Developing good tests and test items in higher education is crucial. Tests determine to a large extent what students learn and how they perceive the domain of study. Especially in first and second year introductory courses, selected response test items are used to keep the assessment and grading task at a manageable level. It is therefore vital for higher education that teachers be able to generate selected response test items that are appropriate for the full range of goals for which test items are needed. Teachers should be able to design test items not only for rote learning purposes but also and more importantly for querying more cognitive demanding goals such as application of knowledge, critical thinking, and problem solving.

However, teachers often develop test items without specific training and with very limited time; they are often forced to engage in this activity alone. The existing literature and support regarding the design of test items is experienced by higher education teachers as unhelpful and offering them little or no useful assistance in their daily practice. This situation is problematic and comes at the expense of the quality of assessment in higher education. Therefore, in this thesis, new ways to improve devising test items by teachers in higher education were explored.

The aim of this thesis was twofold. First, because little is known about how teachers go about designing test items, the objective was to understand and describe how exactly how teachers generate test items. Second, based on this better understanding, support for teachers that leads to better test items was developed and studied in either a practical or an experimental setting. This thesis contains six studies that address these central research questions in different ways.

The results of the studies show that it is productive to consider designing test items as solving an ill-defined design problem. It offers a fundamentally different view of the problem of test item design, leading to new opportunities to support teachers in the test item design task. Specific support resources and interventions were developed. It was found that resources and interventions can each be effective, especially for improving the divergent production of ideas for test items.