Developing FD-MT to investigate students’ mental model on fluid dynamic concept: a Rasch model analysis

F H Dewi*, A Samsudin¹, and D T Chandra¹

¹Departemen Pendidikan Fisika, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudhi No. 229, Bandung 40154, Indonesia

* fannyherliyanadewi@upi.edu

Abstract. Fluid dynamics has some complex and unobservable concept, however, tests are rarely developed to measure students’ understanding of these concepts. This condition make a difficulty to map students’ mental models with existing instruments. Based on that problem, this study aims to develop FD-MT (fluid dynamic-multi tier test) as a diagnostic test for measure students’ mental model on fluid dynamic concept. The ADDIE (Analyzing, Designing, Developing, Implementing, Evaluating) is a research method that used in this study. The data were collected by FD-MT will be analyzed by rasch model analysis, include the reliability, the validity, item fit and the differential item function. Student mental model classified as Scientific, Synthesis-A (Sy-A), Synthesis-B(Sy-B), Synthesis-C (Sy-C), Synthesis-D(Sy-D), and Initial. The participant of this study are 20 students in the eleventh grade (9 boys and 11 girls, 16-17 age in average) at high school in Bandung. Based on result study shows that students’ mental model mostly in mental Sc (8.18%), Sy-A (3.18%), Sy-B (21.36%), Sy-C (19.09%), Sy-D (38.18%) and In (10%), and incompleted answer (Nr = 0%). This conclude that, FD-MT is able to measuring students’ mental model on fluid dynamic concept.

1. Introduction
Mental models represent perceptual reflections of reality and are closely related to one’s knowledge framework in describing, explaining and predicting something [1,2]. Mental models are important to learn because they can provide information about the structure of knowledge that students understand [1,2,3]. Students' mental models can’t be observed directly, one of the parameters used to analyze students' mental models is conceptual understanding [1,4,5,6,7]. The mental models analyzed with conceptual understanding parameters were classified into scientific, synthesis and initial [6,7]. In another study, the type of Synthesis mental model was divided into 4 types, namely the Sy-A, Sy-B, Sy-C and Sy-D [4,5].

![Figure 1. Classification and analysis of students' mental models](image-url)
Based on Figure 1, the classification of mental models in this study is Scientific, Synthesis-A, Synthesis-B, Synthesis-D dan Initial.

In understanding physics concepts, it is necessary to understand macroscopic and microscopic concepts [8,9]. Microscopic concepts are difficult to observe directly, this makes micro-concepts difficult to understand by students [8,9]. One of the microscopic concepts is dynamic fluid. Beside that, there are several difficulties on dynamic fluids concept, for example students are difficult to understand the abstract concept in ideal fluid, also can’t understand the meaning of the Bernoulli equation and its application [10]. The Bernoulli equation is stated:

\[ P_1 + \frac{1}{2} \rho v_1^2 + \frac{1}{2} gh_1 = C \]  

In physics research still very rare to analyze students' mental models on dynamic fluids concept. Various methods of instrument analysis are used both classical theory [4,5] and rasch model theory [6,7]. Fluid Dynamic-Multi-tier Test is a diagnostic test developed to analyze students' mental models based on student conceptual understanding. FD-MT is a change from FMCI which is transformed into a multi-tier test [10,11]. FD-MT is a multi-tiered multiple choice test consisting of 4 tiers which was developed to analyze students' mental models on the concept of dynamic fluids. So, in this study aim to develop FD-MT an instrument for analyze students' mental models on fluid dynamic that using rasch analysis.

2. Experimental Method

2.1. Design study

This research uses ADDIE research method. The ADDIE model is a research model that is systematically developed to solve learning problems according to the needs and characteristics of students [12,13]. ADDIE models consist of 5 stage, there are 1) Analyzing, 2) Designing, 3) Developing, 4) Implementation dan 5) Evaluation, that describe by the figure 2 [6,7,12].

![ADDIE method design](image)

**Figure 2.** ADDIE method design

2.2. Participant

The participant of this study are 20 students in the eleventh grade (9 boys and 11 girls, 16-17 age in average), located at public senior high school in Bandung city. Students became from a various backgrounds, whether domicily, religion, family background, or economic situation. But almost all students are domiciled in the city of Bandung.

