampuan menulis proposal pada Program Pelatihan Pembentukan Jabatan Fungsional Peneliti. Penggunaan rubrik hendaknya diikuti dengan perbaikan dan penyempurnaan agar validitas dan reliabilitas rubrik semakin baik lagi. Selain sebagai instrumen penilaian rubrik juga dapat digunakan sebagai media pembelajaran. Peneliti menyarankan agar dilakukan penelitian yang bertujuan untuk menguji efektivitas rubrik dalam meningkatkan kemampuan dalam menulis proposal penelitian.

### UCAPAN TERIMA KASIH

Ucapan terima kasih kami sampaikan pada Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional yang telah membiayai penelitian ini melalui Program Beasiswa Sainstek Kemenristek/BRIN. Kami juga menyampaikan terima kasih kepada Pusat Pembinaan, Pendidikan, dan Pelatihan, Lembaga Ilmu Pengetahuan Indonesia dan seluruh pihak yang terlibat dalam penelitian ini

### DAFTAR RUJUKAN

- Aji, S. D., Hudha, M. N., Huda, C., Nandiyanto, A. B. D., & Abdullah, A. G. (2018). The Improvement of Learning Effectiveness in The Lesson Study by Using E-Rubric. *Journal of Engineering Science and Technology*, 13(5), 1181–1189.
- Arifin, Z. (2011a). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosda Karya.
- Arifin, Z. (2011b). *Penelitian Pendidikan: Metode dan Paradigma Baru*. Bandung: PT Remaja Rosda Karya.
- Aulet, T. H., Moore, J. S., Callas, P. W., Nicholas, C., & Hulme, M. (2018). (En) trust me : Validating an Assessment Rubric for Documenting Clinical Encounters During a Surgery Clerkship Clinical Skills Exam \*. *The American Journal of Surgery*, (xxxx), 2–6. <https://doi.org/10.1016/j.amjsurg.2018.12.055>
- Brookhart, S. M. (2013). *How to Create and Use Rubrics for Formative Assessment and Grading*. Alexandria: ASCD.
- Febriyanti, A. L., Harsiati, T., & Dermawan, T. (2017). Pengembangan Instrumen Asesemen Menulis Kreatif Cerita Fantasi untuk Siswa Kelas VII SMP. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 2(10), 1399–1408.
- Fitriani, S., & Yarmayani, A. (2018). Pengembangan Rubrik Berpikir Kreatif Siswa Menengah Atas dalam Menyelesaikan Masalah Matematika. *Jurnal Mosharafa*, 7(1), 33-38.
- Furze, J., Gale, J. R., Black, L., Cochran, T. M., & Jensen, G. M. (2015). Clinical Reasoning : Development of a Grading Rubric for Student Assessment. *Journal of Physical Therapy Education*, 29(3).
- Hermawan, Siahaan, P., Suhendi, E., Kaniawati, I., Samsudin, A., Setyadin, A. H., & Hidayat, S. R. (2017). Desain Rubrik Kemampuan Berkolaborasi Siswa SMP dalam Materi Pemantulan Cahaya. *JPPPF - Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3(2005), 167–174.
- Jönsson, A., & Panadero, E. (2017). The Use and Design of Rubrics to Support Assessment for Learning. In *Scaling up Assessment for Learning in Higher Education, The Enabling Power of Assessment 5* (pp. 99–111). Singapore: Springer Nature Singapore Pte Ltd. <https://doi.org/10.1007/978-981-10-3045-1>
- Latifa, A., Rahman, A., Hamra, A., Jabu, B., & Nur, R. (2015). Developing a Practical Rating Rubric of Speaking Test for University Students of English in Parepare, Indonesia. *English Language Teaching*, 8(6), 166–177. <https://doi.org/10.5539/elt.v8n6p166>
- Mckeown, J., & Biss, D. L. (2018). *HEQCO 's Guide to Developing Valid and Reliable Rubrics*. Toronto: Higher Education Quality Council of Ontario.
- Mislevy, R. J., & Knowles, T. (2002). *Performance Assessments for Adult Education*. Washington, DC: National Academy Press.

- O'Donnell, J. A., Oakley, M., Haney, S., O'Neill, P. N., & Taylor, D. (2011). Rubrics 101: A Primer for Rubric Development in Dental Education. *Journal of Dental Education*, 75(September), 1163–1175.
- Rahmawan, E. F., Sumaryanto, T., & Suproyadi. (2016). Pengembangan Instrumen Penilaian Kinerja Kemampuan Bernyanyi Berbasis Android. *Journal of Educational Research and Evaluation*, 5(1), 81–89.
- Rakedzon, T., & Tsabari, A. B. (2017). Assessing Writing to Make a Long Story Short : A Rubric for Assessing Graduate Students ' Academic and Popular Science Writing Skills. *Assessing Writing*, 32, 28–42.  
<https://doi.org/10.1016/j.asw.2016.12.004>
- Sweeney, N. L., Rollins, M. C., Gantt, L., Swanson, M., & Ravitz, J. (2019). Development and Reliability Testing of the Sweeney-Clark Simulation Evaluation Rubric Ó. *Clinical Simulation in Nursing*, 1–11.  
<https://doi.org/10.1016/j.ecns.2019.04.002>
- Tam, C. O. (2017). Evaluating Students ' Performance in Responding to Art : The Development and Validation of an Art Criticism Assessment Rubric. *The International Journal of Art & Design Education*, 0.0.  
<https://doi.org/10.1111/jade.12154>
- Zainul, A. (2001). *Alternative Assesment*. Jakarta: PAU-PPAI Universitas Terbuka.