2.3. Instrument

The instrument used in this study is FD-MT (Fluid Dynamic-Multitier Test) was conducted to analyze students’ mental model. FD-MT consist of 12 item number in the form of four-tier. Student answer all of the question for 35 minutes. The explanation about the instrumen showed on the figure 3. Students’ mental model then analyze by seen combination of their answer following Table [4,5].
Figure 3. The example question of Fluid-Dynamic Multitier Test

Table 1. The combination of MT-FD answer and classification of students’ mental model

| Tier | Sr  | Sy-A | Sy-B | Sy-C | Sy-D | In  | NR  |
|------|-----|------|------|------|------|-----|-----|
| 1    | S   | 1    | 1    | 1    | 1    | 1   |     |
| 2    | S   | S    | NS   | S    | S    | NS  | S   |
| 3    | 1   | 1    | 1    | 1    | 1    | 1   |     |
| 4    | S   | NS   | S    | NS   | S    | NS  | S   |

*1 : Correct Answer  0 : Wrong Answer  S : Sure  NS : Not Sure

2.4. Data analysis technique

The FD-MT instrument was tested on the sample to analyze the reliability, validity, item fit items and DIF using the Rasch model analysis. The reliability test of the FD-MT instrument was analyzed by item and person reliability. The reliability value $r \geq 0.8$ has very good interpretation, $0.7 \leq r < 0$ has a good interpretation, $0.6 \leq r < 0.7$ has enough interpretation, $0.5 \leq r < 0.6$ has a bad interpretation and for $r < 0.5$ has very bad interpretation. The construct validation test using Rasch model analysis is called item undimensionality, which is seen from the raw variance value explained by measures. This value has the interpretation fulfilled ($r > 20\%$), good ($r > 40\%$), and special ($r > 60\%$).

Meanwhile, item fit order analysis shows the quality of the item (the suitability of the item measures what it should be measured) and the level of difficulty. The item fit criteria for the instrument was accepted when it had MNSQ value of $0.5 < \chi < 1.5$; ZSTD $-2 < \chi <+2$; and PT Measure Corr $0.4 < \chi < 0.85$. If these three criteria are met, it can be said that the item is "very suitable". It can be ascertained that the quality of the item is good for use. It is said "suitable" if one or two criteria are met then the item can still be maintained and does not need to be changed, whereas "not suitable". If all three criteria are not met then it is certain that the item is not good so they need to be repaired or replaced. To analyze the bias on the test item, it is analyzed through a $s$ (DIF). The question item is said to have a bias when it has an item probability value below 5% [6,14].

Based on the classification of the mental model in students’ answers, the students' mental model for each concept is expressed in percentage by equation 2. This percentage is used to analyze the level distribution of students’ mental models on each concept and as a whole on the dynamic fluid concept.
3. Result and Discussion

3.1. Analyzing
In the first stage is analysis, researchers conducted a literature study on the development of instruments used to analyze students’ mental models [4,5,6], classification of mental models [4,5,6,7,15] and several analytical methods that could be used to evaluate test instruments [6,14,16]. The analysis stage is continued until analyze competencies and students' conception on fluid dynamic concept by literature study to be developed into blueprint items.

3.2. Designing
At the design stage, the researcher wrote the blueprint till the multiple choices items that were adapted from FMCI. The multiple-choice questions initially had a 2-tier format (reasonable multiple choice) to collect students’ reasons as alternative choice materials on tier-3 in FD-MT. The following is a schematic of the FD-MT question design which was built from 2-tier questions to 4-tier format.

3.3. Developing
On developing stage, researcher develop the third tier based on the answer of students’ reason on the open ended item of fluid dynamics that implemented on the design stage. FD-MT contains of 11 items about fluid dynamic and its’ sub materials that learned on high school. In other words, this stage is to write down the tier-3 from the data of two tier test and finally FD-MT was conducted as shown on the Figure 3.

3.4. Implementation
The implementation, is a stages that distributing the FD-MT to 20 students. Student can access the item test by PC or smartphone on online form. The item test can be accessed by student for 35 minutes and their answer will collect automatically to the answer bank.

3.5. Evaluation
The students’ mental models analyzed consisted of 6 categories. Each category is given a different score adapting from several studies [6,7,17]. Students with an intact mental model (Sc) have a score of 4. Whereas the partial level is separated into 3 scores, specifically 3 for SY-A since this mental model has a complete conception, but isn’t went with by a total level of confidence either. While SY-B and SY-C which were given a score of 2 since it appeared an incomplete mental model whereas SY-D was given a score of 1 because students' mental models had alternative conceptions (misconceptions). After scoring

\[
\text{Students' mental model level (\%)} = \frac{\text{Number of students at one level}}{\text{Total number of all students}}
\]
the mental models of students, the data obtained were analyzed using the Rasch model. The Rasch model test was carried out with the help of Winsteps software form 4.8.0.0.

Reliability results show the value of the item reliability 0.90 so that it has a very good interpretation for the test item. Meanwhile, the score for person reliability shows a low score of 0.49. This condition shows that there is no seriousness of students in doing the test items so that the consistency value of students in answering questions is still low. The figure 5 below appears the comes about of the variable maps of items and persons. Outcome item showed on the right side, while he outline person appeared on the left side. Based on the outline showed that participant almost has a medium ability and some has a low ability on mental model. The value of item reliability shows that the test item used is reliable to measure students’ mental models.

![Figure 5. Item-Person Map of Rasch Analysis](image)

The validity results indicated by the value Raw variance explained by measures showed 50.4% results. These results indicate that the FD-MT is valid and has a good interpretation. To see the suitability of the test item, it is shown by the fit order analysis. The results of the FD-MT fit order analysis are shown in the following table.

| Concept                  | No. Item | MNSQ | ZETD | PT. Measure Corr. | Prob     |
|--------------------------|----------|------|------|-------------------|----------|
| Fluid Dynamic            | 1        | 0.97 | 0.74 | 0.25              | Suitable |
| Fluid Ideal              | 2        | 0.82 | -0.63| 0.77              | Suitable |
|                           | 3        | 1.70 | 1.59 | 0.58              | Suitable |
| Continuity Azas          | 4        | 0.38 | -2   | 0.39              | Not suitable |
|                           | 5        | 0.83 | -0.34| 0.44              | Suitable |
|                           | 6        | 0.56 | -1.55| 0.25              | Suitable |
| Bernoullis’ Principe     | 7        | 0.78 | -0.62| 0.33              | Suitable |
|                           | 8        | 0.93 | -0.08| 0.23              | Suitable |
|                           | 9        | 1.10 | 0.45 | 0.73              | Suitable |
| Bernoullis’ Principe Aplication | 10 | 1.05 | 0.25 | -0.30             | Suitable |
|                           | 11       | 1.43 | 1.44 | 0.26              | Suitable |

The differential item function (DIF) is performed to analyze the bias that occurs in the test item due to certain grouping factors. DIF analysis can be done through gender, age, language differences, and academic background or content knowledge etc. [14]. On this research DIF analysis focuses on the performance of two or more different groups in this study analyzed by the gender group (male and female). The DIF results in knowing which items can be shown in Table 2. DIF analysis appears that, items T2 and T11 have a probability value underneath 5%. So that these two items can cause bias within the gather examined. Based on this value, there are two test items that have DIF. To analyze DIF more
profoundly, DIF Person detection curve examination can moreover be utilized to analyze things particularly for the level of difficulty of the issue with groups of female and male students, showed in Figure 6 bellow.

![PERSON DIF plot (DIF=$S4W1)](image)

**Figure 6.** The result of DIF detection curve

Based on the figure 6 appears that on item test T2 female students contains a higher probability to answer the test compared to male students, whereas on item T11 male students contains a higher probability to answer the test. Based on that curve can conclude that on these items contain DIF since there's a critical contrast between the probability values of male and female students. So that in these items there's one group that benefits more, causing bias within the results. In expansion to the investigation of the instruments utilized, the following analysis of students' mental models on the dynamic fluid concept is appeared in table 3.

| Sub-Content                          | Level Understanding | SC  | SY-A | SY-B | SY-C | SY-D | In  |
|-------------------------------------|---------------------|-----|------|------|------|------|-----|
| Fluid dynamic                       |                     | 10  | 0    | 5    | 60   | 25   | 0   |
| Ideal Fluid                         |                     | 22.5| 12.5 | 35   | 5    | 12.5 | 12.5|
| Debit                               |                     | 0   | 0    | 50   | 25   | 25   | 0   |
| Continuity equation                 |                     | 0   | 0    | 11.7 | 25   | 55   | 8.3 |
| Bernoulli’s Principle               |                     | 17.5| 2.5  | 35   | 2.5  | 30   | 12.5|
| Bernoulli’s Principle and It’s aplication |                 | 3.3 | 5    | 8.3  | 11.7 | 56.7 | 15  |
| Summaries                           |                     | 8.18| 3.18 | 21.36| 19.09| 38.18| 10.00|

Based on the data obtained shows that most students have partial mental model, specially on debit and fluid dynamic concept. Debit, which is a concept that is often encountered in everyday life, is a familiar term for students. However, there is a bias in students' understanding of the concept such another study before [10], some students can predict what will happen or the impact of a phenomenon, but when it is related to scientific explanations, students cannot explain it correctly, this is indicated by the highest percentage of Sy-B. Some of student also have scientific model mental and the largest one is on the ideal fluid concept. This shows that students can understand concepts scientifically and are able to explain and predict what the consequences of these concepts will be so that students have a complete mental model.

But on the continuity equation and Bernoulli’s Principe application the most student have initial and sy-d mental models. Means that almost students have alternative conception on these concept. Based on the same research showed that students still do not understand the relationship between variables in these concepts [10]. In this study most students still have alternative conceptions on the relationship between the variable cross-sectional area (A) and the fluid flow velocity (v), the v relationship with the pressure on the pipe wall (container) (P) and the causal relationship between the velocity of the flowing fluid and the pressure it causes. on the wall or container. In accordance with the distribution of data in
Figure 4, it shows that the participants have a low mental model on the Bernoulli principle concept and its application. This is supported by the percentage of mental models with sy-D and In types that dominate this concept. Whereas in Figure 4 shows several test items that cannot be answered correctly by students are test items with the concept of the Bernoulli equation and its application (T7, T8, T10, T11). Based on the analysis obtained on the feasibility of the FD-MT instrument as a mental model diagnostic test, it shows that the test item is feasible to use, marked by a test of item reliability, item validity, item fit order and DIF which as a whole shows good interpretation results. There are several test items that need to be improved for implementation in further research. While the analysis of students' mental models shows that overall students have the distribution of mental models of Sc (8.18%), Sy-A (3.18%), Sy-B (21.36%), Sy-C (19.09%), Sy-D (38.18%) and in (10%), and incompleted answer (NR = 0%).

4. Conclusion
FD-MT could be a diagnostic test created by researchers to analyze students' mental model levels on dynamic fluid concept. Based on the information gotten from the research most students have a synthesis B and C level of each concept, means that almost student have a partial mental model about some phenomena also have difficulty when clarifying a situation or phenomenon by combining two or more concepts well and experimentally. While the FD-MT also analyze by Rasch model analysis. Based on the result appeared that almost all the instrument item featured a great reliability, validity, and item fit, however there are some items that consist of some bias and need to develop more. So, it can be concluded that FD-MT can be utilized to analyze student metal model on fluid dynamic concept.

5. Acknowledgments
The authors would like to thank all those who helped in completing this investigate, the Ministry of Higher Education, Research and Technology, the republic of Indonesia, also particularly to our research tim and staff of the physics education department lecturer, who continuously given valuable bolster and talk in this research. As consequence, researchers are able to publish in the journal.

6. References
[1] Didiş N, Eryllmaz A and Erkoç S 2014 Phys. Rev. Spec. Top. - Phys. Educ. Res. 10 2
[2] Özcan S B Ö 2016 J. Balt. Sci. Educ. 15 5
[3] Jalmo T and Suwandi T 2018 J. Balt. Sci. Educ. 17 3
[4] Dewi R, Samsudin A, Nugraha M G and Liliawati W 2019 J. Phys.: Conf. Ser. 1280 052029
[5] Kafiyani F, Samsudin A and Saepuzaman D 2019 J. J. Phys.: Conf. Ser. 1280 052030
[6] Nurdini N, Suhandi A, Ramalis T R and Samsudin A 2020 Artic. J. Adv. Res. Dyn. Control Syst. 12 6
[7] Fratitiwi N J, Samsudin A, Ramalis T R and Costu B 2020 Univers. J. Educ. Res. 8 6
[8] Suhandi A, Hermita N, Samsudin A, Maftuh B and Costu B 2017 Turkish Online J. Educ. Technol. pp. 1012–1022
[9] Wibowo F C, Suhandi A, Samsudin A, Darman D R, Suherli Z, Hasani A, Leksono S M, Hendrayana A, Hidayat S, Hamdani D and Costu B 2017 Asia-Pacific Forum Sci. Learn. Teach. 18 2
[10] Dewi F H, Samsudin A and Nugraha M G 2019 J. Phys.: Conf. Ser. 1280 052037
[11] J. Martin, J. Mitchell, and T. Newell 2003 Proc. - Front. Educ. Conf. FIE 1 23
[12] A. K. N. Hess and K. Greer 2016 Commun. Inf. Lit. 10 2
[13] Rahmayani Rahmatsyah and Sani R A 2021 J. Phys.: Conf. Ser. 1811 012069
[14] Karami H 2012 Int. J. Educational Psychol. Assess. 11 2
[15] M. Altan Kurnaz and C. Eksi 2015 Educ. Sci. Theory Pract. 15 3
[16] Huang L, Huang F, Oon P T and Mak M C K 2019 Eurasia J. Math. Sci. Technol. Educ. 15 12
[17] Jiang T, Wang S. Wang and Ma Y 2018 Eurasia J. Math. Sci. Technol. Educ. 14 7